Norwegian University of Life Sciences
Faculty of Veterinary Medicine
Department of Preclinical Sciences and Pathology

Philosophiae Doctor (PhD)
Thesis 2020:19

Genetics of chronic subclinical mastitis in Norwegian Red cows

Genetisk bakgrunn for kronisk subklinisk mastitt hos Norsk Rødt Fe

Elena Kirsanova

Genetics of chronic subclinical mastitis in Norwegian Red cows

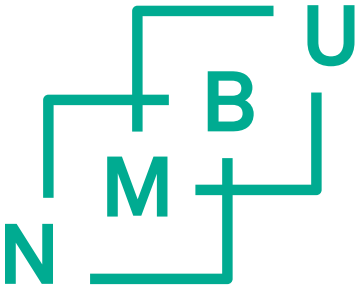
Genetisk bakgrunn for kronisk subklinisk mastitt hos Norsk Rødt Fe

Philosophiae Doctor (PhD) Thesis

Elena Kirsanova

Norwegian University of Life Sciences
Faculty of Veterinary Medicine
Department of Preclinical Sciences and Pathology

Adamstuen (2020)



Thesis number 2020:19
ISSN 1894-6402
ISBN 978-82-575-1682-6

CONTENTS

ACKNOWLEDGEMENT

The work presented here has been a part of the project ”A multidisciplinary approach to the control of pathogen-specific mastitis”. The project was coordinated by the Norwegian University of Life Sciences (NMBU) and was a collaboration between NMBU, Geno SA and TINE SA. The project was funded by the Research Council of Norway project no. 233778.

I am grateful to my supervisors Ingrid Olsaker, Bjørg Heringstad and Anna Lewandowska-Sabat for your guidance, discussions and good humour. Furthermore, I would like to thank employees in Geno SA for making the NDHRS data available to me and all colleagues at the Department of Preclinical Sciences and Pathology and Department of Animal and Aquacultural Sciences at NMBU, for all help, advices and scientific discussions. Finally, my heartfelt thanks go to my family.

Oslo, 2020

Elena Kirsanova

SUMMARY

Chronic subclinical mastitis is an udder inflammation widely defined as milk somatic cell count (SCC) above 200 000 cells/ml for a prolonged period of time, often caused by bacterial infection. Several authors have studied subclinical mastitis (SCM) with the purpose to identify underlying genetic predispositions, while in the case of chronic SCM to date only a few studies exist. Therefore, the genetic mechanisms associated with chronic SCM still remains to be identified. The present thesis focus on chronic SCM in the Norwegian Red (NR) population, aiming to identify underlying biological processes contributing to high SCC in milk.

Firstly (Paper I), the available SCC phenotype records on NR cows in lactation 1 to 3 were used to estimate genetic parameters of novel defined chronic SCM traits and for lactation average somatic cell score (LSCS). Ten SCM traits were defined as binary and scored as 1 if two testday SCC in a row were above SCC thresholds from 50 000 to 400 000 cells/ml (SCM50, -100, -150, -200, -250, -300, -350, -400); and if three testday SCC in a row were above 200 000 and 400 000 cells/ml, respectively (SCM200_3, SCM400_3). Also the number of days before the first case with SCM50=1 (D50) or SCM400=1 (D400) were analysed. The estimated heritability of the SCM traits varied from 0.01 to 0.10 and for LSCS the heritability was 0.3. High genetic correlations were estimated between the alternative SCM traits, ranging from 0.7 to 1.

The estimated genetic parameters from Paper I were used as input in a genome-wide association study (GWAS) together with genotype data from 3795 NR bulls (Paper II). The analyses were based on single nucleotide polymorphism (SNP) data from NR bulls and SCC records from their daughters (n=3 543 764). The GWAS detected 36 significant SNPs associated with one or more of the alternative SCM traits. Based on converted bovine topologically associated domain (TAD) regions with significant SNPs, 181 genes were identified and analysed by Ingenuity

Pathway Analysis (IPA) in order to recognise involved canonical pathways and networks. Several significant pathways and putative candidate genes for the alternative SCM traits and for LSCS were identified, which illustrates the complexity of chronic SCM. A group of chemokine (C-X-C motif) ligand genes (*CXCL1, CXCL2, CXCL3, CXCL5, CXCL8*) were found associated to SCM200_3, in addition to the Fos proto-oncogene, AP-1 transcription factor subunit (*FOS*), that was also found as a top significant gene for the traits SCM150, -200, -250, -300, -350, -400 and D400.

Further, expression profiles of selected candidate genes identified in association with the chronic SCM traits, were analysed in samples of unstimulated peripheral blood mononuclear cells (PBMC) from twenty NR cows; ten cows with high and ten cows with low estimated genomic breeding value (GEBV) for LSCS (Paper III). Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) on isolated RNA revealed differences between the NR cows. Significantly higher expression level of the *CXCL1* gene was found in unstimulated PBMC samples from the group of cows with high GEBV for LSCS compared to the group of cows with low GEBV. Higher expression of the *CXCL1* gene may contribute to more rapid immune response, calling for more neutrophil cells when an invading pathogen enters the udder.

In Paper IV, two groups of NR cows with high (n=6) or low (n=6) GEBV for LSCS were studied. Macrophages isolated from blood samples were challenged *in vitro* with the *Streptococcus agalactiae* strains ST12 or ST10. Challenge with the bacterial strains revealed 5936 and 6443 differentially expressed (DE) genes in the groups infected with ST103 or ST12, respectively, versus unchallenged controls. The results observed in Paper IV, indicates involvement of neutrophils in high SCC in milk during inflammation in the udder. Such as, the observed up-regulation of *Myosin IF* (*MYO1F*) in the ST103 infected macrophages from cows with low GEBV for LSCS, imply increased neutrophil migration in to the infected tissue, which may cause high SCC in milk.

In summary, our study have contributed for better understanding of biological mechanisms underlying the development of chronic SCM in NR cattle. The defined alternative SCM traits are genetically different from the LSCS trait currently used in genetic evaluations for NR and can therefore add information in order to improve breeding for better udder heath. Moreover, the observed cellular expression profiles of genes involved in neutrophil influx indicates regulation of neutrophil movement as an important part of the biological mechanisms behind chronic SCM traits in NR and the number of somatic cells in milk.

NORSK SAMMENDRAG

Kronisk subklinisk mastitt defineres vanligvis som jurbetennelse med høyt celletall i melk (over 200 000 celler/ml) i en lengre periode, ofte forårsaket av en bakterieinfeksjon. Flere har analysert subklinisk mastitt (SCM) og brukt alternative definisjoner av egenskapen for å finne den genetiske bakgrunnen. Men for kronisk SCM finnes det per dags dato kun noen få studier der alternative definisjoner av egenskapen har blitt analysert. Kunnskapen om de genetiske mekanismene som styrer kronisk SCM er derfor fortsatt mangelfull. I denne studien har vi fokusert på kronisk SCM i Norsk Rødt Fe (NRF) populasjonen med mål om å identifisere grunnleggende biologiske prosesser som bidrar til høyt celletall i melk.

I artikkel I, ble data for celletall (SCC) hos NRF kyr i laktasjon 1 til 3 brukt til å estimere genetiske parametere for 12 nye SCM egenskaper og for gjennomsnitt laktasjonscelletall (LSCS). Ti binære SCM egenskaper ble definert som 1 hvis SCC i melk to testdager på rad var høyere enn en gitt terskelverdi fra 50 000 til 400 000 celler/ml (SCM50, -100, -150, -200, -250, -300, -350, -400); og hvis SCC tre testdager på rad var over 200 000 og 400 000 celler/ml, henholdsvis (SCM200_3; SCM400_3). I tillegg ble antall dager før SCM50 (D50) eller SCM400 (D400) oppstår for første gang analysert. Den estimerte arvegraden for SCM egenskapene varierte fra 0,01 til 0,10 og for LSCS var den 0,30. Høye genetiske korrelasjoner, med variasjon fra 0,7 til 1, ble estimert mellom de 12 alternative SCM egenskapene.

De estimerte genetiske parameterne ble videre brukt i artikkel II i en assosiasjonsstudie (GWAS) med genotypedata i form av enkeltbasepolymorfier (SNPer) for NRF okser (n=3795) og SCC fenotypedata for døtrene til oksene (n=3 543 764). Ved GWAS analysen ble 36 SNP–markører identifisert som signifikant assosiert med en eller flere av de alternative SCM egenskapene. Basert på konverterte topologiske områder (TAD) med signifikante SNPer i storfegenomet, ble 181 gener identifisert og analysert med Ingenuity Pathway Analysis (IPA)

for å finne involverte biologiske prosesser og genetiske nettverk. Flere påviste likheter og forskjeller mellom de alternative SCM egenskapene illustrerer kompleksiteten til kronisk SCM. En gruppe av kjemokin (C-X-C motiv) ligand gener (*CXCL1, CXCL2, CXCL3, CXCL5, CXCL8*) ble identifisert for SCM200_3 i tillegg til Fos proto-onkogen, AP-1 transkripsjonsfaktor subenhet (*FOS*). *FOS* genet ble også funnet som ett av de mest signifikante genene for SCM150, -200, -250, -300, -350, -400 og D400.

Uttrykket av selekterte kandidatgener identifisert i forbindelse med de kroniske SCM egenskapene ble videre studert i ustimulerte perifere mononukleære blodceller (PBMC) fra ti NRF kyr med høy genomisk avlsverdi (GEBV) for LSCS og ti kyr med lav GEBV for LSCS (artikkel III). Revers transkripsjon-kvantitativ polymerasekjedereaksjon (RT-qPCR) ble benyttet på RNA isolert fra de ustimulerte PBMC prøvene og forskjeller i genuttrykk ble påvist. Signifikant høyere ekspresjon av genet *CXCL1* ble påvist i prøver fra kyr med høy GEBV sammenlignet med gruppen med lav GEBV. Høyere ekspresjonen av *CXCL1* kan være med på å forklare bedre motstand mot kronisk SCM gjennom en raskere immunrespons ved å tilkalle flere nøytrofile granulocytter for å angripe invaderende patogener i juret.

I artikkel IV ble grupper av NRF kyr med høy (n=6) eller lav (n=6) GEBV for LSCS studert. Makrofagceller isolert fra blodprøver ble i kultur (*in vitro*) infisert med *Streptococcus agalactiae* stammene ST12 eller ST103 i tillegg til ustimulerte kontrollprøver. Det ble funnet totalt 5936 og 6443 gener med signifikant forskjellig uttrykk i gruppene infisert med henholdsvis ST103 og ST12 sammenlignet med ustimulerte kontrollprøver. Resultatene i artikkel IV indikerer at nøyrofile granulocytter er involvert i økt SCC i melk ved en inflammasjon i juret. For eksempel opp-regulering av *Myosin IF* (*MYO1F*) genet i ST103 infiserte makrofager fra kyr med lav GEBV for LSCS antyder øket tilstrømning av nøytrofile granulocytter til infisert vev og kan dermed bidra til høyere SCC i melk.

Kort oppsummert har denne avhandlingen bidratt til en bedre forståelse av de biologiske mekanismene som ligger til grunn for utviklingen av kronisk SCM hos NRF. Nye alternativt definerte SCM egenskaper er genetisk forskjellige fra LSCS egenskapen som brukes ved beregning av avlsverdier for NRF per i dag. De nye egenskapene vil dermed kunne bidra med mer informasjon som kan forbedre avl for bedre jurhelse. I tillegg indikerer de observerte cellulære ekspresjonsprofilene av gener involvert i migrasjon av nøytrofile granulocytter at regulering av nøytrofil vandring fremstår som en viktig del av de biologiske mekanismene bak kronisk SCM i NRF og antall av somatiske celler i melk.

ABBREVIATIONS

bMDMs	bovine Monocyte Derived Macrophages
BTA	*Bos Taurus autosome*
cDNA	complementary Deoxyribonucleic Acid
CM	Clinical Mastitis
Ct	Cycle threshold
DE	Differential Expression
dsDNA	double strand Deoxyribonucleic Acid
DYD	Daughter Yield Deviation
EBV	Estimated Breeding Value
FDR	False Discovery Rate
GEBV	Genomic Estimated Breeding Value
GS	Genomic Selection
GWAS	Genome-Wide Association Study
IL	Interleukin
IPA	Ingenuity Pathway Analysis
LD	Linkage Disequilibrium
lncRNA	long non coding Ribonucleic Acid
LSCS	Lactation mean Somatic Cell Score
miRNA	micro Ribonucleic Acid
mRNA	messenger Ribonucleic Acid
NDHRS	Norwegian Dairy Herd Recording System
NR	Norwegian Red
PBMC	Peripheral Blood Mononuclear Cell

QTL	Quantitative Trait Loci
rRNA	ribosomal Ribonucleic Acid
RT-qPCR	Real Time quantitative Polymerase Chain Reaction
SCC	Somatic Cell Count
SCM	Subclinical Mastitis
SCS	Somatic Cell Score
SD	Standard Deviation
SE	Standard Error
SNP	Single Nucleotide Polymorphism
TAD	Topologically associated domain
tRNA	transfer Ribonucleic Acid

The list of the official full names to bovine gene name abbreviations (*Bos Taurus*; https://www.ncbi.nlm.nih.gov/)

ACOT2 – acyl-CoA thioesterase 2

ACOT4 – acyl-CoA thioesterase 4

ATP5B – ATP synthase F1 subunit beta

C6 – complement C6

CCNB1 – cyclin B1

CDK7 – cyclin-dependent kinase 7

CXCL1, also known as GRO1 – chemokine (C-X-C motif) ligand 1

CXCL2 – chemokine (C-X-C motif) ligand 2

CXCL3 – chemokine (C-X-C motif) ligand 3

CXCL5 – chemokine (C-X-C motif) ligand 5

CXCL8, also known as IL8 – C-X-C motif chemokine ligand 8

EIF2B2 – eukaryotic translation initiation factor 2B subunit beta

FOS – Fos proto-oncogene, AP-1 transcription factor subunit

GAPDH – glyceraldehyde-3-phosphate dehydrogenase

GLI2 – GLI family zinc finger 2

GRO1 – chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity alpha)

MARCO – macrophage receptor with collagenous structure

MYO1F – myosin IF

PADI4 – peptidyl arginine deiminase 4

PLPP1 – phospholipid phosphatase 1

PPIA – peptidylprolyl isomerase A

RAD17 – checkpoint clamp loader component

RPL12 – ribosomal protein L12

SDHA – succinate dehydrogenase complex flavoprotein subunit A

SEL1L – ERAD E3 ligase adaptor subunit

SLC18A2 – solute carrier family 18 member A2

STAT4 – signal transducer and activator of transcription 4

TGFB3 – transforming growth factor beta 3

LIST OF PAPERS

Paper I

Alternative Subclinical Mastitis Traits for Genetic Evaluation in Dairy Cattle

E. Kirsanova, B. Heringstad, A. Lewandowska-Sabat, I. Olsaker.

Journal of Dairy Science, 2019; **102(6):5323-5329** doi: 10.3168/jds.2018-16104

Paper II

Identification of Candidate Genes Affecting Chronic Subclinical Mastitis in Norwegian Red cattle: Combining Genome Wide Association Study, Topologically Associated Domains and Pathway Enrichment Analysis

E. Kirsanova, B. Heringstad, A. Lewandowska-Sabat, I. Olsaker.

Animal Genetics, 2020; **51(1):22-31** doi:10.1111/age.12886

Paper III

Expression Analysis on Candidate Genes for Chronic Subclinical Mastitis in Norwegian Red

E. Kirsanova, P. Boysen, G. M. Johansen, B. Heringstad, A. Lewandowska-Sabat, I. Olsaker

Submitted to Journal of Dairy Science, 7 January 2020

Paper IV

Transcriptional Profiling of Monocyte-Derived Macrophages Infected *in vitro* with Two Strains of *Streptococcus agalactiae* Reveals Candidate Pathways Affecting Subclinical Mastitis in Cattle

A. Lewandowska-Sabat, E. Kirsanova, C. Klopp, T. Solberg, B. Heringstad, O. Østerås, P. Boysen, I. Olsaker.

Frontiers in Genetics, 2019, **10:689** doi: 10.3389/fgene.2019.00689

INTRODUCTION

Background

Chronic subclinical mastitis (SCM) in dairy cattle is widely defined as somatic cell count (SCC) above 200 000 cells/ml in milk for a prolonged period of time (*e.g.* Harmon 1994; Schepers et al. 1997; Pantoja et al. 2009). High SCC can result in reduced milk production and consequently economic loss (Hagnestam-Nielsen et al. 2009). Previously, SCC and clinical mastitis (CM) have been studied in Norwegian Red (NR) cattle in analyses of Quantitative traits loci (QTL), for better understanding of the underlying mechanisms (Klungland et al. 2001; Sodeland et al. 2011; Olsen et al. 2016). Nevertheless, the threshold SCC of chronic SCM has not been defined precisely neither in Norway nor in other countries. Hence, better understanding of the genetic and functional background is needed in order to identify the biological process underlying chronic SCM in dairy cattle and further define this trait by using SCC as indicator. The Norwegian Dairy Herd Recording System (NDHRS) provides records on milk and meat production, fertility and veterinary treatments for CM and other diseases. Phenotypic data on SCC have been routinely collected in the form of testday SCC records since 1978. These data represent a large database available for analyses with solid historical background and a huge amount of records. By using the NDHRS database in combination with genotype and pedigree information on the NR population, novel traits can be identified and thoroughly analysed with the purpose to contribute additional information and improve the genetic evaluation of udder health in NR in order to reduce chronic SCM cases as well as SCC in the population.

The breeding program for Norwegian Red

Geno Breeding and AI Association is responsible for development of the breeding program for the NR population (www.geno.no). CM has been included in the breeding goal for NR since 1978 (Heringstad et al. 1999). While SCC became included in the genetic evaluation of udder health in 2014 as lactation average somatic cell score (LSCS). This trait is derived from testday SCC records as a $\log_e$ transformed somatic cell score (SCS; Schukken et al. 1992). Genomic selection (GS; Meuwissen et al. 2001; see below) is used for breeding of NR. GS is widely used in many dairy cattle breeding schemes (Loberg & Dürr 2009). This selection method is based on genomic estimated breeding values (GEBV). To calculate GEBV in NR a single step genomic best linear unbiased prediction (ssGBLUP) model is used by Geno, where both pedigree (**A**) and single nucleotide polymorphism (SNP) based (**G**) relationship matrices are taken into account, by using a combined **H** matrix (VanRaden 2008; www.geno.no). Relationship estimated based on the SNP data, also called genomic relationship matrix **G,** is based on the realized proportion of the genome that two individuals share (Goddard et al. 2011). The traditionally estimated matrix **A** is based on the expected proportion of the genome that two individuals have in common. Combining **A** and **G** give higher accuracy on the relationship between individuals. The total merit index for NR consists of the following trait groups with relative weights given as percentage in parenthesis (www.norwegianred.com): Production (23.54), udder confirmation (22.56), udder health (15.66), daughter fertility (13.87), beef (7.16), claw health (4.02), foot and leg confirmation (2.18), calving ease (5.40), stillbirth (0.55), milking speed (3.76), leakage (0.53), other diseases resistance (0.35) and temperament (0.42). The udder health index for NR consist of SCC (68 %) and CM (32 %). Information on SCC is recorded through laboratory analysis of testday milk samples, while CM is recorded as veterinary treatment and reported to the NDHRS.

Mastitis in dairy cattle

Mastitis is one of the most frequent and costly diseases in dairy cattle, often characterized by increased SCC (Harmon 1994). CM is described as an inflammation of the udder with clinical symptoms, such as abnormal milk and swelling or pain in the udder. While SCM is present with no visible changes, but milk production decrease, SCC increase and often bacteria are present in the milk secretion. There are several mastitis pathogens that cause udder inflammation *e.g.* *Staphylococcus aureus* and *Streptococcus agalactiae* and coliforms (Smith 2014). In Norway, *S. aureus* is the most common bacteria causing bovine mastitis, while *S. agalactiae* is among the leading causative bacteria of subclinical bovine mastitis. Both bacteria are found in the environment (Zadoks et al. 2002; Jørgensen et al. 2016) and after milking, since the teat canal still remains dilated, they may easily invade the udder and attack tissues lining the milk-collecting ducts and cisterns (Figure 1; Viguier et al. 2009; Bogni et al. 2011). Production of several virulence factors cause swelling and death of milk-secreting cells. By bacterial invasion the leukocytes in the udder signals additional leukocytes to move from the blood into the udder quarters and teat canals to destroy the bacterial infection. However, if the invading bacteria are not completely removed, they continue to attack ducts and alveolar areas. In some cases, infection is cleared and udders return to normal but if the infection persists milk secretory tissue may be destroyed or alternatively the infection become chronical. When an innate immune response is activated, the number of leukocytes together with macrophages and epithelial cells increase in the mammary glands and define the threshold of SCC in milk (Miller et al. 1991; Schwarz et al. 2011).

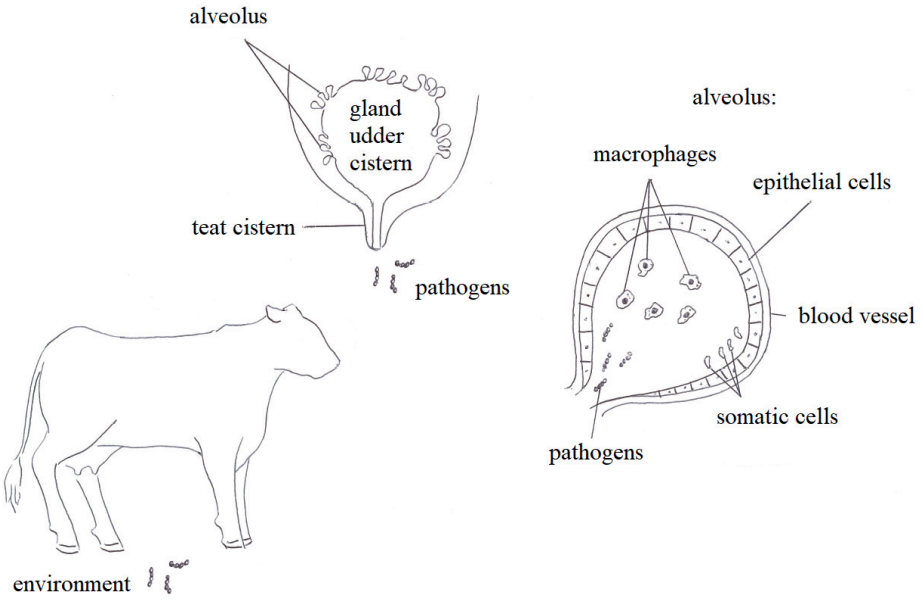


Figure 1. Mastitis infection. Schematic physiological structure of the udder quarter in dairy cattle is presented. On the right picture invading pathogens have entered the udder quarter and alveolus through an open teat canal, causing damage of epithelial cells and leading to atrophy of alveolus. The pathogens recognized by immune cells trigger activation of an innate immune response, increase in somatic cell count and may lead to the inflammation process.

Somatic cell count as indicator of mastitis

SCC in bovine milk is an indicator of an inflammatory process in the udder quarter. CM can easily be detected by observing abnormalities, such as flakes or clots in milk and redness, hardness or pain in the udder. The detection of SCM is more difficult, since no clinical

symptoms are present. Quantification of the number of SCC in milk is often used to determine whether a cow is infected with mastitis pathogens or not. Previously it has been reported that in healthy quarters the somatic cells consist mostly of immune cells, such as lymphocytes and macrophages (Sarikaya et al. 2006) with variation from < 50 000 – 200 000 cells/ml (Schepers et al. 1997; Pantoja et al. 2009). While in the case of an inflammation SCC will increase (Sarikaya et al. 2006) and is expected to be above 200 000 cells/ml (Schepers et al. 1997). However, several authors have analysed different SCC threshold levels and concluded that alternatively defined SCC traits provide information that may improve genetic evaluation and selection on udder health (e.g. de Haas et al. 2008; Winding et al. 2010; Koeck et al. 2012; Bobbo et al. 2018). Furthermore, SCC have higher heritability compared to CM (Rupp and Biochard, 2003). Genetically, SCC is an indicator of both types of mastitis – clinical and subclinical (Martin et al. 2018). SCC records are routinely collected and used as an indicator of mastitis in bovine breeding programs in many countries. In genetic evaluations it is common to use the $\log_e$ transformed lactation-average somatic cell score (Schukken et al. 1992).

Selection against high Somatic Cell Count

Selection against high SCC in milk is a selection against inflammation and neutrophil concentration in milk (Rainard et al. 2018). Selection for lower SCC is often based on the LSCS trait and used to genetically improve udder health. However, in GS one of the methods includes estimation of the GEBV directly from genetic markers located on the genome (Goddard 2011). Based on a reference population effect of specific SNP markers on the trait of interest are estimated. SNP associated with QTL of the trait is probably in linkage disequilibrium (LD) with it. LD means that there is a non-random association of alleles (a variant form of a given gene) at different loci in a given population. Locus (loci in plural) is a fixed position for a gene/genetic marker on a chromosome. The detected markers are correlated with effects on the traits of

interest across families and further can be used for selection. Genome-wide association study (GWAS) is one of the most frequently used method to identify markers associated with trait phenotypes (Delvin & Risch 1995). The trait associated markers will contribute to more precise selection as well as better understanding of the biological mechanisms of complex traits, such as functional traits – fertility, SCC (*e.g.* Kühn et al. 2003; Sahana et al. 2010; Sodeland et al. 2011) and CM (*e.g.* Sahana et al. 2013; Wu et al. 2015; Olsen et al. 2016). In the case of CM in NR significant QTL was previously reported on bovine chromosome (BTA – *Bos Taurus* autosome) 2, 3, 4, 6, 14, 20 and 27 (Klungland et al. 2001; Nilsen et al. 2009; Sodeland et al. 2011; Olsen et al. 2016). While QTL affecting SCC have been reported on BTA8, BTA12, BTA19 and BTA26 (Klungland et al. 2001; Sodeland et al. 2011) for NR, it has been reported on BTA5, BTA6, BTA13, BTA16, BTA19 and BTA20 (Sahana et al. 2014) for other breeds such as Nordic Red cattle. GWAS results can be further analysed using topologically associated domain (TAD) regions to identify candidate genes (Wang et al. 2018; see the chapter "Activity of genes" below for more details). Furthermore, candidate genes can be studied by pathway-based analysis in order to understand biological functions behind complex traits, such as mastitis in dairy cattle (Lewandowska-Sabat et al. 2012; Olsen et al. 2016). In order to activate the biological processes in the udder the immune system of an individual have to recognize invading pathogens.

Immune system

The immune system protects the body from infectious agents entering the organism and is classified into the innate and the adaptive immune system. Innate immune response, also called inherited, identify pathogens and provide the signals necessary to activate an adaptive immune response, also known as acquired (Gallucci & Matzinger 2001). The signals are cytokines, that are proteins, secreted by cells during external stimulus and providing coordination of immune

cells. The first line of innate immune response cells are monocytes and macrophages. When bacteria or other pathogens invade the udder, monocytes migrate from blood to the tissue and differentiate into macrophages that can perform phagocytosis. Earlier it has been reported that *S. aureus* can survive within mammalian host cells and escape from macrophage phagosomes (Bayles et al. 1998). This may lead to chronic inflammation by inducing alternative activation of macrophages (Figure 2; Parham 2015). Normally, when bacteria enter the macrophage, T helper 1 cell type activation start a pro-inflammatory response through interferon gamma (IFN-γ), a cytokine that binds to the IFN-γ receptor on macrophages. The secondary possible activation of macrophages can be performed through a signal delivered when the CD40 ligand, a cytokine bound to the membrane of T helper 1 cells, binds to it's receptor CD40, on macrophages. Further, combination of the signals coming from the CD40 and IFN-γ receptors lead to changes in gene expression in order to activate macrophages. Only macrophages with cytokine receptors bound to the T helper 1 cells become antigen-specific activated and can react on specific antigens from invading pathogens, a mechanism preventing an unnecessary damage of healthy tissue by activated macrophages. However, macrophages may produce a large amount of inflammatory cytokines in response to an antigen, that can contribute to potentially pathological inflammatory processes, hyperinflammation and eventually pathological tissue damage (Hamidzadeh et al. 2017). Alternative activation of the macrophages is characterized with anti-inflammatory effects that may inhibit the pro-inflammatory cytokines (Parham 2015). There are several possible types of alternative activation of macrophages (Avdic et al. 2013). The first type acts through the T-helper 2 cytokines interleukin 4 and 13 (IL-4; IL-13) resulting in tissue repair macrophages. In the second type, macrophages become deactivators of the immune response after stimulation by interleukin 10 (IL-10). The last possible alternative activation is known as induced by immune complexes and agonists of Toll-like receptors.

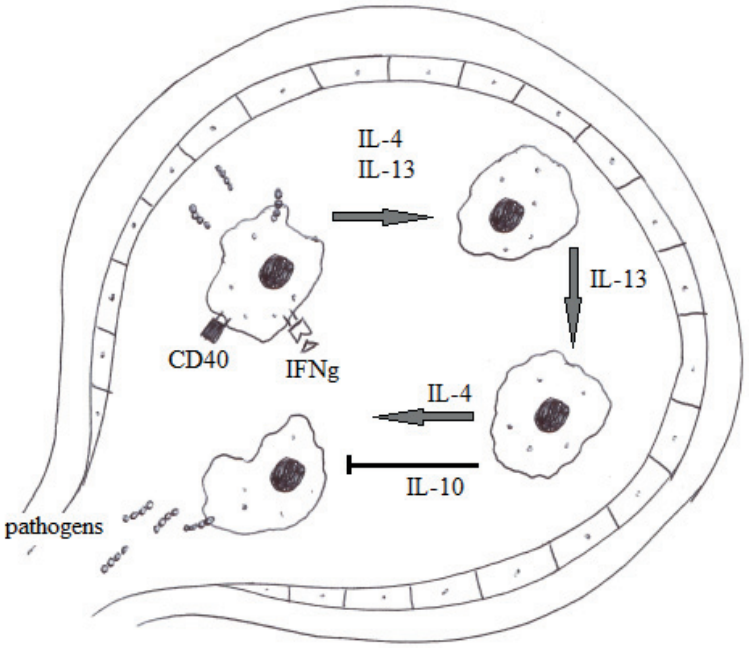


Figure 2. Classically and alternatively activated macrophages. In udder quarters and alveolus macrophages become antigen-specific activated by signals from CD40 and interferon gamma (IFNg) receptors. Such called classically activated macrophages start to produce interleukin 4 and 13 (IL-4, IL-13) to activate other macrophages. Alternatively activated macrophages produce IL-10 that inhibits pro-inflammatory cytokine signalling.

Activity of genes

The central dogma of biology is known as a two-step process, starting with replication of information provided by DNA by transcription into messenger RNA (mRNA) followed by

translation from mRNA to protein (Watson et al. 2008). The best known types of RNA are a mRNA, transfer RNA (tRNA) and ribosomal RNA (rRNA). Function of the mRNA is to carry information from DNA to the ribosomal sites of protein synthesis. tRNA translate the genetic codes of the mRNA into amino acids. Each class of tRNA binds a specific amino acid and transfer it to the ribosome for incorporation in the protein. rRNA combine with protein to form the ribosome (nucleoprotein). In addition, several other types of RNA, such as microRNA (miRNA) and long non coding RNA (lncRNA) exists. These are non coding but have regulatory functions. The miRNA binds to and destabilises mRNA and thereby inhibits protein synthesis (O`Brien et al. 2018). lncRNA can regulate the expression of genes in *cis* (neighbouring genes) or *trans* (distant genes) by binding to protein or other nucleic acids and modulating their function (Bassett et al. 2014). To initiate transcription of a gene the RNA polymerase binds to the promoter upstream of the gene sequence (Lewin 2000). A promoter is a region of DNA that is located close to the start site of transcription of a specific gene. Promoters regulate gene expression together with other regulatory regions, such as enhancers, that activate transcription of a gene, and silencers – that deactivate transcription. Normally, gene expression is controlled by silencers and enhancers, but non-genetic i.e. epigenetic, factors influence the expression of genes in an organism. Epigenetic inheritance denotes phenotypes inherited without any changes in the sequence of DNA. Such epigenetic modifications can change the structure of chromatin resulting in transcriptional activation or repression (Das & Singal 2004). DNA methylation and histone modification are examples of the mechanisms that cause functional changes in the genome without changes in the nucleotide sequence. DNA methylation is a process where methyl groups are added to the DNA molecule and leads to repression of transcription of specific genes. While histone modification may influence all DNA-based processes, such as transcription and chromatin compaction (Lawrence et al. 2016). Chromatin located in the nucleus includes a complex of proteins and DNA and forms chromosomes. In chromatin, long

DNA molecules become packaged in to compact shape, to avoid DNA damage and to contribute to regulation of DNA replication and gene expression (Mondal et al. 2010). A TAD (Figure 3; Dixon et al. 2012) is a genomic region characterised by more frequent self-interaction of DNA sequences within the TAD region compared to sequences outside the region. TAD regions are defined by analyses of chromatin conformation, by using capture methods with following sequencing. TADs are also defined as chromatin domains with important role in gene expression regulation.

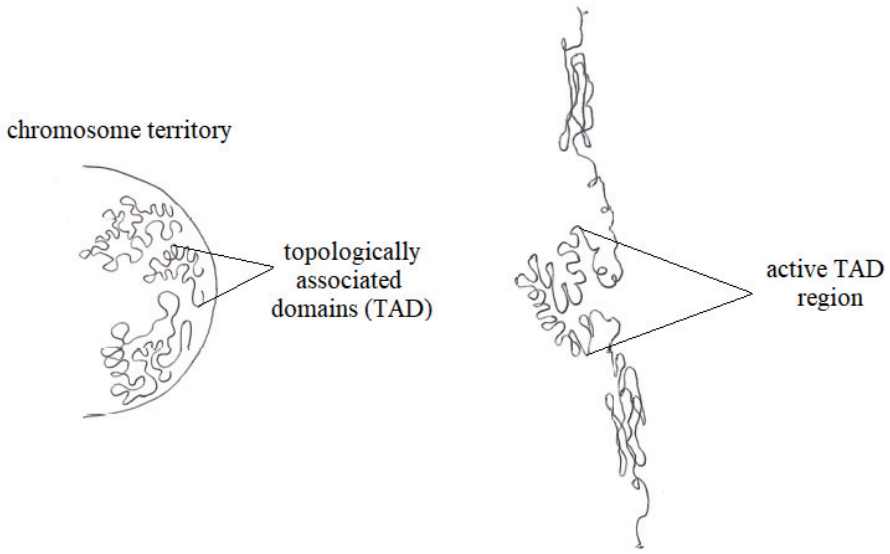


Figure 3. Topologically associated domains (TAD). Schematic structure of chromosome territory is shown with TAD regions. The active TAD has relaxed structure and consist of a region with active genes, surrounded by compacted regions with inactive genes.

TAD (Figure 3) are known from human biology as highly conserved domains with self-interacting chromatin (Krefting et al. 2018). TAD are stable during replication (Pope et al. 2014) and they are conserved between different cell types and between species (Dixon et al. 2012). Such stability and impact on the functional regulation of the genome makes TAD useful for identification of underlying gene regulation and mechanisms of a trait or disease. Moreover, information on TAD can improve the search for candidate gene variants compared to using arbitrarily chosen distances from significant SNPs associated with a trait (Wang et al. 2018). In the bovine genome coordinates for TAD regions based on bovine sequence data are not yet available (September 2019), but they have been described and analysed previously by using a liftover tool (Wang et al. 2018). The TAD coordinates available from several human, mice, dog and macaque cell types were successfully converted to the bovine genome. Furthermore, the most significant QTL associated with variation in expression of a specific gene or alleles of a gene were found to be located within the same TAD as the gene more often than expected. Additionally, Krefting et al. (2018) confirmed that TAD are highly conserved regions of the genome with regulatory functions and are stable against rearrangements during evolution. Therefore, TAD regions may be used as alternative search regions for causative regulatory variants in the bovine genome.

AIM OF THE STUDY

The primary aim of this project was to increase the understanding of the genetic and functional background of chronic SCM in NR dairy cows. Hence, the following sub-aims were formulated:

1. Define alternative chronic SCM traits based on the available information and estimate heritability and genetic correlations for these traits
2. Perform GWAS to identify candidate genes associated with chronic SCM, based on alternative chronic SCM traits
3. Study expression profiles of candidate genes in groups of animals with high and low GEBV for LSCS
4. Analyse expression profiles (mRNA) in bMDM, infected *in vitro* with *S. agalactiae*, an agent causing SCM in cattle

SUMMARY OF THE PAPERS

Paper I

Alternative Subclinical Mastitis Traits for Genetic Evaluation in Dairy Cattle

Chronic SCM in NR cows were analysed as 12 new traits, defined based on SCC thresholds of 50-, 100-, 150-, 200-, 250-, 300-, 350- and 400 x 1000 cells/ml (SCM50 – SCM400). Binary SCM traits were set equal to 1 if at least two consecutive testday SCC records within 2 months were above a fixed threshold, otherwise 0. Moreover, two similar traits based on three testdays within 3 months with SCC above 200- or 400 x 1000 cells/ml, respectively, were analysed in order to identify long lasting subclinical infections. Number of days before first case with SCM50=1 (D50) and SCM400=1 (D400) were analysed to take the time aspect into account. Data from the NDHRS were edited and used (n=701 440) for estimation of variance components by linear animal repeatability model. A smaller dataset was used to estimate genetic correlations (n=243 556). The total mean frequency of SCM traits ranged from 1.2 % (SCM400_3) to 51.8 % (SCM50). The mean number of days were 104 for D50 and 117 for D400. The mean LSCS was 4.4, equivalent to approximately 82 000 SCC in milk.

The estimated heritability for the twelve SCM traits ranged from 0.01 (SCM400_3) to 0.10 (SCM100). Estimated genetic correlation between SCM traits ranged from 0.7 (between D50 and SCM400) to 1 (between SCM350 and SCM400). Additionally, heritability for LSCS was estimated to 0.3 and genetic correlation between LSCS and the SCM traits ranged from 0.9 (with D50) to 1 (with SCM50). Estimated genetic correlation between milk yield trait and SCM ranged from 0.07 (LSCS) to 0.30 (D400). Based on the results defined traits might provide additional information on chronic SCM and can improve breeding for better udder health.

Paper II

Identification of Candidate Genes Affecting Chronic Subclinical Mastitis in Norwegian Red cattle: Combining Genome Wide Association Study, Topologically Associated Domains and Pathway Enrichment Analysis

The 12 new chronic SCM traits from Paper I, were further analysed using GWAS in order to identify candidate genes affecting these traits. Genotype data from NR bulls (n=3 795) were combined with phenotype information from their daughters (n=3 543 764). First, daughter yield deviations (DYD) for each genotyped NR sire were calculated using phenotype data from cows and corresponding pedigree file (n=4 124 678). Secondly, DYD data was combined with SNP (n=613 908) genotypes from each sire and the corresponding genomic relationship matrix, based on SNP data provided by Geno SA. The genomic location of identified significant SNPs (n=36 unique SNPs) were used to choose bovine TAD regions for the following selection of potential candidate genes. The list of genes (unique gene ID n=181) were analysed by the web-based software IPA in order to recognize involved canonical pathways and networks.

The LSCS trait showed similar significant genes as the SCM traits with SCC > 50 000 and >100 000 cells/ml in two testdays in a row (SCM50 and SCM100), i.e. checkpoint clamp loader component (*RAD17*) and cyclin B1 (*CCNB1*). Moreover, the SCM traits with SCC > 250 000; 300 000; 350 000; or 400 000 cells/ml at two testdays in a row and D400 displayed similar top significant genes; acyl-CoA thioesterase 2 and 4 (*ACOT2*; *ACOT4*). For the traits SCM150, SCM200 and SCM200_3, a group of chemokine (C-X-C motif) ligand genes and the Fos proto-oncogene, AP-1 transcription factor subunit (*FOS*) gene were identified. These results illustrates the complexity of chronic SCM. Further functional studies are necessary in order to clarify the actual role of these genes in development of chronic SCM in NR.

Paper III

Expression Analysis of Candidate Genes for Chronic Subclinical Mastitis in Norwegian Red

Several candidate genes for the alternatively defined chronic SCM traits from Paper II were selected to compare RNA expression profiles from unstimulated peripheral blood mononuclear cell (PBMC) in different groups of animals. Blood samples were taken from twenty NR cows selected based on their GEBV for LSCS. RNA expression were studied in association with leukocyte subset in the samples. Using RT-qPCR and flow cytometry methods differences in gene expression were found between animals grouped by high/low GEBV for LSCS and early/late lactation stage. The gene *CXCL1* showed higher expression and the gene *FOS* showed lower expression in the group with high GEBV for LSCS. Grouping by lactation stage revealed significant higher expression of FOS gene in early lactation (two – three month after calving) compared to late (seven-eight month after calving). No significant differences were found in leukocyte subsets among the grouped samples, indicating that the identified gene expression were unlikely to be caused by elevated numbers of cells in the PBMC samples. The identified genes in the current study can be considered as possible candidate genes for chronic SCM in NR cows.

Paper IV

Transcriptional Profiling of Monocyte-Derived Macrophages Infected *in vitro* with Two Strains of *Streptococcus agalactiae* Reveals Candidate Pathways Affecting Subclinical Mastitis in Cattle

For better understanding of the host defence during subclinical mastitis gene expression profiles in bovine monocyte derived macrophages (bMDMs) from blood samples were studied. bMDMs were isolated from twelve NR cows and stimulated *in vitro* with the two *Streptococcus agalactiae* strains ST12 and ST103. The individuals were selected based on their GEBV for LSCS. The cells were challenged with bacteria and left for 6 h. Unchallenged controls were included. Further, mRNA isolated from the bMDMs were sequenced and the results investigated for differentially expressed (DE) genes and pathway analyses. This resulted in 5936 and 6443 DE genes between groups infected with ST103 and ST12 respectively versus unchallenged controls. Ingenuity pathway analysis (IPA) of the up-regulated DE genes showed significant enrichment for type 1 and type 2 T helper cell activation in the samples stimulated with ST12. While significant enrichment of granulocyte adhesion and diapedesis was identified among the up-regulated genes in the samples stimulated with ST103. The gene *MYO1F* was found as significantly higher expressed in the group of low GEBV for LSCS bMDMs challenged by ST103 compared to high GEBV for LSCS. While the *PADI4* gene was lower expressed in bMDMs challenged by ST12 in the group of low GEBV for LSCS. No significantly DE genes were identified between low and high GEBV for LSCS in the control samples. In addition the samples were grouped by high and low SCC in milk given as the geometric mean of three measurements during the last lactation. Six, three and two DE genes were identified in samples challenged by ST103, ST12 and control, respectively. This study reported pathogen-induced regulation of key genes activated in bMDMs during immune response against mastitis causing *Streptococcus agalactiae* in NR.

MATERIALS AND METHODS

Animals and data

In Paper I and II, available data from the NDHRS (TINE Oslo, Norway) from 1979 to 2016 were analysed. The data were used to define traits based on SCC thresholds, defined as described by Svendsen and Heringstad (2006) with 400 000 cells/ml as the highest acceptable level of SCC in bovine milk allowed for human consumption (European Commission Milk Hygiene Directive 92/46, 1992). Ten SCM traits were defined as binary (set to 1 or 0) and scored as 1 if two testday SCC in a row within a 2 month period were above SCC thresholds from 50 000 to 400 000 cells/ml (SCM50, -100, -150, -200, -250, -300, -350, -400); and if three testday SCC in a row within a 3 month period were above 200 000 and 400 000 cells/ml (SCM200_3, SCM400_3). The hypothesis in Paper I was that high SCC during a 2 testdays long period will indicate cows with chronic long-term SCM infection, and more long-lasting stronger chronic SCM cases during a 3 testdays long period. Moreover, two traits: D50 and D400 were defined as the number of days from calving to the first case with SCC threshold above 50 or 400 x 1000 cells/ml, respectively, during 2 testdays within a 2 month period. These traits were supposed to give information on how long cows can stay without any increase in SCC in milk. The hypothesis was that cows who stay longer with low SCC have better ability to resist infections and chronic SCM compared to other NR cows. Finally, LSCS, the trait currently used in routine genetic evaluations of udder health for NR, was analysed. The data were restricted to testday SCC records during the first 3 lactations and between days in milk from 21 to 305 (to avoid the naturally high SCC after calving) and only from herd-years with at least 15 NR cows were included. Two datasets (A and B) from NDHRS were used: year/month of first calving: A) from 2006.01 to 2016.12 and B: from 2014.01 to 2016.12; age at calving in months by lactation number: from 19 to 63 months during 1 – 3 lactations for both

A and B dataset; interval from calving to conception called days open: from 20 to 150 days; herd year: A) 14 860 and B) 13 450. Dataset A including 701 440 cows with corresponding 1 209 128 records and 1 473 837 individuals in a pedigree file, was used for estimation of variance components. The software DmuTrace (Madsen 2012) was used to build the pedigree file by including available relationship information based on the ID-number of animals. The subset of dataset A, the smaller dataset B, including 243 556 cows with 357 203 observations and the corresponding pedigree file of 933 049 individuals, was used for estimation of genetic correlations. It was necessary to reduce the dataset for estimation of the parameters due to computational limitations.

In Paper II, genotype information from 3795 NR bulls with 613 908 SNPs was combined with phenotype records from their 3 543 764 daughters (first calving in 1979 – 2016) with corresponding 7 300 847 observations of testday SCC records, used to define the traits in Paper I and 4 126 678 individuals in the pedigree file. Genotype data on NR sires provided by Geno SA (www.norwegianred.com), the breeding organization for NR, were analysed. Geno SA has typed NR sires using the Affymetrix 25 SNP array (Affymetrix, Santa Clara) or the Illumina BovineSNP50 BeadChip (54K; Illumina Inc., San Diego, CA; Matukumalli et al. 2009) combined with the Illumina BovineHD Genotyping BeadChip (777K; Illumina, San Diego). Further, Geno SA has imputed missing SNP data.

In Paper III, 20 NR cows in the research herd at The Animal Production Experimental Centre, Norwegian University of Life Sciences, were selected for collection of blood samples and test for expression of candidate genes. These cows were chosen based on lactation number and calving date, in addition to maximizing the differences in the cows GEBV for LSCS below and above 100, that is high risk (< 100) and low risk (> 100) to get high SCC in milk during lactations. In Paper IV, twelve NR cows from different herds in Norway were selected based on GEBV for LSCS, maximizing the difference between groups.

Statistical and Genetic Analyses

In Paper I, the SAS software version 9.4 (SAS Inst. Inc., Cary, NC; www.sas.com) was used to edit data. (Co)variance components for the twelve defined SCM traits and the LSCS trait were estimated with the DMU software DMUAI program (Madsen and Jensen 2013) using a linear animal repeatability model. The animal repeatability model is used for estimation of repeated records, such as SCM traits for different lactation numbers (Mrode et al. 2012). For estimation of the heritability the single trait repeatability model was used for each trait, and for estimation of the genetic correlations bivariate repeatability models were used. When traits are defined as binary, the results of the linear model will depend on the frequency in the studied population (Gianola 1982). Therefore, low frequency may result in underestimated heritability for the trait. In Paper I, heritability and genetic correlations were calculated. To estimate heritability the estimated genetic variance was divided by the sum of estimated herd-year, permanent environmental, genetic and residual variances. The covariance between two traits was divided by the geometric mean of standard deviations (SD) for those traits to calculate genetic correlations. Finally, standard errors (SE) were calculated by the DMU software as the asymptotic SE.

The variance components estimated in Paper I were used in Paper II, as input for the estimation of breeding values (EBV) by the DMU5 procedure in the DMU software (Madsen and Jensen 2013). Solutions for the fixed and random effects were then used for calculation of the daughter yield deviations (DYD; VanRaden & Wiggans 1991) for the NR sires, DYD being the average of daughter phenotypes corrected for fixed and non-genetic random effects. DYD were used as input in the GWAS. Identified significant SNPs were further analysed by using converted bovine genome TAD coordinates. TAD regions are known to be conserved between the different cell types and species (Dixon et al. 2012; Rao et al. 2014; Rudan et al. 2015) as well

as stable during replication (Pope et al. 2014). The percentage of TADs successfully converted to the bovine genome ranged between 92.6 – 93.2 % and 86.0 – 86.8 % from human and mouse respectively. Then, the lists of identified genes were analysed using the network based software IPA (www.qiagenbioinformatics.com), known as one of the most updated pathway analysis software, primary designed for human pharmaceutical industry. IPA has specific focus on human diseases such as cancer, and this need to be taken into consideration when interpreting the results in relation to chronic SCM in cattle.

In Paper III, statistical analyses of the expression results for the identified candidate genes was performed using GraphPad Prism version 8.1.1 software (GraphPad Software, La Jolla California USA, www.graphpad.com). To estimate differences in gene expression levels between the groups of cows with high and low GEBV for LSCS, multiple t-test with following False Discovery Rate (FDR) correction of the estimated p-values was used. The significance level was defined at 5 % FDR P-value (Q-value).

To estimate statistically significant differences between DE transcripts in control and treatment samples in Paper IV, the packages DESeq2 v.1.4.5 (Love et al. 2014) and edgeR v.3.20.9 (Robinson & Oshlack 2010) in the R software v.3.4.1 (http://www.R-project.org) were used. DE genes were defined as significant when both FDR was < 5 % in edgeR and the Benjamini and Hochberg corrected p-value was < 5 % in DESeq2. In Paper II and IV, using IPA to analyse gene function, Fisher Extract-Test at 5 % level and Benjamini and Hochberg corrected p-values ≤ 5 %, respectively were used to define the significance level of candidate gene expression. Bonferroni correction was used to control the family-wise error rate when testing multiple hypotheses (Vickerstaff et al. 2019). Bonferroni correction is very conservative and ignore correlation between markers that become higher with higher density of markers (Han e al. 2009). In the case of low power of QTL, identifying Bonferroni correction will result in a high number of false negatives. In Paper II, the significance level of the identified SNPs was not

high and in order to include most SNPs of interest in further analyses the 5 % chromosome-wide significance threshold (Sahana et al. 2010) was implemented.

Genome-wide Association Study

In Paper II, GWAS was used to analyse the association between the traits defined in Paper I and SNPs to identify involved genes/chromosomal regions. To perform the GWAS analyses the package GenABEL v.1.8-0 (Aulchenko et al. 2007) in R software v.3.4.1 (http://www.r-project.org/) was used. GWAS is commonly used to identify genes associated with phenotypes of interest. In Paper II, DYD were used as phenotypes in a GWAS model with GRAMMAR-Gamma function, a variance component-based two-step method used for unbiased estimation of SNP effects for samples with genetic sub-structure, such as relatives (Svischeva et al. 2012). During the first step residuals and GRAMMAR-Gamma factor are estimated, using the polygenic function. The polygenic function give the inverse of the covariance matrix and estimates the residuals of the analysed traits by using kinship matrix of relatedness. In Paper II, the kinship relationship matrix was estimated from SNP genotype data provided by Geno SA. To avoid biased estimates of regression coefficient it is recommended to use the GRAMMAR-Gamma function (Svischeva et al. 2012). During the first step in the analysis the GRAMMAR-Gamma factor have been estimated, which gives correct estimation of test statistic and unbiased estimates of SNP effects. Further, in the second step the estimated effect of SNPs and test statistic values are divided by the GRAMMAR-Gamma genomic inflation factor.

Analyses of gene expression

To perform analysis of gene expression reverse transcription quantitative polymerase chain reaction (RT-qPCR) and mRNA sequencing was used. These methods are used for detection and quantification of specific genes in the form of mRNA sequences. In the current project RT-

qPCR were performed to identify the expression level of selected candidate genes for chronic SCM (Paper III).

The RT-qPCR method is divided in three steps. The first step is the reverse transcription (RT) where the enzyme reverse transcriptase produce a complementary DNA strand (cDNA) from the RNA. During the second step, amplification, cDNA is amplified through repeated cycles of denaturation and polymerization producing copies of the sequence targeted by specific primers. Additionally, at the end of each cycle fluorescence is emitted to quantify the level of double strand DNA (dsDNA). The final step, quantification, includes relative quantification by defining the number of amplification cycles needed for the fluorescence signal to cross a pre-set threshold. The cycle threshold (Ct) value is calculated as the cycle number at which the amplification curve cross the defined fluorescence threshold. The Ct value is related to the level of target RNA in the analysed sample, i.e. the higher Ct value, the lower level of target RNA.

In Paper III, 40 cycles were run in the qPCR using the EXPRESS SYBR™ GreenER™ qPCR Supermix, with premixed ROX (Invitrogen) dye that intercalates in all dsDNA and emits a fluorescent signal, followed by melting curve analyses. The EXPRESS SYBR™ GreenER™ qPCR Supermix kit (https://www.thermofisher.com/) provides higher sensitivity, specificity and reproducibility by producing a brighter signal, less PCR inhibition and reduced contamination in qPCR compared to the original SYBR Green I. However, the SYBR Green detection is not specific for the target sequence, since it binds to all dsDNA including erroneous amplicons. Hence, additional melting curve analysis for each RT-qPCR run is necessary to test for product purity. In addition, PCR products for each of the designed primers were analysed for size on agarose gel and sequenced to ensure that the correct amplicons were produced. Moreover, for each sample no RT controls were included to test for contamination of the RNA samples with genomic DNA. Without the RT reaction pure samples should only consist of RNA

and the PCR reactions will not work. For each primer pair negative controls with no added template (no template control) were included to test for contamination of the reaction mix .

In Paper III the RT-qPCR method was used to evaluate expression of genes selected as possible candidates with effect on chronic subclinical mastitis in NR cows, based on the GWAS analyses. All gene specific primers were designed using the Primer BLAST tool (Ye et al. 2012) and checked for polymorphisms as far as possible in the NR genome to ensure their specificity. To quantify the level of gene expression normalizing the Ct values against the expression of reference (housekeeping) genes, with the same level of expression in all experimental situations, was performed. In addition, to examine if the differences in expression of the analysed genes were caused by differences in cellular expression and not by different numbers of monocytes, dendritic cells or other fractions of PBMC, flow cytometry was performed. Mixes of primary monoclonal antibodies and secondary isotype-specific antibodies were used for specific designation of the types of leukocyte cells in the PBMC samples.

The bovine geNORM kit combined with the software qbase+ 3.2 (http://www.qbaseplus.com) were used (Paper III) to identify the most stable reference genes expressed in the bovine PBMC samples. This kit includes primers for six housekeeping genes (*GAPDH*, *PPIA*, *ATP5B*, *EIF2B2*, *RPL12*, *SDHA*) that are candidates for reference genes to be used in the samples. Several RNA samples isolated from PBMC (representing each group of high and low LSCS GEBV) were tested for all of the six reference genes and the most stable results were obtained for the *EIF2B2* and *RPL12* genes. The average expression stability value (M) were calculated as the lowest among the six genes and were below 0.3 for both *EIF2B2* and *RPL12*. Hence, these two genes were chosen as optimal reference targets and calculated as the geometric mean for normalization of the expression level of the tested genes. To calculate the changes in expression level of the genes of interest the $2^{-\Delta Ct}$ method was used. This method performs correction of gene expression against reference genes in the same run, i.e.

$\Delta Ct = Ct_{(target\ gene)} - Ct_{(reference\ gene)}$. Additionally, the efficiency of all primer pairs were tested by template dilution series. Three samples from cows with high GEBV for LSCS and three from low GEBV were used. The number of molecules of the target sequences should double during each replication cycle and in our study (Paper III) the efficiency of all the primer pairs were 100 % (±10).

Alternatively, RNA sequencing and differential expression analysis can be performed. In Paper IV mRNA isolated from bMDM challenged *in vitro* by the mastitis causing *S. agalactiae* strains ST12 and ST103 were sequenced in order to analyse for DE genes among defined groups. The samples with bMDM were exposed to bacteria for 1 hour before addition of 1% penicillin/streptomycin to the cultures in order to stop further growth of extracellular bacteria which would kill the cells within a couple of hours. The control cultures without bacteria were treated likewise. By sequencing RNA isolated from the samples, all the involved genes in the bovine genome may be identified. Moreover, the sequence data can be analysed in order to identify SNPs by comparing data from several individuals. This method, called variant calling, are used to identify SNP variants from RNA sequencing data (Piskol et al. 2013; Olson et al. 2015). But this is a multistep process that may give wrongly mapped reads and could hamper accurate identification of genomic variants. Microarray analysis is an older method that also may reveal differentially expressed genes (Lewandowska-Sabat et al. 2013). But by using a microarray analysis only the genes printed on the array can be identified, while sequencing is not depending on known information and can reveal unknown genes and splice variants expressed in the analysed samples.

KEY RESULTS AND GENERAL DISCUSSION

The aim of this project was to increase the understanding of the genetic and functional background of chronic SCM in NR dairy cows. The hypothesis is that cows with low SCC in milk have different gene variants that manage a quick and effective immune response reacting on an invading pathogen. Those cows are likely to have better ability to resist infections, including development of chronic SCM, compared to cows who stay with high SCC in milk for a prolonged period of time. Moreover, individuals with high and low GEBV for LSCS were expected to have different gene expression profiles in PBMC. Statistical and genetic analyses were performed in order to test the hypothesis.

The 12 alternative traits for chronic SCM were analysed in order to investigate different thresholds of SCC in milk. Several traits were defined, since it is still not clear which threshold is optimal to use to define the chronic SCM cases. Strong genetic correlations were estimated among the SCM traits, ranging from 0.7 (D50 and SCM400) to 1 (SCM350 and SCM400) as described in Paper I. The estimated heritability was low and ranged from 0.01 (SCM400_3 and D400) to 0.12 (SCM100) among the defined SCM traits. Similar traits have been defined for other dairy cattle populations (*e. g.* Koeck et al. 2010; Bobbo et al. 2018; Narayana et al. 2018; Bobbo et al. 2019). Several authors have studied alternative SCC traits (reviewed by Martin et al. 2018), in order to identify the most informative traits for udder health. However, estimated genetic parameters, such as the heritability and genetic correlations of the defined traits cannot be compared directly, due to use of different models for estimation of the parameters and different trait definitions. Estimated genetic correlations between alternatively defined SCC traits and SCM reported previously vary from a positive correlation of 0.99 to favourable negative correlation of -0.85, depending on definition of the traits (de Haas et al. 2008; Windig et al. 2010; Urioste et al. 2012). The estimated heritability was reported to vary from 0.05 to

0.09 for traits defined similar to SCM200 – SCM50, respectively (Svendsen and Heringstad 2006), from 0.09 to 0.14 for a trait similar to SCM150 (Koeck et al. 2012; Urioste et al. 2012; Bobbo et al. 2018) and from 0.11 to 0.19 for LSCS (Luttinen & Juga 1997; Ødegård et al. 2004; Haugaard et al. 2013). In Paper I, cows with high SCC for a long-term period, that is two and three testdays in a row above different SCC thresholds was analysed. We assumed that such traits may indicate chronic SCM events. The advantage of studying SCC is that SCC is an indicator for both CM and SCM (Martin et al. 2018) and more heritable than the direct record of CM (Rupp and Biochard, 2003). Increase of SCC in udder is mainly caused by influx of white blood cells required for response to invading pathogens (Coffey et al. 1986). Therefore, SCC thresholds is often used as indicator of udder health and a combination of several defined SCM traits could better represent the chronic SCM events. Including these traits in genetic evaluations of udder health may improve resistance to chronic SCM. Most studies have used a linear model for estimation of SCC and SCM heritability (Koivula et al. 2004; Svendsen and Heringstad 2006; de Haas et al. 2008; Urioste et al. 2012; Narayana et al. 2018; Bobbo et al. 2019). This model fits for routine estimation based on a large dataset, due to less computational demand. While, the alternative threshold model that can be used for categorical traits (0 or 1) may pose a problem in routine genetic evaluation using large dataset (Narayana et al. 2018). In the current study the linear animal repeatability model was used because it fits best for further estimation of DYD used in the GWAS in Paper II. GWAS (Delvin & Risch 1995) is a widely used tool to identify significant SNPs associated with traits of interest. In Paper II, using GWAS with DYD as the bulls phenotype, a total of 36 significant SNPs were identified across all the defined traits, with 16 SNPs significant for more than one trait and 20 significant for one trait only. The chronic SCM traits were associated with several SNPs located on BTA2 (2), BTA5 (1), BTA6 (1), BTA7 (1), BTA10 (5), BTA12 (8), BTA20 (12) and BTA26 (6), with number of SNPs in parenthesis. To date, little information on sources of variation in novel

chronic SCM traits has been published. Corresponding QTL for SCC has previously been reported on BTA5, BTA6, BTA7, BTA10, BTA12, BTA20 and BTA26, but with different chromosomal regions or significant SNPs for Nordic Holstein, Nordic Red, Danish Jersey, Holstein-Fresian and German Holstein cattle populations (Klungland et al. 2001; Kuhn et al. 2003; Sahana et al. 2010; Sodeland et al. 2011; Meredith et al. 2012; Meredith et al. 2013; Raven et al. 2014; Sahana et al. 2014). For NR significant QTL for SCC and SCS were reported by Klungland et al. (2001) on BTA8 by using microsatellites and by Sodeland et al. (2011) on BTA12, BTA19 and BTA26 by using SNPs, respectively. Microsatellites are short DNA fragments with repetitive motifs (one – six basepair(s)) yielding highly polymorphic fragment lengths. Microsatellites are scattered in the genome and has been widely used as markers to map genome-wide LD (Farnir et al. 2000). In the current study, to the best of our knowledge, we used the highest number of markers (n=613 908 SNPs) in NR that has been analysed in order to identify candidate genes involved in chronic SCM. The differences in the GWAS results compared to the literature may be explained by the present use of a better bovine reference genome and different sets of animals with more comprehensive data on phenotypes and genotypes. In the present work, the most updated version of the bovine genome at the time of the study UMD3.1.1 (UCSC Genome Browser assembly ID: bosTau8) was used. Further, the SNP chip of 777K was used including a larger number of SNPs, compared to 25K SNPs and 288 markers studied by Sodeland et al. (2011) and Klungland et al. (2001), respectively. The number of animals analysed is also higher in our study and might represent more genetic variation within the NR population. Variability both at individual and breed level may explain differences in the results (Begley et al. 2009; Mielczarek et al. 2018). The NR breed is known to have more genetic variation due to cross breeding and historical intercrossing of old Norwegian breeds (www.norwegianred.com). For CM QTL in the NR population have been reported on BTA2, BTA3, BTA4, BTA6, BTA14, BTA20, BTA27 (Klungland et al. 2001;

Nilsen et al. 2009; Sodeland et al. 2011; Olsen et al. 2016). It is well known that CM has a positive genetic correlation with SCC (e.g. Lund et al. 1999; de Haas et al. 2008). Genetic correlations between alternative SCM traits and CM were estimated by Svendsen and Heringstad (2006) and ranged from 0.26 to 0.62. CM have lower heritability than SCC, ranging from 0.02 – 0.04 as estimated with the linear model (Heringstad et al. 2000) and from 0.06 – 0.12 when the threshold models were used (Lund et al. 1999; Heringstad et al. 2003; Zwald et al. 2004). The heritability for SCC in the current study (Paper I) was estimated to 0.26 for LSCS and ranged from 0.01 to 0.12 for the alternatively defined SCM traits, using the linear model. SCM and SCC are not the same trait genetically and the alternatively defined traits displayed different correlations with LSCS, which is used in genetic evaluation of udder health for NR. Therefore, including alternatively defined traits from the current study in genetic evaluations will give more information in order to improve genetic selection against chronic SCM, since neither CM nor LSCS represent the full spectrum of the occurrence of chronic subclinical mastitis in the population.

In our study (Paper II) we have combined the results from GWAS with converted bovine TAD positions, that revealed 181 unique genes with 168 genes mapped by the following IPA analyses. A reason for using TAD regions instead of arbitrarily chosen distances from significant SNPs to identify potential genes involved in the analysed trait is that TAD represent functional domains of the genome (Figure 3) which in turn can improve the search for causative variants as reported by Wang et al. (2018). The resulting list of genes in the current study was further analysed and the most frequent genes from top canonical pathways or the genes that were unique for only one trait were selected for expression analysis in unstimulated cells from blood samples (Paper III). Among other results in Paper II, the genes *RAD17*, *CDK7* and *CCNB1* were involved in SCM50 – SCM200 and LSCS, traits with not necessarily very high SCC in milk. While *CXCL8* and *CXCL1* were unique for SCM200_3, the most representative

trait for long term chronic SCM, since based on definition the threshold of SCC remain above 200 000 cells/ml in milk within three continuous testday measurements within a 3 month period. The genes *GLI2*, *SLC18A2*, *SEL1L* and *STAT4* were only found in association with the SCM400_3 trait and the *C6* gene was associated with D50. The genes *ACOT2*, *ACOT4*, *TGFB3*, and *FOS* were associated with SCM150 – SCM400 and D400, and the *FOS* gene was in addition associated with the SCM200_3 trait. For better understanding of the function of the identified genes we investigated gene expression profiles from unstimulated samples (Paper III) and samples challenged with *S. agalactiae* (Paper IV), a bacterium causing subclinical mastitis in dairy cattle. Reports on expression in blood cells were not available for several of the selected genes, hence we decided to test all. The genes *GLI2*, *C6* and *SCL18A2* turned out to not be expressed in bovine PBMC. The *GLI2* gene was previously found to be expressed in the liver of Qinchuan cattle (Liu et al. 2014) and *C6* was found to be expressed in plasma samples of Chinese Holstein cows (Min et al. 2016). We did not find any information on *SLC18A2* and expression in bovine blood samples. The remaining tested genes were expressed in the unstimulated bovine PBMC samples (Paper III). Two samples tested for expression of the *TGFB3* gene had to be excluded due to large differences (SD > 0.3) between repeated triple reactions. This might be due to undiscovered polymorphisms in the regions for primer binding making binding of primers unstable. Therefore, 18 samples were analysed for this gene. In order to identify significant results samples were grouped by official GEBV for LSCS as high (>100, n=10) vs low (<100, n=10) and stage of lactation as early period (one – two month after calving, n=12) and late period (seven-eight month after calving, n=8). Two (*CXCL1* and *FOS*) of nine genes displayed significant differential expression in the unstimulated PBMC samples. *CXCL1* was significantly higher expressed in samples from high GEBV for LSCS based on q-value (0.02). The *FOS* gene was higher expressed in the group of cows with low GEBV for LSCS and in the early lactating period, based on p-value (0.02) and q-value (0.05), correspondingly.

The chemokine ligand *CXCL1* gene situated 637 998 bp from a significant SNP located on BTA6 (at 90 184 750 bp, Paper II), was previously identified to be included in udder inflammatory response to *S.aureus* causing clinical and subclinical mastitis (Rainard et al. 2008). *CXCL1* is a neutrophil-oriented chemokine, involved in massive influx of neutrophils to infected tissue (Sipka et al. 2014). Neutrophils play a major role in the mammary gland immunity (Verbeke et al. 2015). Neutrophils together with macrophages, natural killer cells and cytokine signals constitute the first line defence of the innate immune system. During bovine mastitis the number of neutrophils and macrophages increase in bovine milk, through migration from the bloodstream to the mammary gland when foreign invading pathogens are recognized (Oviedo-Boyso et al. 2007). Both macrophages and neutrophils can perform phagocytosis of bacteria and be responsible for the production of several cytokines required for the immune response (Stelwagen et al. 2009). The ability to perform a quick immune response define the susceptibility or resistance to infection. In bovine mastitis bacteria can survive within the host cells and escape phagocytosis by triggering alternative activation of macrophages (Gordon et al. 2010). This may lead to chronic inflammation (Parham 2015) with high SCC in milk for a prolonged period of time. Therefore, the difference between NR cows in how they handle chronic SCM may be due to higher cellular expression of *CXCL1*. This might result in more effective influx of neutrophils and other immune cells from blood to the infected tissue in order to respond against invading pathogens, and quickly kill the majority of the bacteria. Further, the *FOS* gene is a transcription factor reported previously in association with *S. aureus* infection in NR (Lewandowska-Sabat et al. 2013). This gene play a main role in proliferation, differentiation and survival of cells (Bahrami and Drabløs 2016) and can be activated and react within minutes after stimulation (Healy et al. 2013). Such potential for quick reaction may explain low SCC in milk for cows with high GEBV for LSCS. In the UMD3.1.1 genome assembly *FOS* was found to be situated on BTA10 (at 86 887 kb) more than 1 Mb from a

significant SNP position (87 958 kb). Such results make the identified genes interesting candidates for regulation of chronic SCM in NR and this merits further investigation. Additionally, we have performed flow cytometry analyses (Paper III) to investigate the leukocyte subsets in the PBMC samples and identify number and type of cells as potential sources for differential gene expression. However, no significant differences in lymphocyte cells were found between the groups of high and low GEBV for LSCS or stage of lactation, i.e. the results reported for the *CXCL1* and *FOS* genes must be caused by higher cellular expression. Most studies have analysed bovine gene expression in samples stimulated with different bacteria (*e.g.* Casey et al. 2015; Hop et al. 2017; Roy et al. 2018). In our study we also studied differences in immune response of isolated bMDMs from cows with low (n=6) and high (n=6) GEBV for LSCS challenged *in vitro* with *S. agalactiae* strains ST12 and ST103 (Paper IV). Additionally, the cows were grouped by phenotypes, high (n=3) and low (n=9) SCC in milk. To identify DE genes mRNA isolated from the bMDMs were sequenced and mapped against the bovine genome assembly UMD3.1. The following DE analysis revealed 5936 and 6443 significantly DE genes (p-value < 0.05) in bMDMs challenged by ST103 and ST12 respectively, vs unchallenged controls. While between ST12 and ST103, 588 DE genes were identified. There were no difference between high and low GEBV for LSCS in the control samples. The myosin IF (*MYO1F*) gene was up-regulated in low GEBV for LSCS samples challenged with ST103 and peptidyl arginine deiminase 4 (*PADI4*) down-regulated in low GEBV for LSCS samples challenged with ST12. The gene *MYO1F*, has an important role in neutrophil trafficking during inflammation, through neutrophil migration (Salvermoser et al. 2018) and *PADI4* has a role in neutrophil regulation (Suzuki et al. 2003). Interesting results were found in the groups of SCC phenotype in milk. In the unstimulated control groups the genes phospholipid phosphatase 1 (*PLPP1*) and macrophage receptor with collagenous structure (*MARCO*) were up-regulated in samples with high SCC. The *MARCO* gene have a

role in recruitment of mononuclear cells and pro-inflammatory cytokines in case of bacterial infection, while the *PLPP1* gene is important for the hydrolysis and uptake of lipids from extracellular space. Moreover, the *MARCO* gene were up-regulated in samples stimulated with both ST12 and ST103 in cows with high SCC compared to low. Hence, the gene *MARCO* is suggested to be used as a potential biomarker for persistent chronic mastitis, since transcription of this gene was found in macrophages long after the exposure (Paper IV). The *PADI4* and *MYO1F* genes may be considered as candidates with important roles in chronic bovine mastitis. Indicating that NR cows with low GEBV might have reduced neutrophil response due to the down-regulation of *PADI4* and higher SCC in milk because of up-regulated *MYO1F*.

To summarise, several genes have been identified in association with alternatively defined chronic SCM traits in NR dairy cows. These traits were genetically different from LSCS, the trait currently used in routine genetic evaluations of udder health for NR cattle, i.e. the traits analysed in the current study will add valuable information in order to improve udder health. The threshold of 200 000 cells/ml (SCM200) commonly used to define SCM have shown equally significant genes to 150 000 cells/ml (SCM150), indicating similar background mechanisms for these threshold. Further, the results of gene expression (*CXCL1*, *MYO1F* and *PADI4*) points to a role of neutrophils in development of chronic SCM, since the *CXCL1* gene was found in association with the SCM200_3 trait. However, further investigation of the candidate genes is necessary for improved understanding of their function in relation to chronic SCM in NR cattle. Our analyses of unstimulated blood cells in healthy animals were performed during mid lactation, that is a period with low level of stress to the cow. Analyses of gene expression profiles during higher stress levels, such as close to parturition, might be useful for better understanding of the function of the candidate genes in development of chronic SCM and SCC in bovine milk.

CONCLUSIONS

This thesis provides new knowledge on the genetic and functional background of chronic subclinical mastitis in the Norwegian Red cattle population.

- Novel alternative chronic SCM traits have been defined and analysed. The new traits were heritable and had a genetic correlation with milk yield from 0.07 to 0.34 and to the LSCS varying from 0.88 to 0.99 (Paper I).
- Thirty-six significant SNPs were revealed among the defined SCM traits (Paper II). Identical SNPs were identified for SCM150 and SCM200 and for SCM250 and SCM300. Otherwise the SCM traits displayed different sets of SNPs, demonstrating the complexity of the chronic SCM traits.
- *CXCL1* with higher cellular expression for cows with high GEBV for LSCS was revealed as a candidate gene based on TAD regions including or nearby significant SNP (Paper II, III). The *FOS* gene was revealed significant when cows were grouped by lactation stage with a higher expression for cows in an early period of lactation.
- Candidate genes (Paper III, IV) central in regulation of neutrophil trafficking (*CXCL1*, *MYO1F* and *PADI4*) and recruitment of mononuclear cells and pro-inflammatory cytokine production in response to bacterial infection (*MARCO*) have been identified.
- Our results indicate that cows with high GEBV for LSCS have potential for a quicker and more effective immune response by calling more neutrophil cells from blood in order to respond to pathogens invading the udder.

PROPOSITIONS FOR FUTURE STUDIES

- Further analyses of the identified candidate genes *CXCL1*, *FOS, MYO1F*, *PADI4* and *MARCO* in order to understand the function of these genes in association with chronic SCM in NR.
- Evaluation of a new udder health index for NR including novel chronic SCM traits to find an optimal combination of the defined traits that might be most informative.
- Expand investigation on differences in the genetic background between individuals with high and low GEBV for udder health with focus on the identified candidate genes.
- Gene expression analyses of a higher number of NR cows to identify genes differentially expressed between the groups.
- Analyses of blood and milk samples of same individuals for a prolonged period of time (throughout lactation periods) in order to study changes over time in gene expression profiles.

REFERENCES

1. Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. van Duijn. 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23(10):1294-1296. doi: https://doi.org/10.1093/bioinformatics/btm108

2. Avdic, S., J. Z. Cao, B. P. McSharry, L. E. Clancy, R. Brown, M. Steain, D. J. Gottlieb, A. Abendroth, and B. Slobedman. 2013. Human cytomegalovirus interleukin-10 polarizes monocytes toward a deactivated M2c phenotype to repress host immune responses. J Virol 87(18):10273-10282. doi: https://doi.org/10.1128/JVI.00912-13

3. Bahrami, S. and F. Drablos. 2016. Gene regulation in the immediate-early response process. Adv Biol Regul 62:37-49. doi: https://doi.org/10.1016/j.jbior.2016.05.001

4. Bassett, A. R., A. Akhtar, D. P. Barlow, A. P. Bird, N. Brockdorff, D. Duboule, A. Ephrussi, A. C. Ferguson-Smith, T. R. Gingeras, W. Haerty, D. R. Higgs, E. A. Miska, and C. P. Ponting. 2014. Considerations when investigating lncRNA function in vivo. Elife 3:e03058. doi: https://doi.org/10.7554/eLife.03058

5. Bayles, K. W., C. A. Wesson, L. E. Liou, L. K. Fox, G. A. Bohach, and W. R. Trumble. 1998. Intracellular Staphylococcus aureus escapes the endosome and induces apoptosis in epithelial cells. Infect Immun 66(1):336-342.

6. Begley, N., F. Buckley, K. M. Pierce, A. G. Fahey, and B. A. Mallard. 2009. Differences in udder health and immune response traits of Holstein-Friesians, Norwegian Reds, and their crosses in second lactation. J Dairy Sci 92(2):749-757. doi: https://doi.org/10.3168/jds.2008-1356

7. Bobbo, T., M. Penasa, R. Finocchiaro, G. Visentin, and M. Cassandro. 2018. Alternative somatic cell count traits exploitable in genetic selection for mastitis resistance in Italian

Holsteins. J Dairy Sci 101(11):10001-10010. doi: https://doi.org/10.3168/jds.2018-14827

8. Bobbo, T., C. Roveglia, M. Penasa, G. Visentin, R. Finocchiaro, and M. Cassandro. 2019. Genetic relationships of alternative somatic cell count traits with milk yield, composition and udder type traits in Italian Jersey cows. Anim Sci J 90(7):808-817. doi: https://doi.org/10.1111/asj.13204

9. Bogni, C., L. Odierno, C. Raspanti, J. Giraudo, A. Larriestra, E. Reinoso, M. Lasagno, M. Ferrari, E. Ducrós, and C. Frigerio. 2011. War against mastitis: Current concepts on controlling bovine mastitis pathogens. Science against microbial pathogens: Communicafing current research and technological advances:483-494.

10. Casey, M. E., K. G. Meade, N. C. Nalpas, M. Taraktsoglou, J. A. Browne, K. E. Killick, S. D. E. Park, E. Gormley, K. Hokamp, D. A. Magee, and D. E. MacHugh. 2015. Analysis of the bovine monocyte-derived macrophage response to Mycobacterium avium subspecies paratuberculosis infection using RNA-seq. Frontiers in Immunology 6. doi: https://doi.org/ARTN2310.3389/fimmu.2015.00023

11. Coffey, E. M., W. E. Vinson, and R. E. Pearson. 1986. Somatic cell counts and infection rates for cows of varying somatic cell count in initial test of first lactation. J Dairy Sci 69(2):552-555. doi: https://doi.org/10.3168/jds.S0022-0302(86)80437-4

12. Das, P. M. and R. Singal. 2004. DNA methylation and cancer. J Clin Oncol 22(22):4632-4642. doi: https://doi.org/10.1200/JCO.2004.07.151

13. de Haas, Y., W. Ouweltjes, J. ten Napel, J. J. Windig, and G. de Jong. 2008. Alternative somatic cell count traits as mastitis indicators for genetic selection. J Dairy Sci 91(6):2501-2511. doi: https://doi.org/10.3168/jds.2007-0459

14. Devlin, B. and N. Risch. 1995. A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 29(2):311-322. doi: https://doi.org/10.1006/geno.1995.9003

15. Dixon, J. R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485(7398):376-380. doi: https://doi.org/10.1038/nature11082

16. Farnir, F., W. Coppieters, J. J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, M. Mni, C. Nezer, P. Simon, P. Vanmanshoven, D. Wagenaar, and M. Georges. 2000. Extensive genome-wide linkage disequilibrium in cattle. Genome Res 10(2):220-227. doi: https://doi.org/10.1101/gr.10.2.220

17. Gallucci, S. and P. Matzinger. 2001. Danger signals: SOS to the immune system. Current opinion in immunology 13(1):114-119.

18. Gianola, D. 1982. Theory and Analysis of Threshold Characters. Journal of Animal Science 54(5):1079-1096.

19. Goddard, M. E., B. J. Hayes, and T. H. E. Meuwissen. 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. Journal of Animal Breeding and Genetics 128(6):409-421. doi: https://doi.org/10.1111/j.1439-0388.2011.00964.x

20. Gordon, S. and F. O. Martinez. 2010. Alternative activation of macrophages: mechanism and functions. Immunity 32(5):593-604. doi: https://doi.org/10.1016/j.immuni.2010.05.007

21. Hagnestam-Nielsen, C., U. Emanuelson, B. Berglund, and E. Strandberg. 2009. Relationship between somatic cell count and milk yield in different stages of lactation. J Dairy Sci 92(7):3124-3133. doi: https://doi.org/10.3168/jds.2008-1719

22. Hamidzadeh, K., S. M. Christensen, E. Dalby, P. Chandrasekaran, and D. M. Mosser. 2017. Macrophages and the Recovery from Acute and Chronic Inflammation. Annu Rev Physiol 79:567-592. doi: https://doi.org/10.1146/annurev-physiol-022516-034348

23. Han, B., H. M. Kang, and E. Eskin. 2009. Rapid and accurate multiple testing correction and power estimation for millions of correlated markers. PLoS Genet 5(4):e1000456. doi: https://doi.org/10.1371/journal.pgen.1000456

24. Harmon, R. J. 1994. Symposium - Mastitis and Genetic Evaluation for Somatic-Cell Count - Physiology of Mastitis and Factors Affecting Somatic-Cell Counts. Journal of Dairy Science 77(7):2103-2112. doi: https://doi.org/DOI 10.3168/jds.S0022-0302(94)77153-8

25. Haugaard, K., B. Heringstad, and A. C. Whist. 2013. Genetic associations between somatic cell score and pathogen-specific subclinical mastitis in Norwegian Red cows. Journal of Animal Breeding and Genetics 130(2):98-105. doi: https://doi.org/10.1111/jbg.12019

26. Healy, S., P. Khan, and J. R. Davie. 2013. Immediate early response genes and cell transformation. Pharmacol Ther 137(1):64-77. doi: https://doi.org/10.1016/j.pharmthera.2012.09.001

27. Heringstad, B., G. Klemetsdal, and J. Ruane. 1999. Clinical mastitis in Norwegian cattle: frequency, variance components, and genetic correlation with protein yield. J Dairy Sci 82(6):1325-1330. doi: https://doi.org/10.3168/jds.S0022-0302(99)75356-7

28. Heringstad, B., G. Klemetsdal, and J. Ruane. 2000. Selection for mastitis resistance in dairy cattle: a review with focus on the situation in the Nordic countries. Livestock Production Science 64(2-3):95-106. doi: https://doi.org/Doi 10.1016/S0301-6226(99)00128-1

29. Heringstad, B., R. Rekaya, D. Gianola, G. Klemetsdal, and K. A. Weigel. 2003. Bivariate analysis of liability to clinical mastitis and to culling in first-lactation cows. J Dairy Sci 86(2):653-660. doi: https://doi.org/10.3168/jds.S0022-0302(03)73643-1

30. Hop, H. T., L. T. Arayan, A. W. B. Reyes, T. X. N. Huy, W. Min, H. J. Lee, J. S. Son, and S. Kim. 2017. Simultaneous RNA-seq based transcriptional profiling of intracellular

Brucella abortus and B. abortus-infected murine macrophages. Microb Pathogenesis 113:57-67. doi: https://doi.org/10.1016/j.micpath.2017.10.029

31. Jørgensen, H. J., A. B. Nordstoga, S. Sviland, R. N. Zadoks, L. Solverod, B. Kvitle, and T. Mork. 2016. Streptococcus agalactiae in the environment of bovine dairy herds--rewriting the textbooks? Vet Microbiol 184:64-72. doi: https://doi.org/10.1016/j.vetmic.2015.12.014

32. Klungland, H., A. Sabry, B. Heringstad, H. G. Olsen, L. Gomez-Raya, D. I. Vage, I. Olsaker, J. Odegard, G. Klemetsdal, N. Schulman, J. Vilkki, J. Ruane, M. Aasland, K. Ronningen, and S. Lien. 2001. Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle. Mamm Genome 12(11):837-842.

33. Koeck, A., B. Heringstad, C. Egger-Danner, C. Fuerst, P. Winter, and B. Fuerst-Waltl. 2010. Genetic analysis of clinical mastitis and somatic cell count traits in Austrian Fleckvieh cows. J Dairy Sci 93(12):5987-5995. doi: https://doi.org/10.3168/jds.2010-3451

34. Koeck, A., F. Miglior, D. F. Kelton, and F. S. Schenkel. 2012. Alternative somatic cell count traits to improve mastitis resistance in Canadian Holsteins. J Dairy Sci 95(1):432-439. doi: https://doi.org/10.3168/jds.2011-4731

35. Koivula, M., E. Negussie, and E. A. Mantysaari. 2004. Genetic parameters for test-day somatic cell count at different lactation stages of Finnish dairy cattle. Livestock Production Science 90(2-3):145-157. doi: https://doi.org/10.1016/j.livprodsci.2004.03.004

36. Krefting, J., M. A. Andrade-Navarro, and J. Ibn-Salem. 2018. Evolutionary stability of topologically associating domains is associated with conserved gene regulation. BMC Biol 16(1):87. doi: https://doi.org/10.1186/s12915-018-0556-x

37. Kuhn, C., J. Bennewitz, N. Reinsch, N. Xu, H. Thomsen, C. Looft, G. A. Brockmann, M. Schwerin, C. Weimann, S. Hiendleder, G. Erhardt, I. Medjugorac, M. Forster, B.

Brenig, F. Reinhardt, R. Reents, I. Russ, G. Averdunk, J. Blumel, and E. Kalm. 2003. Quantitative trait loci mapping of functional traits in the German Holstein cattle population. J Dairy Sci 86(1):360-368.

38. Lawrence, M., S. Daujat, and R. Schneider. 2016. Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends Genet 32(1):42-56. doi: https://doi.org/10.1016/j.tig.2015.10.007

39. Lewandowska-Sabat, A. M., G. M. Boman, A. Downing, R. Talbot, A. K. Storset, and I. Olsaker. 2013. The early phase transcriptome of bovine monocyte-derived macrophages infected with Staphylococcus aureus in vitro. BMC Genomics 14:891. doi: https://doi.org/10.1186/1471-2164-14-891

40. Lewandowska-Sabat, A. M., J. Gunther, H. M. Seyfert, and I. Olsaker. 2012. Combining quantitative trait loci and heterogeneous microarray data analyses reveals putative candidate pathways affecting mastitis in cattle. Anim Genet 43(6):793-799. doi: https://doi.org/10.1111/j.1365-2052.2012.02342.x

41. Lewin, B. 2000. Genes VII. Oxford University Press Inc., New York.

42. Liu, Y., H. Wang, G. Cheng, B. Jiang, and L. Zan. 2014. Expression analysis of Gli1 and Gli2 in different tissues and muscle-derived cells of Qinchuan cattle. Genet Mol Res 13(4):8767-8775. doi: https://doi.org/10.4238/2014.October.27.18

43. Loberg and Dürr. 2009. Interbull survey on the use of genomic information. Interbull Bulletin no 39.

44. Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550. doi: https://doi.org/10.1186/s13059-014-0550-8

45. Lund, M. S., J. Jensen, and P. H. Petersen. 1999. Estimation of genetic and phenotypic parameters for clinical mastitis, somatic cell production deviance, and protein yield in

dairy cattle using Gibbs sampling. J Dairy Sci 82(5):1045-1051. doi: https://doi.org/10.3168/jds.S0022-0302(99)75325-7

46. Luttinen, A. and J. Juga. 1997. Genetic Relationships between Milk Yield, Somatic Cell Count, Mastitis, Milkability and Leakage in Finish Dairy Cattle Population. Interbull Bulletin (15):pp. 78-83.

47. Madsen, P. 2012. DmuTrace. A program for extracting the pedigree for a subset of animals from a larger pedigree. Version 2. Aarhus University.

48. Madsen, P. and J. Jensen. 2013. A User`s Guide to DMU. A Package for Analysing Multivariate Mixed Models. Version 6, release 5.2. Aarhus University.

49. Martin, P., H. W. Barkema, L. F. Brito, S. G. Narayana, and F. Miglior. 2018. Symposium review: Novel strategies to genetically improve mastitis resistance in dairy cattle. J Dairy Sci 101(3):2724-2736. doi: https://doi.org/10.3168/jds.2017-13554

50. Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. F. Allan, M. P. Heaton, J. O'Connell, S. S. Moore, T. P. Smith, T. S. Sonstegard, and C. P. Van Tassell. 2009. Development and characterization of a high density SNP genotyping assay for cattle. PLoS One 4(4):e5350. doi: https://doi.org/10.1371/journal.pone.0005350

51. Meredith, B. K., D. P. Berry, F. Kearney, E. K. Finlay, A. G. Fahey, D. G. Bradley, and D. J. Lynn. 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. Front Genet 4:229. doi: https://doi.org/10.3389/fgene.2013.00229

52. Meredith, B. K., F. J. Kearney, E. K. Finlay, D. G. Bradley, A. G. Fahey, D. P. Berry, and D. J. Lynn. 2012. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. BMC Genet 13:21. doi: https://doi.org/10.1186/1471-2156-13-21

53. Meuwissen, T. H., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157(4):1819-1829.

54. Mielczarek, M., M. Fraszczak, E. Nicolazzi, J. L. Williams, and J. Szyda. 2018. Landscape of copy number variations in Bos taurus: individual - and inter-breed variability. BMC Genomics 19(1):410. doi: https://doi.org/10.1186/s12864-018-4815-6

55. Miller, R. H., M. J. Paape, and L. A. Fulton. 1991. Variation in milk somatic cells of heifers at first calving. J Dairy Sci 74(11):3782-3790. doi: https://doi.org/10.3168/jds.S0022-0302(91)78570-6

56. Min, L., J. B. Cheng, S. G. Zhao, H. Tian, Y. D. Zhang, S. L. Li, H. J. Yang, N. Zheng, and J. Q. Wang. 2016. Plasma-based proteomics reveals immune response, complement and coagulation cascades pathway shifts in heat-stressed lactating dairy cows. Journal of Proteomics 146:99-108. doi: https://doi.org/10.1016/j.jprot.2016.06.008

57. Mondal, T., M. Rasmussen, G. K. Pandey, A. Isaksson, and C. Kanduri. 2010. Characterization of the RNA content of chromatin. Genome Res 20(7):899-907. doi: https://doi.org/10.1101/gr.103473.109

58. Mrode, R., T. Pritchard, M. Coffey, and E. Wall. 2012. Joint estimation of genetic parameters for test-day somatic cell count and mastitis in the United Kingdom. J Dairy Sci 95(8):4618-4628. doi: https://doi.org/10.3168/jds.2011-4971

59. Narayana, S. G., F. Miglior, S. A. Naqvi, F. Malchiodi, P. Martin, and H. W. Barkema. 2018. Genetic analysis of subclinical mastitis in early lactation of heifers using both linear and threshold models. J Dairy Sci 101(12):11120-11131. doi: https://doi.org/10.3168/jds.2018-15126

60. Nilsen, H., H. G. Olsen, B. Hayes, T. Nome, E. Sehested, M. Svendsen, T. H. Meuwissen, and S. Lien. 2009. Characterization of a QTL region affecting clinical mastitis and

protein yield on BTA6. Anim Genet 40(5):701-712. doi: https://doi.org/10.1111/j.1365-2052.2009.01908.x

61. O'Brien, J., H. Hayder, Y. Zayed, and C. Peng. 2018. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. Front Endocrinol (Lausanne) 9:402. doi: https://doi.org/10.3389/fendo.2018.00402

62. Olsen, H. G., T. M. Knutsen, A. M. Lewandowska-Sabat, H. Grove, T. Nome, M. Svendsen, M. Arnyasi, M. Sodeland, K. K. Sundsaasen, S. R. Dahl, B. Heringstad, H. H. Hansen, I. Olsaker, M. P. Kent, and S. Lien. 2016. Fine mapping of a QTL on bovine chromosome 6 using imputed full sequence data suggests a key role for the group-specific component (GC) gene in clinical mastitis and milk production. Genet Sel Evol 48(1):79. doi: https://doi.org/10.1186/s12711-016-0257-2

63. Olson, N. D., S. P. Lund, R. E. Colman, J. T. Foster, J. W. Sahl, J. M. Schupp, P. Keim, J. B. Morrow, M. L. Salit, and J. M. Zook. 2015. Best practices for evaluating single nucleotide variant calling methods for microbial genomics. Front Genet 6:235. doi: https://doi.org/10.3389/fgene.2015.00235

64. Oviedo-Boyso, J., J. J. Valdez-Alarcon, M. Cajero-Juarez, A. Ochoa-Zarzosa, J. E. Lopez-Meza, A. Bravo-Patino, and V. M. Baizabal-Aguirre. 2007. Innate immune response of bovine mammary gland to pathogenic bacteria responsible for mastitis. J Infect 54(4):399-409. doi: https://doi.org/10.1016/j.jinf.2006.06.010

65. Pantoja, J. C., C. Hulland, and P. L. Ruegg. 2009. Dynamics of somatic cell counts and intramammary infections across the dry period. Prev Vet Med 90(1-2):43-54. doi: https://doi.org/10.1016/j.prevetmed.2009.03.012

66. Parham, P. 2015. The immune system. Garland Science, Taylor & Francis Group, LLC.

67. Piskol, R., G. Ramaswami, and J. B. Li. 2013. Reliable identification of genomic variants from RNA-seq data. Am J Hum Genet 93(4):641-651. doi: https://doi.org/10.1016/j.ajhg.2013.08.008

68. Pope, B. D., T. Ryba, V. Dileep, F. Yue, W. Wu, O. Denas, D. L. Vera, Y. Wang, R. S. Hansen, T. K. Canfield, R. E. Thurman, Y. Cheng, G. Gulsoy, J. H. Dennis, M. P. Snyder, J. A. Stamatoyannopoulos, J. Taylor, R. C. Hardison, T. Kahveci, B. Ren, and D. M. Gilbert. 2014. Topologically associating domains are stable units of replication-timing regulation. Nature 515(7527):402-405. doi: https://doi.org/10.1038/nature13986

69. Rainard, P., G. Foucras, D. Boichard, and R. Rupp. 2018. Invited review: Low milk somatic cell count and susceptibility to mastitis. J Dairy Sci 101(8):6703-6714. doi: https://doi.org/10.3168/jds.2018-14593

70. Rainard, P., A. Fromageau, P. Cunha, and F. B. Gilbert. 2008. Staphylococcus aureus lipoteichoic acid triggers inflammation in the lactating bovine mammary gland. Veterinary Research 39(5):52. doi: https://doi.org/10.1051/vetres:2008034

71. Rao, S. S. P., M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden. 2014. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159(7):1665-1680. doi: https://doi.org/10.1016/j.cell.2014.11.021

72. Raven, L. A., B. G. Cocks, and B. J. Hayes. 2014. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. BMC Genomics 15:62. doi: https://doi.org/10.1186/1471-2164-15-62

73. Robinson, M. D. and A. Oshlack. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11(3):R25. doi: https://doi.org/10.1186/gb-2010-11-3-r25

74. Roy, S., S. Schmeier, B. Kaczkowski, E. Arner, T. Alam, M. Ozturk, O. Tamgue, S. P. Parihar, H. Kawaji, M. Itoh, T. Lassmann, P. Carninci, Y. Hayashizaki, A. R. R. Forrest, R. Guler, V. B. Bajic, F. Brombacher, and H. Suzuki. 2018. Transcriptional landscape of Mycobacterium tuberculosis infection in macrophages. Scientific Reports 8. doi: https://doi.org/10.1038/s41598-018-24509-6

75. Rudan, M. V., C. Barrington, S. Henderson, C. Ernst, D. T. Odom, A. Tanay, and S. Hadjur. 2015. Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. Cell Reports 10(8):1297-1309. doi: https://doi.org/10.1016/j.celrep.2015.02.004

76. Rupp, R. and D. Boichard. 2003. Genetics of resistance to mastitis in dairy cattle. Vet Res 34(5):671-688. doi: https://doi.org/10.1051/vetres:2003020

77. Sahana, G., B. Guldbrandtsen, C. Bendixen, and M. S. Lund. 2010. Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. Anim Genet 41(6):579-588. doi: https://doi.org/10.1111/j.1365-2052.2010.02064.x

78. Sahana, G., B. Guldbrandtsen, B. Thomsen, L. E. Holm, F. Panitz, R. F. Brondum, C. Bendixen, and M. S. Lund. 2014. Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle. J Dairy Sci 97(11):7258-7275. doi: https://doi.org/10.3168/jds.2014-8141

79. Sahana, G., B. Guldbrandtsen, B. Thomsen, and M. S. Lund. 2013. Confirmation and fine-mapping of clinical mastitis and somatic cell score QTL in Nordic Holstein cattle. Animal Genetics 44(6):620-626. doi: https://doi.org/10.1111/age.12053

80. Salvermoser, M., R. Pick, L. T. Weckbach, A. Zehrer, P. Lohr, M. Drechsler, M. Sperandio, O. Soehnlein, and B. Walzog. 2018. Myosin 1f is specifically required for neutrophil

migration in 3D environments during acute inflammation. Blood 131(17):1887-1898. doi: https://doi.org/10.1182/blood-2017-10-811851

81. Sarikaya, H., G. Schlamberger, H. H. Meyer, and R. M. Bruckmaier. 2006. Leukocyte populations and mRNA expression of inflammatory factors in quarter milk fractions at different somatic cell score levels in dairy cows. J Dairy Sci 89(7):2479-2486. doi: https://doi.org/10.3168/jds.S0022-0302(06)72322-0

82. Schepers, A. J., T. J. Lam, Y. H. Schukken, J. B. Wilmink, and W. J. Hanekamp. 1997. Estimation of variance components for somatic cell counts to determine thresholds for uninfected quarters. J Dairy Sci 80(8):1833-1840. doi: https://doi.org/10.3168/jds.S0022-0302(97)76118-6

83. Schukken, Y. H., K. E. Leslie, A. J. Weersink, and S. W. Martin. 1992. Ontario Bulk Milk Somatic-Cell Count Reduction Program .2. Dynamics of Bulk Milk Somatic-Cell Counts. J Dairy Sci 75(12):3359-3366. doi: https://doi.org/DOI10.3168/jds.S0022-0302(92)78112-0

84. Schwarz, D., U. S. Diesterbeck, S. Konig, K. Brugemann, K. Schlez, M. Zschock, W. Wolter, and C. P. Czerny. 2011. Microscopic differential cell counts in milk for the evaluation of inflammatory reactions in clinically healthy and subclinically infected bovine mammary glands. J Dairy Res 78(4):448-455. doi: https://doi.org/10.1017/S0022029911000574

85. Sipka, A., S. Klaessig, G. E. Duhamel, J. Swinkels, P. Rainard, and Y. Schukken. 2014. Impact of intramammary treatment on gene expression profiles in bovine Escherichia coli mastitis. PLoS One 9(1):e85579. doi: https://doi.org/10.1371/journal.pone.0085579

86. Smith, B. P. 2014. Large Animal Internal Medicine. Elsevier Health Sciences, 4th Edition:1015-1043.

87. Sodeland, M., M. P. Kent, H. G. Olsen, M. A. Opsal, M. Svendsen, E. Sehested, B. J. Hayes, and S. Lien. 2011. Quantitative trait loci for clinical mastitis on chromosomes 2, 6, 14 and 20 in Norwegian Red cattle. Anim Genet 42(5):457-465. doi: https://doi.org/10.1111/j.1365-2052.2010.02165.x

88. Stelwagen, K., E. Carpenter, B. Haigh, A. Hodgkinson, and T. T. Wheeler. 2009. Immune components of bovine colostrum and milk. J Anim Sci 87(13 Suppl):3-9. doi: https://doi.org/10.2527/jas.2008-1377

89. Suzuki, A., R. Yamada, X. Chang, S. Tokuhiro, T. Sawada, M. Suzuki, M. Nagasaki, M. Nakayama-Hamada, R. Kawaida, M. Ono, M. Ohtsuki, H. Furukawa, S. Yoshino, M. Yukioka, S. Tohma, T. Matsubara, S. Wakitani, R. Teshima, Y. Nishioka, A. Sekine, A. Iida, A. Takahashi, T. Tsunoda, Y. Nakamura, and K. Yamamoto. 2003. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 34(4):395-402. doi: https://doi.org/10.1038/ng1206

90. Svendsen, M. and B. Heringstad. 2006. Somatic cell count as an indicator of sub-clinical mastitis. Genetic parameters and correlations with clinical mastitis. Interbull Bulletin (35):12.

91. Svishcheva, G. R., T. I. Axenovich, N. M. Belonogova, C. M. van Duijn, and Y. S. Aulchenko. 2012. Rapid variance components-based method for whole-genome association analysis. Nature Genetics 44(10):1166-1170. doi: https://doi.org/10.1038/ng.2410

92. Urioste, J. I., J. Franzen, J. J. Windig, and E. Strandberg. 2012. Genetic relationships among mastitis and alternative somatic cell count traits in the first 3 lactations of Swedish Holsteins. J Dairy Sci 95(6):3428-3434. doi: https://doi.org/10.3168/jds.2011-4739

93. VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J Dairy Sci 91(11):4414-4423. doi: https://doi.org/10.3168/jds.2007-0980

94. VanRaden, P. M. and G. R. Wiggans. 1991. Derivation, calculation, and use of national animal model information. J Dairy Sci 74(8):2737-2746. doi: https://doi.org/10.3168/jds.S0022-0302(91)78453-1

95. Verbeke, J., K. Piccart, S. Piepers, M. Van Poucke, L. Peelman, A. De Visscher, and S. De Vliegher. 2015. Somatic cell count and milk neutrophil viability of dairy heifers with specific CXCR1 genotypes following experimental intramammary infection with Staphylococcus chromogenes originating from milk. Vet J 204(3):322-326. doi: https://doi.org/10.1016/j.tvjl.2015.04.010

96. Vickerstaff, V., R. Z. Omar, and G. Ambler. 2019. Methods to adjust for multiple comparisons in the analysis and sample size calculation of randomised controlled trials with multiple primary outcomes. BMC Med Res Methodol 19(1):129. doi: https://doi.org/10.1186/s12874-019-0754-4

97. Viguier, C., S. Arora, N. Gilmartin, K. Welbeck, and R. O'Kennedy. 2009. Mastitis detection: current trends and future perspectives. Trends Biotechnol 27(8):486-493. doi: https://doi.org/10.1016/j.tibtech.2009.05.004

98. Wang, M., T. P. Hancock, A. J. Chamberlain, C. J. Vander Jagt, J. E. Pryce, B. G. Cocks, M. E. Goddard, and B. J. Hayes. 2018. Putative bovine topological association domains and CTCF binding motifs can reduce the search space for causative regulatory variants of complex traits. BMC Genomics 19(1):395. doi: https://doi.org/10.1186/s12864-018-4800-0

99. Watson, J. D., T. A. Baker, S. P. Bell, A. Gann, M. Levine, and R. Losick. 2008. Molecular biology of the gene. Pearson Education Inc., publishing as Benjamin Cummings, San Francisco.

100. Windig, J. J., W. Ouweltjes, J. Ten Napel, G. de Jong, R. F. Veerkamp, and Y. De Haas. 2010. Combining somatic cell count traits for optimal selection against mastitis. J Dairy Sci 93(4):1690-1701. doi: https://doi.org/10.3168/jds.2009-2052

101. Wu, X., M. S. Lund, G. Sahana, B. Guldbrandtsen, D. Sun, Q. Zhang, and G. Su. 2015. Association analysis for udder health based on SNP-panel and sequence data in Danish Holsteins. Genet Sel Evol 47:50. doi: https://doi.org/10.1186/s12711-015-0129-1

102. Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. L. Madden. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13:134. doi: https://doi.org/10.1186/1471-2105-13-134

103. Zadoks, R. N., W. B. van Leeuwen, D. Kreft, L. K. Fox, H. W. Barkema, Y. H. Schukken, and A. van Belkum. 2002. Comparison of Staphylococcus aureus isolates from bovine and human skin, milking equipment, and bovine milk by phage typing, pulsed-field gel electrophoresis, and binary typing. J Clin Microbiol 40(11):3894-3902.

104. Zwald, N. R., K. A. Weigel, Y. M. Chang, R. D. Welper, and J. S. Clay. 2004. Genetic selection for health traits using producer-recorded data. I. Incidence rates, heritability estimates, and sire breeding values. J Dairy Sci 87(12):4287-4294. doi: https://doi.org/10.3168/jds.S0022-0302(04)73573-0

105. Ødegard, J., B. Heringstad, and G. Klemetsdal. 2004. Short communication: bivariate genetic analysis of clinical mastitis and somatic cell count in Norwegian dairy cattle. Journal of Dairy Science 87(10):3515-3517. doi: https://doi.org/10.3168/jds.S0022-0302(04)73487-6

PAPERS I-IV

Paper I

J. Dairy Sci. 102:5323–5329
https://doi.org/10.3168/jds.2018-16104

Alternative subclinical mastitis traits for genetic evaluation in dairy cattle

E. Kirsanova,[1]* B. Heringstad,[2,3] A. Lewandowska-Sabat,[1] and I. Olsaker[1]
[1]Department of Basic Sciences and Aquatic Medicine, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, NO-0102 Oslo, Norway
[2]Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway
[3]Geno Breeding and AI Association, NO-2317 Hamar, Norway

ABSTRACT

Chronic subclinical mastitis (SCM), characterized by changes in milk composition and high somatic cell count (SCC) in milk for a prolonged period of time, is often caused by a bacterial infection. Different levels of SCC have been suggested and used as threshold to identify subclinical infection. The aim of this study was to examine different definitions of SCM based on test-day SCC and estimate genetic parameters for these traits and their genetic correlation to milk production. Test-day SCC records from 1,209,128 Norwegian Red cows in lactation 1 to 3 were analyzed. Twelve SCM traits were defined as binary with 2 test-day SCC in a row above SCC thresholds from 50,000 to 400,000 cells/mL (SCM50, SCM100, SCM150, SCM200, SCM250, SCM300, SCM350, and SCM400), with 3 test-day SCC in a row above 200,000 and 400,000 cells/mL (SCM200_3 and SCM400_3), and the number of days before the first case with SCM50 (D50) or SCM400 (D400). The heritability and genetic correlations were estimated for SCM traits and the mean lactation-average somatic cell score (LSCS) using linear animal repeatability models. The total mean frequency of SCM ranged from 1.2% to 51.8%, for different trait definitions, high for low SCC threshold (SCM50) and low for the highest SCC threshold (SCM400_3). For the 2 traits based on number of days, the mean values were 104 (D50) and 117 (D400) days. The mean LSCS was 4.4 (equivalent to around 82,000 SCC). Heritabilities for the 12 alternative SCM traits were low and varied from 0.01 (SCM400_3) to 0.1 (SCM100), whereas for LSCS the estimated heritability was 0.3 and standard error varied from 0.001 to 0.003. Genetic correlations among the SCM traits ranged from 0.7 (D50 and SCM400) to 1 (SCM350 and SCM400), whereas between SCM traits and milk production the correlation ranged from 0.07 (LSCS) to 0.3 (D400). The standard error for genetic correlations varied from 0.001 to 0.06. The heritability was low and the genetic correlations were strong among SCM traits. Genetic correlations lower than 1 suggest that the alternative SCM traits are genetically different from LSCS, the trait currently used in genetic evaluations for Norwegian Red. Hence, the alternative traits will add information and improve breeding for better udder health.
Key words: somatic cell count, Norwegian Red, heritability, genetic correlation

INTRODUCTION

Mastitis is one of the most frequent and costly diseases in dairy cattle, characterized as an inflammatory process in the mammary gland, often with increased SCC in milk (Harmon, 1994). Clinical mastitis (**CM**) is characterized by clinical symptoms such as severe udder inflammation, whereas chronic subclinical mastitis (**SCM**) is recognized by changes in milk composition such as high SCC for prolonged periods of time (Harmon, 1994; Østerås et al., 2007). Normal SCC in healthy udder quarters has been reported to be lower than 100,000 cells/mL (Schwarz et al., 2011). High SCC, commonly used to characterize SCM, is defined as more than 200,000 cells/mL, and will often lead to decreased milk production (Dohoo and Leslie, 1991; Harmon, 1994; Schepers et al., 1997; Østerås et al., 2007; Pantoja et al., 2009). *Staphylococcus aureus* and *Streptococcus agalactiae* are known as some of the most common bacteria causing udder inflammation. According to the health status report for Norwegian Red (**NR**) cows in 2017, *S. aureus* is the major causative bacterium of mastitis, found to be present in 25.4% of milk samples (TINE SA, 2017). Somatic cells consist of epithelial cells together with polymorphonuclear leucocytes and macrophages involved in defense against infections in the udder (Miller et al., 1991; Schwarz et al., 2011). Hence, the amount of these cells in milk gives an indicator of the inflammatory process in the udder.

Alternative definitions of the SCC-based mastitis traits have been analyzed previously. Windig et al.

Received December 5, 2018.
Accepted February 11, 2019.
*Corresponding author: elena.kirsanova@nmbu.no

(2010) defined SCM as binary traits based on whether or not 2 consecutive SCC test days were above 150,000 and 250,000 cells/mL by using Dutch herd records from the Nederlands Rundvee Syndicaat. Koeck et al. (2012) defined 7 SCC traits, for different parts of the lactation, for genetic evaluation of mastitis in Canadian Holsteins. Urioste et al. (2012) defined SCM in Swedish Holsteins as number of periods from days in milk >45 with SCC >150,000 cells/mL. de Haas et al. (2008) estimated high genetic correlation (0.98) between SCM, defined as a binary trait (0/1), as absence or presence of test-day SCC >150,000 cells/mL in lactation, and suspicion of IMI. Svendsen and Heringstad (2006) defined SCM traits as present if 2 test days with SCC were above a fixed threshold of 50,000 to 200,000 cells/mL and reported genetic correlation among these SCM traits was between 0.89 and 0.92. Estimated heritability was lower for higher thresholds and ranged from 0.05 to 0.09 across the 3 first lactations.

Which SCC threshold to use to classify chronic SCM is still not clear. Also, most of the literature is on Holsteins, and breed differences have been reported for udder health and immune response (e.g., Begley et al., 2009). Hence, the aim of this study was to examine novel alternative definitions of chronic SCM traits in NR based on prolonged high SCC, and estimate heritability of and genetic correlations among these SCM traits together with their genetic correlation to milk yield.

MATERIALS AND METHODS

Animals and Phenotypes

Phenotypic data have been collected routinely in the Norwegian Dairy Herd Recording System since 1978, which is a complete historical database for the NR population that is available for analyses. Records of test-day SCC for NR cows with calving in the years 2006 to 2016 were obtained from the Norwegian Dairy Herd Recording System. Data set A with 1,209,128 observations from 701,440 cows with a corresponding pedigree file of 1,473,837 individuals was used for estimation of variance components and heritability. Records from the first 3 lactations of cows from herd-years with at least 15 NR cows were included. The data set was further restricted to only test-day SCC records from DIM between 21 and 305, and from lactations with 2 or more test day records. Due to computational limitations, a smaller data set B, including information from 2014 to 2016, was used for estimation of genetic correlations. The last data set contained 357,203 observations from 243,556 cows and had a corresponding pedigree file of 933,049 individuals.

Trait Definition and Distribution

Four SCM traits were defined with the same SCC thresholds of 50, 100, 150, and 200 $\times$ 10^3 cells/mL as Svendsen and Heringstad (2006). Additional traits with higher SCC thresholds up to 400 $\times$ 10^3 cells/mL, the highest acceptable level of SCC in bovine milk allowed for human consumption (European Commission Milk Hygiene Directive 92/46, 1992) were also included. Thus, 8 binary SCM traits were defined based on SCC thresholds of 50, 100, 150, 200, 250, 300, 350, and 400 $\times$ 10^3 cells/mL (**SCM50**, **SCM100**, **SCM150**, **SCM200**, **SCM250**, **SCM300**, **SCM350**, and **SCM400**, respectively). The SCM50 to SCM400 were set equal to 1 if SCC was above the given threshold at 2 test days in a row within a 2-mo period, otherwise 0. The hypothesis is that high SCC at 2 test days will indicate cows with chronic long-term SCM infection.

Two additional novel SCM traits were defined for SCC thresholds 200 and 400 $\times$ 10^3 cells/mL; **SCM200_3** and **SCM400_3** were set to 1 if SCC was above the threshold at 3 test days in a row within a 3-mo period, otherwise 0. By requiring 3 test days with high SCC, which indicates stronger, more long-lasting chronic SCM cases, we assume these traits will identify cows unable to overcome chronic subclinical mastitis, neither going into a clinical phase nor able to reduce SCC in milk.

The number of days from calving to the first case of 2 test-day SCC records above 50 or 400 $\times$ 10^3 cells/mL during a 2-mo period were defined as **D50** and **D400**, respectively. These traits are of interest because they could give information on how long cows can stay uninfected without any increase in SCC. We assume that cows who stay longer with SCC below 50 $\times$ 10^3 cells/mL have better ability to resist chronic SCM compared with other cows. Thus, these traits can provide additional information for genetic evaluations and indicate cows with either better or worse ability to get SCM for a prolonged period of time.

Additionally, **LSCS** (the lactation-average SCS; $\log_e$ SCC; Schukken et al., 1992) was included in the analysis. The mean frequencies for the alternative SCM traits from first to third lactation based on 559,988, 401,848, and 247,292 phenotypic records, respectively, for data set A and 163,513, 119,727, and 73,963 for data set B are given in Table 1. The mean 305-d milk yield was 7,441 kg in data set A and 7,844 kg in data set B.

Statistical Analyses

Software packages SAS, version 9.4 (SAS Inst. Inc., Cary, NC) and DMU (Madsen and Jensen, 2013) were used to perform editing and genetic analyses, respec-

tively. The pedigree files were built by the DmuTrace software (Madsen, 2012). (Co)variance components were estimated with REML using the DMUAI program (Madsen and Jensen, 2013). Heritability was estimated by single trait repeatability models, whereas genetic correlations between the traits were estimated using bivariate models.

The following linear animal repeatability model was used for estimation of the variance components for all the SCM traits:

$$y = year\_month + age + days\_open + herd\_year + pe + animal + e,$$

where y is the observations for each trait, *year_month* is the fixed effect of calving year/month (A: 396 levels: 2006.01 to 2016.12; B: 108 levels: 2,014.01 to 2,016.12), *age* is the fixed effect of age at calving in months by lactation number (58 levels: from 19 to 63 mo during 1–3 lactations), *days_open* is the fixed effect of days open (i.e., calving-to-conception interval; number of days from 20 to 150 grouped each 10 d for each lactation), *herd_year* is the random effect of herd year (A: 49,860; B: 13,450 levels), *pe* is the permanent environmental effect of repeated cow measurements, *animal* is the additive genetic effect of animal, and e is the residual. An exception was made for the LSCS trait, by changing the *herd_year* effect to fixed in the model.

Heritability was calculated using the formula

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{hy}^2 + \sigma_{pe}^2 + \sigma_e^2},$$

where σ_g^2 is the estimated genetic variance, and σ_{hy}^2, σ_{pe}^2, and σ_e^2 are the estimated herd-year, permanent environmental, and residual variance, respectively.

Genetic correlations between the traits were calculated by the following formula:

$$corr = \frac{Cov(g_1, g_2)}{\sqrt{\sigma_{g1}\sigma_{g2}}},$$

where $Cov(g_1,g_2)$ is covariance between 2 traits and $\sqrt{\sigma_{g1}\sigma_{g2}}$ is the geometric mean of standard deviations for those traits.

Standard errors (**SE**) estimated from asymptotic SE calculated by DMU, based on the AI-information matrix including the asymptotic correlation matrix parameter vector.

RESULTS AND DISCUSSION

Descriptive Statistics

The mean frequency of the SCM traits ranged from 0.8 to 64.7%, presented in Table 1. For all traits, the frequency increased with higher lactation number, but

Table 1. The mean frequency of alternative chronic subclinical mastitis (SCM) traits,[1] mean lactation-average somatic cell score (LSCS), and 305-d milk yield overall and for lactation 1 to 3 (1st, 2nd, and 3rd) in data sets A and B

	A, test-day SCC records (2006–2016)				B, test-day SCC records (2014–2016)			
Trait	1st	2nd	3rd	Total	1st	2nd	3rd	Total
SCM50, %	41.6	57.6	64.5	51.6	41.5	57.8	64.7	51.8
SCM100, %	23.8	35.9	42.7	31.7	23.8	36.4	43.5	32.1
SCM150, %	15.3	24.1	29.7	21.2	15.4	24.6	30.7	21.7
SCM200, %	10.5	17.0	21.5	14.9	10.7	17.6	22.6	15.4
SCM250, %	7.5	12.5	16.1	10.9	7.7	12.9	17.1	11.4
SCM300, %	5.6	9.6	12.5	8.3	5.8	9.9	13.3	8.7
SCM350, %	4.3	7.5	9.9	6.5	4.5	7.8	10.5	6.9
SCM400, %	3.4	5.9	8.0	5.2	3.6	6.3	8.4	5.5
SCM200_3, %	3.1	5.1	6.5	4.5	4.1	6.7	8.7	5.9
SCM400_3, %	0.8	1.4	1.9	1.2	1.0	1.8	2.5	1.6
D50, d	104	106	104	105	104	105	103	104
D400, d	123	115	114	117	124	116	113	117
LSCS	4.3	4.4	4.6	4.4	4.3	4.4	4.6	4.4
Milk yield, kg	7,472	7,364	7,498	7,441	7,875	7,779	7,880	7,844

[1]Traits: SCM50, SCM100, SCM150, SCM200, SCM250, SCM300, SCM350, and SCM400 = subclinical mastitis above the threshold in 2 test days at 50,000, 100,000, 150,000, 200,000, 250,000, 300,000, 350,000, and 400,000 cells/mL, respectively; SCM200_3 and SCM400_3 = subclinical mastitis traits above the threshold of 200,000 or 400,000 in 3 test days; D50 and D400 = number of days before the first case with SCM50 and SCM400, respectively; LSCS = lactation-average somatic cell score; milk yield = milk production during 305 d.

decreased with higher SCC threshold. The mean number of days before the first case with 2 test-day SCC >50,000 or SCC >400,000 was similar across parities, with range of 103 to 106 and 113 to 124 d, respectively. Mean LSCS increased from 4.3 in 1st lactation to 4.6 in 3rd lactation, corresponding to 72,000 and 100,000 cells/mL. The mean frequency of SCM, based on SCC threshold 50,000 to 200,000, corresponded with the results reported for NR by Svendsen and Heringstad (2006). The highest frequency for SCM50 follows directly from the definition. The lowest frequency was calculated for SCM200_3 and SCM400_3 because of the requirement of 3 test days in a row above the fixed threshold; moreover, the risk of culling increase with high SCC and the frequency may be underestimated. The number of days before SCM50 and SCM400 displays small variation between lactations, but increases with higher threshold. Number of days varies from 21 to 289 (D50) and 21 to 288 (D400).

Heritability

The estimated heritabilities presented in Table 2, ranged from 0.04 (SCM400) to 0.12 (SCM100), with SE of 0.002, across SCM traits based on 2 SCC test days. For SCM200_3 and SCM400_3, based on 3 SCC test days, the heritability (SE) was 0.04 (0.002) and 0.01 (0.001), for D50 and D400 it was 0.02 (0.001) and 0.01 (0.001), respectively. For LSCS and 305-d milk yield the heritability was 0.26 (0.003) and 0.26 (0.002). The estimated heritability of SCM150 was the same as reported for Canadian Holsteins (Koeck et al., 2012). Slightly higher heritability for SCC150 (0.14) was reported for Swedish Holsteins (Urioste et al., 2012) and lower (0.09) for Italian Holsteins (Bobbo et al., 2018). Our results were in agreement with estimated heritabilities reported for SCM in NR based on test-day SCC thresholds of 50 to 200 $\times$ 10^3 cells/mL for each of the 3 first lactations, where heritability varied from 0.05 for threshold 200,000 to 0.09 for 50,000 cells/mL (Svendsen and Heringstad, 2006). For LSCS, previous estimated heritabilities for NR vary from 0.11 (Ødegård et al., 2004) to 0.17 (Haugaard et al., 2013). For other populations, such as Finnish dairy cattle (Ayrshire and Holstein-Friesian), Austrian Fleckvieh cows, and Holstein-Friesians, LSCS heritabilities of 0.19 (Luttinen and Juga, 1997) and 0.13 (de Haas et al., 2008; Koeck et al., 2010) have been reported. However, SCM traits were defined as binary and variance components estimated by linear models. Heritability estimated on binary traits from linear models depends on frequency, and results from different studies therefore cannot be compared directly. Moreover, SCM can be caused by several different bacterial species giving different SCC. Haugaard et al. (2013) showed that mastitis caused by different pathogens is not the same trait genetically and have different heritabilities. Subclinical mastitis caused by *S. aureus* had a heritability of 0.04, whereas mastitis caused by *Streptococcus uberis* or unspecific had a heritability of 0.11. The variation in the heritability found in the current study may indicate differences in the causative pathogen and association with different threshold.

Genetic Correlation

Estimated genetic correlations among the SCM traits are presented in Table 2. High genetic correlations were observed among the SCM traits, with low SE from 0.001 to 0.06. An unfavorable low or moderate genetic correlation was found to milk yield (0.07–0.34). The correlation between LSCS and milk yield was 0.07, which is in agreement with a correlation of 0.08 reported by Luttinen and Juga (1997) for Finnish dairy cattle. However, a weighted average genetic correlation between SCC and milk production of 0.14 was reported in a review (Mrode and Swanson, 1996). The high genetic correlations among SCM traits corresponded to a previous study for NR (Svendsen and Heringstad, 2006). In other studies, genetic correlations between SCM and alternative SCC traits have been reported to range from a negative favorable correlation such as −0.85 to positive 0.99 (de Haas et al., 2008; Windig et al., 2010; Urioste et al., 2012). A high genetic correlation close to 1 indicates that the same genetic mechanisms affect the traits, otherwise lower genetic correlation indicates that even if traits have some common background they are not exactly the same genetically. The high genetic correlation between SCM50 and LSCS reveals that the threshold of 50,000 cells/mL is most similar to LSCS in the NR population, the trait used in the current genetic evaluation of NR. The D50 trait, based on the same threshold of 50,000 cells/mL, showed lower genetic correlations to other alternative SCM traits with variation from 0.67 to 0.93. The threshold of 400,000 was most different from LSCS and showed the highest (moderate) unfavorable genetic correlation with milk yield. This corresponds to loss in milk production caused by SCM as reported by Hagnestam-Nielsen et al. (2009) and reviewed by Ruegg (2017).

Genetic correlation to CM was not estimated in the current study; however, a positive genetic correlation between SCC and CM was reported previously by several authors (e.g., Lund et al., 1999; de Haas et al., 2008). Svendsen and Heringstad (2006) estimated genetic correlations ranging from 0.26 to 0.62 between CM and SCM traits, and found a stronger correlation between SCM based on higher threshold of SCC and

Table 2. Genetic correlation[1] between the alternative chronic subclinical mastitis (SCM) traits,[2] lactation-average somatic cell score (LSCS) during 1 to 3 lactations, and 305-d milk yield (heritability with SE on the diagonal)

Item	Milk yield	LSCS	SCM50	SCM100	SCM150	SCM200	SCM250	SCM300	SCM350	SCM400	SCM200_3	SCM400_3	D50	D400
Milk yield	0.26 (0.002)	0.07	0.11	0.15	0.16	0.21	0.23	0.24	0.25	0.27	0.26	0.32	0.19	0.34
LSCS		0.26 (0.003)	0.99	0.98	0.98	0.96	0.95	0.95	0.94	0.92	0.96	0.89	0.88	0.92
SCM50			0.12 (0.002)	0.98	0.96	0.93	0.92	0.90	0.88	0.86	0.93	0.84	0.93	0.86
SCM100				0.12 (0.002)	0.99	0.98	0.97	0.96	0.94	0.93	0.98	0.91	0.87	0.93
SCM150					0.10 (0.002)	1.00	0.99	0.98	0.96	0.96	0.99	0.94	0.84	0.96
SCM200						0.08 (0.002)	1.00	0.99	0.98	0.97	1.00	0.97	0.79	0.97
SCM250							0.07 (0.002)	1.00	0.99	0.98	1.00	0.98	0.77	0.98
SCM300								0.06 (0.002)	1.00	0.99	0.99	0.99	0.76	0.99
SCM350									0.05 (0.002)	1.00	0.98	0.99	0.72	0.99
SCM400										0.04 (0.002)	0.97	0.99	0.69	0.99
SCM200_3											0.04 (0.002)	0.97	0.77	0.97
SCM400_3												0.01 (0.001)	0.67	0.99
D50													0.02 (0.001)	0.70
D400														0.01 (0.001)

[1]SE for genetic correlations ranged from 0.001 to 0.06.

[2]SCM traits: SCM50, SCM100, SCM150, SCM200, SCM250, SCM300, SCM350, and SCM400 = subclinical mastitis above the threshold in 2 test days at 50,000, 100,000, 150,000, 200,000, 250,000, 300,000, 350,000, and 400,000 cells/mL, respectively; SCM200_3 and SCM400_3 = subclinical mastitis traits above the threshold 200,000 and 400,000 in 3 test days; D50 and D400 = number of days before the first case with SCM50 and SCM400, respectively.

CM in late lactation. Haugaard et al. (2012, 2013) estimated genetic correlations lower than 1 between pathogen-specific CM and SCM in NR dairy cows and concluded that they should be considered as partly different traits.

Trait Definitions

The definition of the traits in the current study were based on previous genetic analyses of SCC and SCM in the NR population (Svendsen and Heringstad, 2006). Eight SCC thresholds from 50,000 to 400,000 cells/mL were evaluated. A healthy mammary gland usually has a SCC below 50,000 cells (Barbano et al., 2006), but is often defined as SCC <100,000 cells/mL (Schwarz et al., 2011). The discussion on possible effects of too low SCC in bovine milk has increased in recent years (Rainard et al., 2018). For example, Schalm et al. (1971) reported that any mid-lactation milk samples with SCC >20,000 cells/mL are a sign of inflammation. Other studies considered SCC <100,000 cells/mL to be defined as healthy quarters (Schwarz et al., 2011) and others use below 200,000 cells/mL (Schepers et al., 1997; Pantoja et al., 2009). Sarikaya et al. (2006) reported that SCC in healthy quarters consists mostly of immune cells, such as lymphocytes and macrophages, whereas in case of increased SCC all inflammatory factors will increase. Several authors concluded that alternatively defined SCC traits provide additional information that may improve genetic evaluation and selection on mastitis resistance and udder health (de Haas et al., 2008; Windig et al., 2010; Koeck et al., 2012; Bobbo et al., 2018). The SCC limit of 400,000 cells/mL in bovine milk allowed for human consumption was applied by the European Commission Milk Hygiene Directive (92/46) in 1992 and used as the upper limit in the current study. To identify chronic SCM cases, 2 and 3 subsequent SCC records were evaluated. The number of affected cows with SCC above 200,000 cells/mL was reduced from 14.9% (SCM200) to 4.5% (SCM200_3) between 2 and 3 subsequent test day records, respectively, and from 5.2% (SCM400) to 1.2% (SCM400_3) for 400,000 cells/mL. The main reason for the very low frequency of SCM with 3 test days with high SCC in the NR population is probably censoring (i.e., culling of cows with high SCM) because SCM led to reduced milk production followed by economic loss. Moreover, by doubling SCC above 50,000 cells/mL, production losses of 91 and 181 kg of milk per lactation for parity 1 and >1, respectively, will be expected (reviewed by Ruegg, 2017). Low frequency will affect the estimated parameters as mentioned above, but by requiring 3 subsequent test days with high SCC we will identify stronger cases with chronic inflammations and possibly have a more precise definition of chronic SCM cases. Traits D50 and D400 take the time aspect into account, namely how long cows can stay with SCC below 50,000 cells/mL or below 400,000 cells/mL. A higher number of days indicates better ability to resist SCM. The number of days for D50 displayed little or no variation between parities, indicating no effect of lactation number on the D50 trait. However, for D400 the number of days was lower with increased lactation number and compared with D50 were higher on average. Several of the alternative SCM traits (D50, D400, and SCM400_3) included in the current study are novel in definition. Relatively high frequency and lower correlation to the other traits indicate that these novel traits should be taken into account to improve breeding strategies against chronic SCM. However, udder health is a complex trait and further research is needed before we can make any recommendations on how to combine the many aspects of udder health in a selection index.

CONCLUSIONS

The 12 alternatively defined traits for chronic SCM display genetic variation, and the estimated genetic correlations among the traits were strong. The trait used in genetic evaluation for NR, LSCS, had the highest genetic correlation with SCM50, whereas milk yield had the strongest unfavorable genetic correlation with SCM400_3 and D400, which were the most different from the LSCS trait. Based on the genetic correlations that were lower than 1, all traits can be considered as partly different traits, which provide additional information on chronic subclinical mastitis and may be used for genetic evaluation to improve breeding for better udder health.

ACKNOWLEDGMENTS

Geno Breeding and AI Association (Hamar, Norway) and the Norwegian Dairy Herd Recording System are acknowledged for providing access to data. This work is part of the “Multimast” project (no. 233778) funded by the Research Council of Norway, Tine, and Geno and coordinated by Norwegian University of Life Sciences (NMBU).

REFERENCES

Barbano, D. M., Y. Ma, and M. V. Santos. 2006. Influence of raw milk quality on fluid milk shelf life. J. Dairy Sci. 89(Suppl 1):E15–E19. https://doi.org/10.3168/jds.S0022-0302(06)72360-8.

Begley, N., F. Buckley, K. M. Pierce, A. G. Fahey, and B. A. Mallard. 2009. Differences in udder health and immune response traits of Holstein-Friesians, Norwegian Reds, and their crosses in second lactation. J. Dairy Sci. 92:749–757. https://doi.org/10.3168/jds.2008-1356.

Bobbo, T., M. Penasa, R. Finocchiaro, G. Visentin, and M. Cassandro. 2018. Alternative somatic cell count traits exploitable in genetic selection for mastitis resistance in Italian Holsteins. J. Dairy Sci. https://doi.org/10.3168/jds.2018-14827.

de Haas, Y., W. Ouweltjes, J. ten Napel, J. J. Windig, and G. de Jong. 2008. Alternative somatic cell count traits as mastitis indicators for genetic selection. J. Dairy Sci. 91:2501–2511. https://doi.org/10.3168/jds.2007-0459.

Dohoo, I. R., and K. E. Leslie. 1991. Evaluation of changes in somatic-cell counts as indicators of new intramammary infections. Prev. Vet. Med. 10:225–237. https://doi.org/10.1016/0167-5877(91)90006-N.

Hagnestam-Nielsen, C., U. Emanuelson, B. Berglund, and E. Strandberg. 2009. Relationship between somatic cell count and milk yield in different stages of lactation. J. Dairy Sci. 92:3124–3133. https://doi.org/10.3168/jds.2008-1719.

Harmon, R. J. 1994. Symposium - Mastitis and genetic evaluation for somatic cell count - Physiology of mastitis and factors affecting somatic cell counts. J. Dairy Sci. 77:2103–2112. https://doi.org/10.3168/jds.S0022-0302(94)77153-8.

Haugaard, K., B. Heringstad, and A. C. Whist. 2012. Genetic analysis of pathogen-specific clinical mastitis in Norwegian Red cows. J. Dairy Sci. 95:1545–1551. https://doi.org/10.3168/jds.2011-4522.

Haugaard, K., B. Heringstad, and A. C. Whist. 2013. Genetic associations between somatic cell score and pathogen-specific subclinical mastitis in Norwegian Red cows. J. Anim. Breed. Genet. 130:98–105. https://doi.org/10.1111/jbg.12019.

Koeck, A., B. Heringstad, C. Egger-Danner, C. Fuerst, P. Winter, and B. Fuerst-Waltl. 2010. Genetic analysis of clinical mastitis and somatic cell count traits in Austrian Fleckvieh cows. J. Dairy Sci. 93:5987–5995. https://doi.org/10.3168/jds.2010-3451.

Koeck, A., F. Miglior, D. F. Kelton, and F. S. Schenkel. 2012. Alternative somatic cell count traits to improve mastitis resistance in Canadian Holsteins. J. Dairy Sci. 95:432–439. https://doi.org/10.3168/jds.2011-4731.

Lund, M. S., J. Jensen, and P. H. Petersen. 1999. Estimation of genetic and phenotypic parameters for clinical mastitis, somatic cell production deviance, and protein yield in dairy cattle using Gibbs sampling. J. Dairy Sci. 82:1045–1051. https://doi.org/10.3168/jds.S0022-0302(99)75325-7.

Luttinen, A., and J. Juga. 1997. Genetic relationships between milk yield, somatic cell count, mastitis, milkability and leakage in Finish dairy cattle population. Interbull Bull. 15:78–83.

Madsen, P. 2012. DmuTrace. A program for extracting the pedigree for a subset of animals from a larger pedigree. Version 2. Aarhus University, Aarhus, Denmark.

Madsen, P., and J. Jensen. 2013. A User's Guide to DMU. A Package for Analysing Multivariate Mixed Models. Version 6, release 5.2. Aarhus University, Aarhus, Denmark.

Miller, R. H., M. J. Paape, and L. A. Fulton. 1991. Variation in milk somatic cells of heifers at first calving. J. Dairy Sci. 74:3782–3790. https://doi.org/10.3168/jds.S0022-0302(91)78570-6.

Mrode, R., and G. Swanson. 1996. Genetic and statistical properties of somatic cell count and its suitability as an indirect means of reducing the incidence of mastitis in dairy cattle. Animal Breeding Abstracts (United Kingdom).

Ødegård,, J., B. Heringstad, and G. Klemetsdal. 2004. Short communication: bivariate genetic analysis of clinical mastitis and somatic cell count in Norwegian dairy cattle. J. Dairy Sci. 87:3515–3517. https://doi.org/10.3168/jds.S0022-0302(04)73487-6.

Østerås, O., H. Solbu, A. O. Refsdal, T. Roalkvam, O. Filseth, and A. Minsaas. 2007. Results and evaluation of thirty years of health recordings in the Norwegian dairy cattle population. J. Dairy Sci. 90:4483–4497. https://doi.org/10.3168/jds.2007-0030.

Pantoja, J. C., C. Hulland, and P. L. Ruegg. 2009. Dynamics of somatic cell counts and intramammary infections across the dry period. Prev. Vet. Med. 90:43–54. https://doi.org/10.1016/j.prevetmed.2009.03.012.

Rainard, P., G. Foucras, D. Boichard, and R. Rupp. 2018. Invited review: Low milk somatic cell count and susceptibility to mastitis. J. Dairy Sci. 101:6703–6714. https://doi.org/10.3168/jds.2018-14593.

Ruegg, P. L. 2017. A 100-Year Review: Mastitis detection, management, and prevention. J. Dairy Sci. 100:10381–10397. https://doi.org/10.3168/jds.2017-13023.

Sarikaya, H., G. Schlamberger, H. H. Meyer, and R. M. Bruckmaier. 2006. Leukocyte populations and mRNA expression of inflammatory factors in quarter milk fractions at different somatic cell score levels in dairy cows. J. Dairy Sci. 89:2479–2486. https://doi.org/10.3168/jds.S0022-0302(06)72322-0.

Schalm, O. W., E. J. Carroll, and N. C. Jain. 1971. Number and types of somatic cells in normal and mastitic milk. Pages 94–127 in Bovine Mastitis. Lea & Febiger, Philadelphia, PA.

Schepers, A. J., T. J. Lam, Y. H. Schukken, J. B. Wilmink, and W. J. Hanekamp. 1997. Estimation of variance components for somatic cell counts to determine thresholds for uninfected quarters. J. Dairy Sci. 80:1833–1840. https://doi.org/10.3168/jds.S0022-0302(97)76118-6.

Schukken, Y. H., K. E. Leslie, A. J. Weersink, and S. W. Martin. 1992. Ontario bulk milk somatic cell count reduction program. 2. Dynamics of bulk milk somatic cell counts. J. Dairy Sci. 75:3359–3366. https://doi.org/10.3168/jds.S0022-0302(92)78112-0.

Schwarz, D., U. S. Diesterbeck, S. Konig, K. Brugemann, K. Schlez, M. Zschock, W. Wolter, and C. P. Czerny. 2011. Microscopic differential cell counts in milk for the evaluation of inflammatory reactions in clinically healthy and subclinically infected bovine mammary glands. J. Dairy Res. 78:448–455. https://doi.org/10.1017/S0022029911000574.

Svendsen, M., and B. Heringstad. 2006. Somatic cell count as an indicator of sub-clinical mastitis. Genetic parameters and correlations with clinical mastitis. Interbull Bull. 35:12–16.

TINE. 2017. TINE Rådgivning. Statistikksamling 2017 [TINE Annual Statistics 2017]. Accessed October 12, 2018. https://medlem.tine.no/fagprat/helse/_attachment/427411?_ts=16167bec8ab. (in Norwegian, with English Tables).

Urioste, J. I., J. Franzen, J. J. Windig, and E. Strandberg. 2012. Genetic relationships among mastitis and alternative somatic cell count traits in the first 3 lactations of Swedish Holsteins. J. Dairy Sci. 95:3428–3434. https://doi.org/10.3168/jds.2011-4739.

Windig, J. J., W. Ouweltjes, J. Ten Napel, G. de Jong, R. F. Veerkamp, and Y. De Haas. 2010. Combining somatic cell count traits for optimal selection against mastitis. J. Dairy Sci. 93:1690–1701. https://doi.org/10.3168/jds.2009-2052.

Paper II

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12886

Identification of candidate genes affecting chronic subclinical mastitis in Norwegian Red cattle: combining genome-wide association study, topologically associated domains and pathway enrichment analysis

E. Kirsanova*, B. Heringstad†,‡, A. Lewandowska-Sabat* and I. Olsaker*

*Department of Basic Sciences and Aquatic Medicine, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Oslo, Norway. †Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, Oslo, Norway. ‡Geno Breeding and A.I. Association, Hamar, Norway.

Summary

The aim of this study was to identify genes associated with chronic subclinical mastitis (SCM) in Norwegian Red (NR) cattle. Twelve SCM traits defined based on fixed threshold for test-day somatic cell count (SCC) were, together with lactation-average somatic cell score (LSCS) used for association and pathway enrichment analyses. A GWAS was performed on 3795 genotyped NR bulls with 777K SNP data and phenotypic information from 7 300 847 test-day SCC observations from 3 543 764 cows. At 5% chromosome-wide significance level 36 unique SNP were detected to be associated with one or more of the traits. These SNPs were analysed for linked genes using genomic positions of topologically associated domains (TAD). For the SCM traits with SCC >50 000 and >100 000 cells/ml on two test-days in a row and LSCS, the same top significant genes were identified – checkpoint clamp loader component (*RAD17*) and cyclin B1 (*CCNB1*). The SCM traits with SCC >250 000, 300 000, 350 000 or 400 000 cells/ml on two test-days in a row and D400 (number of days before the first case with SCC >400 000 cells/ml) displayed similar top significant genes: acyl-CoA thioesterase 2 and 4 (*ACOT2*; *ACOT4*). For the traits SCM200_3 (SCC >200 000 cells/ml on three test-days in a row) and SCM150, SCM200 (SCC >150 000; 200 000 cells/ml on two test-days in a row) a group of chemokine (C–X–C motif) ligand genes and the Fos proto-oncogene, AP-1 transcription factor subunit (*FOS*) gene, were identified. Further functional studies of these identified candidate genes are necessary to clarify their actual role in development of chronic SCM in NR cattle.

Keywords genome-wide association study, somatic cell count

Introduction

Chronic subclinical mastitis (SCM) is a complex trait characterised by high somatic cell count (SCC) for a prolonged period of time, often with SCC> 200 000 cells/ml (Harmon, 1994). In Norway, SCC has been recorded since the 1970s in the Norwegian Dairy Herd Recording System (Østerås *et al.* 2007). The data have been used in genetic analyses (e.g. Ødegård *et al.* 2003; Ødegård *et al.* 2004; Svendsen & Heringstad 2006; Haugaard *et al.* 2013) and associations studies (e.g. Sodeland *et al.* 2011; Olsen *et al.* 2016). GWASs (Delvin & Risch 1995) have been widely used for identification of chromosomal regions and for mapping of functional traits such as SCC (*e.g.* Kuhn *et al.* 2003; Sahana *et al.* 2010; Sodeland *et al.* 2011) and clinical mastitis (CM) in dairy cattle (*e.g.* Sahana *et al.* 2013; Wu *et al.* 2015; Olsen *et al.* 2016). In Norwegian Red (NR) cattle, significant QTL that affect somatic cell score have been reported on bovine chromosome (BTA–*Bos Taurus* autosome) 12, 19 and 26 (Sodeland *et al.* 2011), and for NR cattle on BTA5, -6, -13, -16, -19 and -20 (Sahana *et al.* 2014). GWAS combined with pathway-based analysis is often used to understand biological functions involved in complex traits. There are several methods used to select the genes for pathway analysis. Using topologically associated

Address of correspondence

E. Kirsanova, Department of Basic Sciences and Aquatic Medicine, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Oslo, Norway.
E-mail: elena.kirsanova@nmbu.no

Accepted for publication 10 November 2019

domains (TADs; Dixon *et al.* 2012), representing functional regulatory regions of the genome can improve the search for causative variants compared with using arbitrarily chosen distances from significant SNPs (Wang *et al.* 2018). TADs are highly conserved domains with self-interacting chromatin, and constitute one type of functional annotation of the regulatory structure of the genome (Krefting *et al.* 2018). TAD regions are known to be stable during replication (Pope *et al.* 2014) and they are conserved between different cell types and between species (Dixon *et al.* 2012; Rao *et al.* 2014; Rudan *et al.* 2015). The stability and impact on gene regulation makes TADs an important tool for identification of underlying gene regulation and mechanisms of a trait or disease. In the bovine genome, TADs were first described and analysed by Wang *et al.* (2018). TAD coordinates from available humans, mice, dogs and macaques from several cell types were converted to the bovine genome with 79.63–98.88% mapped. TADs at least 200 kb wide were included to avoid small homological genomic fragments incapable of forming TADs in the bovine genome. Coordinates for TAD regions based on bovine sequence data are not yet available (September 2019). However, the analysis by Wang *et al* (2018) showed that the liftover TAD regions provide useful information about causative regulatory variants in the bovine genome and constitute a sensible search space.

The aim of this study was to identify genes with potential effect on chronic SCM in NR cattle using GWAS on 777K SNP data and the alternative SCM traits described by Kirsanova *et al.* (2019) and subsequently perform enrichment analysis using ingenuity pathway analysis (IPA) with genes identified within TAD regions.

Materials and methods

Phenotypic data

Phenotypic test-day SCC records from 1979 to 2016 for NR cows were obtained from the national Norwegian Dairy Herd Recording System and edited using SAS version 9.4 (SAS Inst. Inc., Cary, NC, USA; www.sas.com). Lactations 1–3 with two or more test-days in a row within a 2-month period, from 3 543 764 cows with 7 300 847 observations, were analysed. The DMU trace (Madsen 2012) was used to build the corresponding pedigree file of 4 126 678 animals.

Traits and statistical analyses

Kirsanova *et al.* (2019) defined 12 alternative SCM traits, and estimated genetic parameters for these together with lactation-average somatic cell score (LSCS). Briefly, there were eight binary SCM traits defined based on whether or not two test-day SCCs in a row were above the given SCC threshold, with threshold from 50 000 to 400 000 cells/ml (SCM50, -100, -150, -200, -250, -300, -350 and -400); two similar traits were based on three test-day SCC records in a row above 200 000 and 400 000 cells/ml respectively (SCM200_3 and SCM400_3); and finally, two traits were defined as the number of days before the first case with SCM50 or SCM400 (D50 and D400). All of the defined traits were within the days in-milk period between 21 and 305 days after calving. Linear animal repeatability models with estimated variance components from Kirsanova *et al.* (2019) were used for analyses in the DMU5 procedure in the DMU (Madsen & Jensen 2013). Solutions for fixed and random effects were used in calculations of daughter-yield-deviation (DYD; VanRaden & Wiggans 1991) for NR sires. DYDs are daughter phenotypes corrected for fixed and non-genetic random effects. The following model was used to calculate yield deviation (YD) for each cow for each trait:

$$\begin{aligned} \mathrm{YD} = {} & \mathrm{trait} - \mathrm{year\_month} - \mathrm{age} - \mathrm{days\_open} \\ & - \mathrm{herd\_year} - pe - 0.5\,\mathrm{EBV_{mor}} \end{aligned}$$

where year_month is calving year/month; age is age at calving in month by lactation number; and days_open is days open grouped (each 10 days for each lactation number). Non-genetic random effects (herd_year, herd year of calving; *pe*, permanent environment effect) and 0.5 of dam's estimated breeding value ($\mathrm{EBV_{mor}}$) were used to correct for genetic effect of dam in the model. All the fixed and random effects have been described by Kirsanova *et al.* (2019). The DYDs for sires were calculated as the arithmetic mean of the daughter's YD.

Genotypes

Sires were genotyped with the Affymetrix 25K SNP array (Affymetrix, Santa Clara, CA, USA), or the Illumina BovineSNP50 BeadChip (54K) (Illumina Inc., San Diego, CA, USA; Matukumalli *et al.* 2009) and combined with the Illumina BovineHD Genotyping BeadChip (777K). Geno SA (www.geno.no), the breeding organisation for NR cattle, performed imputation of missing genotypes. Genotypes from 3795 NR bulls with 613 908 SNP located on 29 chromosomes passed the quality control. A MAF of 1%, a minimum call rate of 95% and an individual call rate of 85% were used. The genome assembly *Bos taurus* UMD 3.1.1 (Zimin *et al.* 2009) was used for location of the SNP positions.

GWAS

GWAS analyses were performed using the R version 3.4.1 package GenABEL version 1.8-0 (http://www.r-project.org/; Aulchenko *et al.* 2007). DYDs were used as phenotypes in the GWAS. GRAMMA-Gamma function, a variance component-based two-step method, was used for unbiased estimation of SNP effects (Svischeva *et al.* 2012). The polygenic function included the kinship relationship matrix calculated and provided by Geno SA, based on 35 652 genotyped SNPs. The *P*-values were corrected by dividing

the observed test statistic by genomic inflation factor estimated in GenABEL. Chromosome-wide significance level was defined according to Sahana *et al.* (2010). The 5% chromosome-wide significance threshold ranged from the point-wise *P*-value of 1.29×10^{-6} on BTA1 to 4.31×10^{-6} on BTA25. Therefore *P*-values of $<5 \times 10^{-6}$ were required for significance. To create Manhattan plots the qqman package version 0.1.4 was used (Turner 2014).

TAD and pathway analysis

Genomic TAD coordinates from mouse cortex (Shen *et al.* 2012) and embryonic stem cells in the mm9 genome (Dixon *et al.* 2012) in addition to TAD coordinates from human embryonic stem cells and foetal lung IMR90 fibroblasts (IMR90) in the hg18 (Dixon *et al.* 2012) were obtained from Bing Ren's Lab, University of California, San Diego (downloaded from: http://chromosome.sdsc.edu/mouse/hi-c/download.html). Using the UCSC Batch Conversion program (liftOver tool; Kuhn *et al.* 2013) provided on the UCSC Genome Browser page (http://genome.ucsc.edu/), all of the TAD coordinates were converted (default settings, as described by Wang *et al.* 2018) to the bovine reference genome *Bos taurus* UMD3.1.1 (UCSC Genome Browser assembly ID: bosTau8). The genome version UMD3.1.1 has more correctly closed gaps and 173 contigs defined as contaminants and suppressed, compared with the version UMD3.1 used in Wang *et al.* 2018. Both human and mouse genomes were converted from hg18 and mm9 to the bosTau8 (NCBI Assembly ID 189361) through hg38 (NCBI Assembly ID 5800238) and mm10 (NCBI Assembly ID 327618) respectively. Bedtools coverage version 2.29.0 (Quinlan & Hall 2010) was used to compute the percentage of converted TADs that covered the reference genome *Bos Taurus* UMD3.1.1 (presented in Supplementary Appendix S1). Additional summary statistics of TAD mapping are described in Table 1. The length of the bovine TADs varied from 275 805 to 5 775 052 bp and from 271 661 to 6 763 506 bp, converted from the human and mouse genome respectively. TAD coordinates with at least 200 kb length (Wang *et al.* 2018) that matched converted TAD from the human and mouse genomes for all types of tissue were defined as stable and used for further analysis. For two SNP no annotated genes were found within the TAD regions containing the SNP. In order to include these in further analyses close TAD regions were used. For trait SCM400_3, SNP BTA2:84213972, BovineHD0200024041, the TAD region was located 100 158 bp from the SNP and for trait SCM50, SNP BTA20:10493654, BovineHD2000003312, the TADs were located 475 073 bp from the SNP. The number of genes detected using the UCSC Genome Browser across the defined traits ranged from 11 to 28 and from 4 to 76, based on the human and mouse TAD position respectively.

The list of genes for each trait was analysed using IPA (https://www.qiagenbioinformatics.com/). IPA is a database based on human gene ID, presenting top canonical pathways, upstream regulators and functional networks. In IPA, Fisher's extract-test at 5% level was used for calculation of significant *P*-values.

Results

A total of 36 significant SNPs across all 13 traits were identified by GWAS; 20 of these were unique for only one trait, whereas 16 were significantly associated with more than one trait (Table 2). Significant SNPs for SCM in NR cattle were found on BTA2, -5, -6, -7, -10, -12, -20 and -26 for the alternative SCM traits, whereas for LSCS there were found only on BTA20. The highest number of SNPs was observed on BTA20 (12 SNPs), followed by BTA12 (eight SNPs), BTA26 (six SNPs) and BTA10 (five SNPs). Only one or two SNPs were identified on BTA2, -5, -6 and -7. The SNPs that showed significant association to most traits was BovineHD1000025012 on BTA10, affecting eight of the 13 traits, followed by BovineHD2600010426 on BTA26, affecting seven traits. The rest of the identified SNPs showed significant association with one to six traits. SNPs located on BTA10 affected 10 of the analysed traits, followed by nine on BTA26 and eight on BTA12 and BTA20. Manhattan plots for each trait at 5 and 1% significant chromosome-wide thresholds are presented in Fig. 1.

In total 181 unique genes were identified using the bovine TAD approach. Many genes were associated with more than one SCM trait, and the number of identified associated genes ranged from 15 (LSCS) to 100 (SCM150; SCM200) per trait. Subsequently IPA mapped 168 genes. The genes found by IPA in the top five significant pathways for each trait are presented in Table 3. Additional information on the top five canonical pathways with upstream regulators, *P*-values and gene ID for all the traits are presented in Table S1 and Table S2.

Some of the analysed traits showed highly overlapping IPA results, with the same genes and pathways identified. The results for SCM150 were identical to those for SCM200, and so were results for SCM250 and SCM300. For five traits (SCM250, SCM300, SCM350, SCM400 and D400) acyl coenzyme A (CoA) hydrolysis, part of the metabolism of fatty acids, was identified as the top significant pathway involving the acyl-CoA thioesterase 2 (*ACOT2*) and acyl-CoA thioesterase 4 (*ACOT4*) genes. For three traits (SCM50, SCM100 and LSCS) DNA damage-induced 14-3-3 σ signalling, responsible for control of the biological activity in the cell cycle, was identified as the top pathway involving the checkpoint clamp loader component enzyme (*RAD17*) and cyclin B1 kinase (*CCNB1*) genes. Retinoic acid receptor activation, which may act as transcription factor, was one of the genes in the top pathway detected for SCM150/SCM200. For the SCM200_3 the top pathway was a role of interleukin (IL) 17A in psoriasis with four significant genes from the chemokines family. For the last two traits,

Table 1 Topologically associated domains (TAD) statistics mapped to bovine reference genome *Bos taurus* UMD3.1.1.

Genome	Tissue	Reference assembly	Number of TADs	Percentage of TADs converted successfully	Percentage of the UMD3.1.1 covered by converted TADs
Human[1]	hESC[2]	hg18	3127		
		bostau8	2914	93.2 %	89.1 %
	IMR90[3]	hg18	2349		
		bostau8	2174	92.6 %	87.6 %
Mouse[1]	mESC[4]	mm9	2200		
		bostau8	1893	86.0 %	80.7 %
	Cortex	mm9	1519		
		bostau8	1319	86.8 %	77.4 %

The cell and tissue with input data on TADs from human and mouse genomes were converted to the latest version of the bovine reference genome *Bos taurus* UMD3.1.1. The number and percentage of TADs successfully converted to the bovine genome and percentage of the genome covered by converted TADs are shown.
[1]Human/mouse – data from study by Dixon *et al.* (2012).
[2]hESC – human embryonic stem cells.
[3]IMR90 – foetal lung IMR90 fibroblasts.
[4]mESC – mouse embryonic stem cells.

SCM400_3 and D50, only one gene for each trait was identified within any canonical pathways, hence no significant pathways could be assigned (Table 3 with corresponding abbreviations in List S1). The top networks for the traits (presented with score in Table S3) were mainly associated with developmental disorders (SCM150/SCM200, SCM350), abnormalities (SCM250/SCM300, D50, D400, SCM400) and cellular growth (LSCS) in addition to cell death and survival (SCM50, SCM100) and cell-to-cell signalling (SCM200_3).

Discussion

Associated SNP positions

QTL located on eight different chromosomes were identified across the 13 analysed SCM traits. Other authors reported QTL for SCC at the same chromosomes BTA5, -6, -7, -10, -12, -20 and -26 for different breeds (Klungland *et al.* 2001; Kuhn *et al.* 2003; Sahana *et al.* 2010; Sodeland *et al.* 2011; Meredith *et al.* 2012; Meredith *et al.* 2013; Raven *et al.* 2014; Sahana *et al.* 2014). However, even if QTL on the same chromosome were reported in other studies and breeds, positions were different in NR cattle. Possible reasons include slightly different traits, different breeds and populations. To the best of our knowledge, only the previous GWAS analysis by Kirsanova *et al.* (2018) for the equally defined SCM50–SCM400 traits based on 35 605 SNP showed similar results. Some false-negative associations are also expected, even if a significant level applied (Korte & Farlow 2013). The GRAMMAR-Gamma function used in this study is a mixed-model-based method that is supposed to estimate unbiased SNP effects (Svischeva *et al.* 2012). Ekine *et al.* (2014) reported the GRAMMAR function to have the lowest type 1 error rate compared with other methods used for GWASs. That method can be recognised as conservative, and the false-positive rate in this case was lower than 0.04.

Analyses of biological processes and pathways underlying the traits

The genes identified in this study were based on TAD positions converted to bovine from mouse and human. Owing to the limited amount of available data and few detected significant genes, TAD data from both these two species were combined. The first study describing TAD positions successfully converted from several species to the bovine genome was Wang *et al.* (2018). Moreover, Wang *et al.* (2018) concluded that TAD positions can be used as a search space for causative regulatory variants and also around significant SNPs. In the current study TAD positions from the human genome mapped better to the bovine genome compared with those from mouse, in agreement with Wang *et al.* (2018).

In Norway *Staphylococcus aureus* is the major bacterium causing SCM (TINE SA 2017). This bacterium is Gram-positive and the cell wall components – lipoteichoic acid or peptidoglycan – are assumed to trigger activation of immune response through expression of interleukin 8 (IL-8; Bannerman *et al.* 2004). In the current study, *IL-8* or C–X–C motif chemokine ligand 8 also known as *CXCL8* was identified as an associated gene for the SCM200_3 trait in all of the five top canonical pathways (Table S1). Moreover, *IL-8* is located within a TAD region and 375 132 bp from a significant SNP identified by GWAS. Also, IPA has defined the *IL-1* gene as a one of the central genes in the top networks for SCM200_3 (Table S3). Moreover, *CXCL1*, *CXCL3* and *CXCL5* genes in addition to *CXCL8* were found for the SNPs placed on BTA6 within the TAD region. The top canonical pathway involving those genes was a role of *IL-17A* in psoriasis. *IL-17A* is a pro-inflammatory cytokine,

Table 2 Significant SNP detected by GWAS for alternative subclinical mastitis (SCM) traits[1].

BTA[2]	*N*[3]	SNP	Position (bp)[4]	Minor allele A1[5]	Major allele A2[6]	Trait – see text for further explanations
2	1	BovineHD0200021808	75996043	C	T	SCM400_3
2	2	BovineHD0200024041	84213972	G	A	SCM400_3
5	3	BovineHD0500024021	84931066	G	A	SCM350
6	4	BovineHD0600024697	90184750	T	C	SCM200_3
7	5	BovineHD0700010862	37686316	A	G	D400
10	6	BovineHD1000025012	87958006	T	C	SCM150, SCM200, SCM250, SCM300, SCM350, SCM400, SCM200_3, D400
10	7	BovineHD1000025014	87970294	A	G	SCM400, SCM200_3, D400
10	8	BovineHD1000027284	94432205	T	C	SCM200_3
10	9	ARS-BFGL-NGS-54789	94425952	A	G	SCM200_3
10	10	BovineHD1000027289	94462757	T	C	SCM100, SCM200_3, SCM400_3
12	11	BovineHD1200027006	54131183	T	C	SCM100, SCM150, SCM200, SCM250, SCM300, SCM200_3
12	12	BovineHD1200014960	54126754	C	T	SCM100, SCM150, SCM200, SCM300, SCM200_3
12	13	BovineHD1200014962	54143111	C	T	SCM100, SCM150, SCM200, SCM300, SCM200_3
12	14	BovineHD4100009623	54124043	A	G	SCM100, SCM150, SCM200, SCM200_3
12	15	BovineHD1200006048	20069484	A	C	SCM50, D50
12	16	BovineHD4100009623	19138863	G	A	D50
12	17	BovineHD1200005799	19142767	G	T	D50
12	18	BovineHD1200006023	19980842	C	T	SCM50
20	19	BovineHD2000003312	10493654	C	T	SCM50, SCM100, SCM150, LSCS
20	20	BovineHD2000003308	10486199	A	G	SCM50, SCM100, SCM150, SCM200, LSCS
20	21	BTB-00772821	10486993	C	T	SCM50, SCM100, SCM150, SCM200, LSCS
20	22	BovineHD2000003309	10488085	C	T	SCM50, SCM100, SCM150, SCM200, LSCS
20	23	BovineHD2000003310	10489403	A	G	SCM50, SCM100, SCM150, SCM200, LSCS
20	24	BTB-00772795	10435385	A	C	SCM50, SCM100
20	25	BovineHD2000003311	10491752	T	C	SCM50, SCM100, SCM150, SCM200
20	26	BovineHD4100014652	32962559	A	C	D50
20	27	BovineHD2000009441	32966444	C	T	D50
20	28	BovineHD2000017367	61885151	C	A	SCM400_3
20	29	BovineHD2000017376	61894720	A	G	SCM400_3
20	30	BovineHD2000017384	61902955	T	C	SCM400_3
26	31	BovineHD2600010426	37927026	G	A	SCM100, SCM150, SCM200, SCM250, SCM300, SCM350, SCM400
26	32	BovineHD2600010455	38009502	G	A	SCM400_3
26	33	BovineHD2600010690	39029294	T	C	SCM400_3
26	34	BovineHD2600012556	44579917	G	A	D400
26	35	BovineHD2600012565	44604386	T	C	D400
26	36	BovineHD2600012933	45869170	T	C	D400

[1]Traits – SCM50, -100, -150, -200, -250, -300, -350 and -400, subclinical mastitis above the threshold on two test-days at 50 000, 100 000, 150 000, 200 000, 250 000, 300 000, 350 000 and 400 000 cells/ml respectively; SCM200_3 and SCM400_3, subclinical mastitis traits above the threshold 200 000 and 400 000 on three test-days; D50 and D400, number of days before the first case with SCM50 and SCM400 respectively. LSCS, Lactation-average somatic cell score during 1–3 lactations.
[2]BTA, *Bos Taurus* autosome.
[3]*N*, Number of significant SNP.
[4]Position (bp) –position in base pair.
[5]A1, Minor allele.
[6]A2, Major allele.

known to take part in human chronic skin inflammation, contributing to the inflow of neutrophils, dendritic cells and memory T-cells (Nograles *et al.* 2008; Pfaff *et al.* 2017). This corresponds to the activation of CXC chemokines for SCM200_3, involved in the recruitment of neutrophils to the inflamed or infected tissue (Kobayashi 2006). The *IL-8* gene corresponds with the assumed activation of immune response by the bacterium as described by Bannerman *et al.* (2004). *CXCL1*, also known as *GRO1*, has previously been found to be induced in response to *S. aureus*, together with *CXCL2*, *CXCL3* and *CXCL8* (Rainard *et al.* 2008; Bougarn *et al.* 2010). The last associated gene, *CXCL5*, was previously identified in a microarray analysis by Pareek *et al.* (2005) as an important gene in initiating the innate immune response to *Escherichia coli* infection in bovine mammary epithelial cells. Additionally, for the SCM200_3 trait, *IL-1* was found to be associated. *IL-1* is an important receptor on macrophages, providing signalling for initiation of immunity and the inflammation processes.

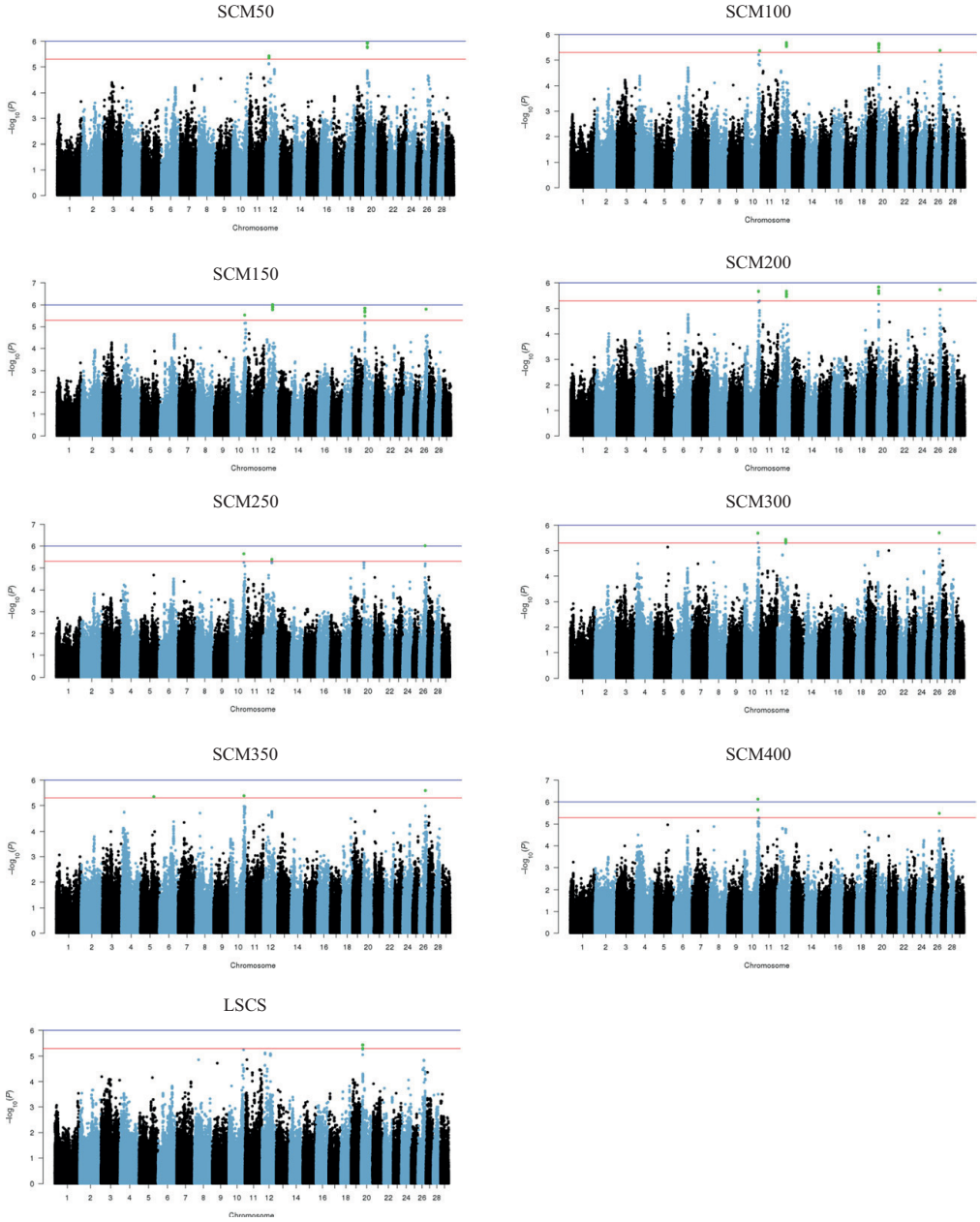


Figure 1 Manhattan plots from GWAS for alternative subclinical mastitis traits showing significant SNPs with genome-wide corrected *P*-values. The red horizontal line indicates defined chromosome-wide significance threshold on −log10(*P*) = 5.30103 ($P = 5 \times 10^{-6}$) at 5%. The blue horizontal line indicates threshold on −log10(*P*) = 6.0 ($P = 1 \times 10^{-6}$) at 1%. Number of significant SNPs – see Table 1 for further explanations.

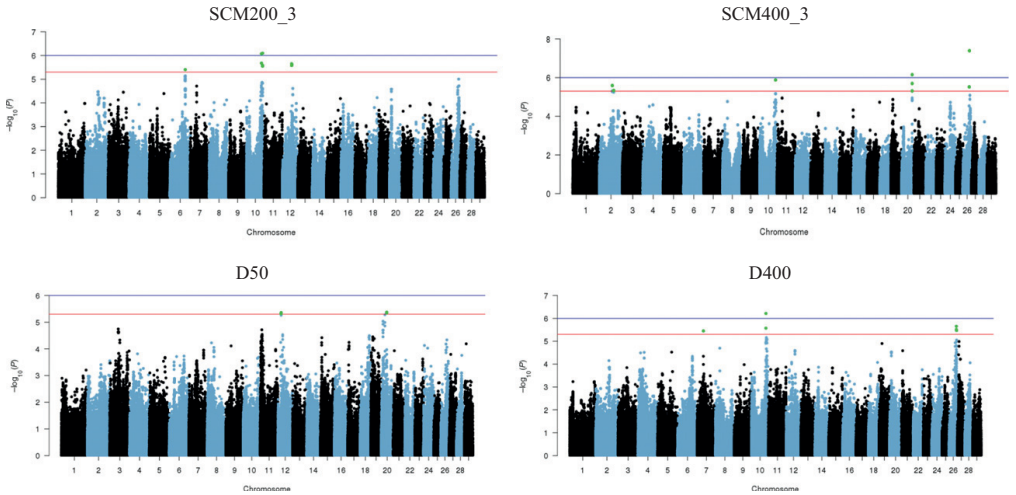


Figure 1 (Continued)

Interesting results were obtained for almost all of the alternative SCM traits with the thresholds higher than 250 000 cells/ml, which definitely indicate SCM. Acyl-CoA, a main associated canonical pathway for five of the analysed traits (SCM250, SCM300, SCM350, SCM400 and D400) involves the *ACOT2* and *ACOT4* genes, both placed 2.5 Mb from the significant SNP on BTA10. These genes, found to have an important role in bovine milk fat synthesis regulation during lactation (Strillacci *et al.* 2014), may be associated with decreased milk production in chronic SCM cases. For other traits with SCC threshold lower than 100 000 cells/ml (SCM50 and SCM100) and for LSCS, the canonical pathway of DNA damage-induced 14-3-3 σ signalling appears with two associated genes, *RAD17* and *CCNB1*, on BTA20 at 193 and 52 kb respectively, away from the significant SNP. Both genes are involved in regulation of the cell cycle process (Griffiths *et al.* 1995). *RAD17* is required for repair in response to DNA damage, which is important for correct duplication of DNA before cell division, whereas activation of the *CCNB1* gene decides that mitosis will be activated (http://www.ncbi.nlm.nih.gov/). Moreover, DNA repair genes are known to be involved in the inflammatory response during infections (Roth *et al.* 2014; Spanou *et al.* 2017). The phosphoinositide 3-kinase (*PI3K*) complex is a central molecule in the networks identified to be associated to LSCS, SCM50, SCM100 and SCM350. *PI3K* is most known in association with cancer, but it is a part of cellular function, e.g. cell growth, survival, motility and migration, which also corresponds to the increasing number of cells during udder inflammation (Harmon 1994). Nuclear factor kappa B (*NFkB*) was shown to be one of the central molecules in the identified networks for SCM100 as well as for SCM400_3 and D50. *NFkB* is a part of the classic macrophages activated by *S. aureus* (Lewandowska-Sabat *et al.* 2013). By entering the macrophages the bacterium initiates activation of *NFkB*, which activates the pro-inflammatory cytokines and the following inflammation response. Boutet *et al.* (2007) confirmed that inflammatory responses in bovine mammary epithelial cells occur via *NFkB* activation. For the traits SCM250, SCM400, SCM200_3 and D400, the most central molecules were found to be extracellular signal-regulated kinase 1 and 2 (*ERK 1/2*). *ERK* are members of MAPK1 (*ERK2*) and MAPK3 (*ERK1*), which are involved in regulation of the *NFkB* (complex) and *FOS* genes, in addition to having a role in survival, migration and cell death. Yonezawa *et al.* (2008) reported the importance of *ERK 1/2* genes in combination with *Akt* gene (serine/threonine kinase) in the bovine mammary epithelial cells via unsaturated fatty acid proliferation. In the current study, the *Akt* gene was also identified in one of the networks for the SCM200_3 and SCM250 traits as well as in a combination of *Akt*, *FOS* and *NFkB* as central genes in the network detected by IPA for D400. Moreover, *ERK 1/2* is activated by *S. aureus* infection (Ellington *et al.* 2001). Finally, for SCM150/200 the *FOS* gene appears as a main component of the top network and as one of the associated genes for SCM250/300, SCM350, SCM400, SCM200_3 and D400. *FOS* is placed at 86 887 kb on BTA10, which is more than 1 Mb from the significant SNP position (87 958 kb). The *FOS* gene is a transcription factor, and has previously been found in several Ingenuity canonical pathways associated with *S. aureus* infection in macrophages (Lewandowska-Sabat *et al.* 2013). This gene was significant for

Table 3 Alternative subclinical mastitis (SCM) traits[1] with detected associated genes by ingenuity pathway analysis.

N	Trait	Gene ID
1	SCM50	RAD17, CCNB1, TAF9, CDK7, KPNA3, PIK3R1
2	SCM100	RAD17, CDK7, CCNB1, TAF9, GRK5
3[2]	SCM150/200	FOS, PIK3R1, CDK7, TGFB3, SNW1, ACOT2, ACOT4, RAD17, CCNB1, ARHGEF12, PSEN1
4[2]	SCM250/300	ACOT2, ACOT4, NUMB, PSEN1, FOS, TGFB3, ALDH6A1, EDNRB, PGF
5	SCM350	ACOT2, ACOT4, FOS, ITPR2, KRAS, TGFB3, PGF
8	SCM400	ACOT2, ACOT4, NUMB, PSEN1, FOS, TGFB3, ALDH6A1, SNW1
9	LSCS	RAD17, CCNB1, CDK7, TAF9, PIK3R1
10	SCM200_3	CXCL8, CXCL3, CXCL1, CXCL5, CXCL2, PF4, FOS
11	SCM400_3	GLI2, SLC18A2, SEL1L, GLI2, STAT4
12	D50	NADP
13	D400	ACOT2, ACOT4, NUMB, PSEN1, FOS, TGFB3, ALDH6A1, SNW1

[1]Traits – SCM50, -100, -150, -200, -250, -300, -350, -400, subclinical mastitis above the threshold in two test-days at 50 000, 100 000, 150 000, 200 000, 250 000, 300 000, 350 000 and 400 000 cells/ml respectively; LSCS, lactation-average somatic cell score during 1–3 lactations; SCM200_3 and SCM400_3, subclinical mastitis traits above the threshold 200 000 and 400 000 on three test-days; D50 and D400, number of days before the first case with SCM50 and SCM400 respectively.
[2]3/4, Traits with identical results.

eight of the 13 analysed traits in the current study, which makes it a promising candidate involved in regulation of chronic SCM.

However, IPA is a database primary designed for the human pharmaceutical industry with specific weight on cancer, owing to a high level of activity in this area. Hence, it can be difficult to interpret the IPA results in relation to bovine chronic SCM. In addition, some of the bovine genes identified in the UCSC Genome Browser were unmapped by IPA and have not been included in the analysis.

Conclusion

Possible causative genes affecting chronic SCM in NR cattle were identified. The most significant genes were not the same across traits, which illustrates the complexity of chronic SCM. Alternatively defined traits can cover additional genetic information to improve genetic evaluation and selection for improved udder health in NR cattle. Moreover, TAD regions proved to make sense as alternative search space for candidate genes linked to significant SNP.

Acknowledgements

The Geno Breeding and AI Association and the Norwegian Dairy Herd Recording System are acknowledged for providing access to data. Dr Christos Dadousis is acknowledged for help with GWAS analyses. We thank Arne B. Gjuvsland for help on the SNP and TAD analysis strategies. The current study was funded by the Research Council of Norway project no. 233778.

Conflict of interest

The authors declare that they have no conflict of interest.

Availability of data

The phenotypes and SNP data used in this study are the property of Geno SA, the breeding organisation of Norwegian Red cattle. The phenotypes and SNP data will be made available by Geno SA after signing a Material Transfer Agreement, in which Geno SA can impose reasonable conditions, such as confidentiality.

References

Aulchenko Y.S., Ripke S., Isaacs A., van Duijn C.M. (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–6.

Bannerman D.D., Paape M.J., Lee J.W., Zhao X., Hope J.C., Rainard P. (2004) Escherichia coli and Staphylococcus aureus elicit differential innate immune responses following intramammary infection. *Clinical and Diagnostic Laboratory Immunology* **11**, 463–72.

Bougarn S., Cunha P., Harmache A., Fromageau A., Gilbert F.B., Rainard P. (2010) Muramyl dipeptide synergizes with Staphylococcus aureus lipoteichoic acid to recruit neutrophils in the mammary gland and to stimulate mammary epithelial cells. *Clinical and Vaccine Immunology* **17**, 1797–809.

Boutet P., Sulon J., Closset R., Detilleux J., Beckers J.F., Bureau F., Lekeux P. (2007) Prolactin-induced activation of nuclear factor kappaB in bovine mammary epithelial cells: role in chronic mastitis. *Journal of Dairy Science* **90**, 155–64.

Devlin B., Risch N. (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29**, 311–22.

Dixon J.R., Selvaraj S., Yue F., Kim A., Li Y., Shen Y., Hu M., Liu J.S., Ren B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–80.

Ekine C.C., Rowe S.J., Bishop S.C., de Koning D.J. (2014) Why breeding values estimated using familial data should not be used for genome-wide association studies. *G3 (Bethesda)* **4**, 341-7.

Ellington J.K., Elhofy A., Bost K.L., Hudson M.C. (2001) Involvement of mitogen-activated protein kinase pathways in Staphylococcus aureus invasion of normal osteoblasts. *Infection and Immunity* **69**, 5235–42.

Griffiths D.J., Barbet N.C., McCready S., Lehmann A.R., Carr A.M. (1995) Fission yeast rad17: a homologue of budding yeast RAD24 that shares regions of sequence similarity with DNA polymerase accessory proteins. *The EMBO Journal* **14**, 5812–23.

Harmon R.J. (1994) Symposium - mastitis and genetic evaluation for somatic-cell count - physiology of mastitis and factors affecting somatic-cell counts. *Journal of Dairy Science* **77**, 2103–12.

Haugaard K., Heringstad B., Whist A.C. (2013) Genetic associations between somatic cell score and pathogen-specific subclinical mastitis in Norwegian Red cows. *Journal of Animal Breeding and Genetics* **130**, 98–105.

Kirsanova E., Heringstad B., Lewandowska-Sabat A., Olsaker I. (2018) *Genome-wide Association Study for Alternative Subclinical Mastitis Traits in Norwegian Red.* World Congress on Genetics Applied to Livestock Production, http://www.wcgalp.org/system/files/proceedings/2018/genome-wide-association-study-alternative-subclinical-mastitis-traits-norwegian-red.pdf.

Kirsanova E., Heringstad B., Lewandowska-Sabat A., Olsaker I. (2019) Alternative subclinical mastitis traits for genetic evaluation in dairy cattle. *Journal of Dairy Science*;**102**:5323-9.

Klungland H., Sabry A., Heringstad B. *et al.* (2001) Quantitative trait loci affecting clinical mastitis and somatic cell count in dairy cattle. *Mammalian Genome* **12**, 837–42.

Kobayashi Y. (2006) Neutrophil infiltration and chemokines. *Critical reviews in immunology* **26**, 307–16.

Korte A., Farlow A. (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* **9**, 29.

Krefting J., Andrade-Navarro M.A., Ibn-Salem J. (2018) Evolutionary stability of topologically associating domains is associated with conserved gene regulation. *BMC Biology* **16**, 87.

Kuhn C., Bennewitz J., Reinsch N. *et al.* (2003) Quantitative trait loci mapping of functional traits in the German Holstein cattle population. *Journal of Dairy Science* **86**, 360–8.

Kuhn R.M., Haussler D., Kent W.J. (2013) The UCSC genome browser and associated tools. *Briefings in Bioinformatics* **14**, 144–61.

Lewandowska-Sabat A.M., Boman G.M., Downing A., Talbot R., Storset A.K., Olsaker I. (2013) The early phase transcriptome of bovine monocyte-derived macrophages infected with Staphylococcus aureus in vitro. *BMC Genomics* **14**, 891.

Madsen P. (2012) *DmuTrace. A program for extracting the pedigree for a subset of animals from a larger pedigree. Version 2*. Aarhus University.

Madsen P., Jensen J. (2013) *A User's Guide to DMU. A Package for Analysing Multivariate Mixed Models. Version 6, release 5.2*. Aarhus University.

Matukumalli L.K., Lawley C.T., Schnabel R.D. *et al.* (2009) Development and characterization of a high density SNP genotyping assay for cattle. *PLoS ONE* **4**, e5350.

Meredith B.K., Berry D.P., Kearney F., Finlay E.K., Fahey A.G., Bradley D.G., 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., Lynn D.J. (2012) Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. *BMC Genetics* **13**, 21.

Nograles K.E., Zaba L.C., Guttman-Yassky E. *et al.* (2008) Th17 cytokines interleukin (IL)-17 and IL-22 modulate distinct inflammatory and keratinocyte-response pathways. *British Journal of Dermatology* **159**, 1092–102.

Olsen H.G., Knutsen T.M., Lewandowska-Sabat A.M. *et al.* (2016) Fine mapping of a QTL on bovine chromosome 6 using imputed full sequence data suggests a key role for the group-specific component (GC) gene in clinical mastitis and milk production. *Genetics Selection Evolution* **48**, 79.

Pareek R., Wellnitz O., Van Dorp R., Burton J., Kerr D. (2005) Immunorelevant gene expression in LPS-challenged bovine mammary epithelial cells. *Journal of Applied Genetics* **46**, 171–7.

Pfaff C.M., Marquardt Y., Fietkau K., Baron J.M., Luscher B. (2017) The psoriasis-associated IL-17A induces and cooperates with IL-36 cytokines to control keratinocyte differentiation and function. *Scientific Reports* **7**, 15631.

Pope B.D., Ryba T., Dileep V. *et al.* (2014) Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**, 402–5.

Quinlan A.R., Hall I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–2.

Rainard P., Fromageau A., Cunha P., Gilbert F.B. (2008) Staphylococcus aureus lipoteichoic acid triggers inflammation in the lactating bovine mammary gland. *Veterinary Research* **39**, 52.

Rao S.S.P., Huntley M.H., Durand N.C. *et al.* (2014) A 3D Map of the human genome at Kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–80.

Raven L.A., Cocks B.G., Hayes B.J. (2014) Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics* **15**, 62.

Roth S., Rottach A., Lotz-Havla A.S. *et al.* (2014) Rad50-CARD9 interactions link cytosolic DNA sensing to IL-1beta production. *Nature Immunology* **15**, 538–45.

Rudan M.V., Barrington C., Henderson S., Ernst C., Odom D.T., Tanay A., Hadjur S. (2015) Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Reports* **10**, 1297–309.

Sahana G., Guldbrandtsen B., Bendixen C., Lund M.S. (2010) Genome-wide association mapping for female fertility traits in Danish and Swedish Holstein cattle. *Animal Genetics* **41**, 579–88.

Sahana G., Guldbrandtsen B., Thomsen B., Holm L.E., Panitz F., Brondum R.F., Bendixen C., 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**, 7258–75.

Sahana G., Guldbrandtsen B., Thomsen B., Lund M.S. (2013) Confirmation and fine-mapping of clinical mastitis and somatic cell score QTL in Nordic Holstein cattle. *Animal Genetics* **44**, 620–6.

Shen Y., Yue F., McCleary D.F. *et al.* (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–20.

Sodeland M., Kent M.P., Olsen H.G., Opsal M.A., Svendsen M., Sehested E., Hayes B.J., Lien S. (2011) Quantitative trait loci for clinical mastitis on chromosomes 2, 6, 14 and 20 in Norwegian Red cattle. *Animal Genetics* **42**, 457–65.

Spanou E., Kalisperati P., Pateras I.S., Papalampros A., Barbouti A., Tzioufas A.G., Kotsinas A., Sougioultzis S. (2017) Genetic Variability as a regulator of TLR4 and NOD signaling in response to bacterial driven DNA damage response (DDR) and inflammation: focus on the gastrointestinal (GI) tract. *Frontiers in Genetics* **8**, 65.

Strillacci M.G., Frigo E., Canavesi F. *et al.* (2014) Quantitative trait loci mapping for conjugated linoleic acid, vaccenic acid and (9) - desaturase in Italian Brown Swiss dairy cattle using selective DNA pooling. *Animal Genetics* **45**, 485–99.

Svendsen M., Heringstad B. (2006) *Somatic cell count as an indicator of sub-clinical mastitis.* Genetic parameters and correlations with clinical mastitis. *Interbull Bulletin*, 12.

Svishcheva G.R., Axenovich T.I., Belonogova N.M., van Duijn C.M., Aulchenko Y.S. (2012) Rapid variance components-based method for whole-genome association analysis. *Nature Genetics* **44**, 1166–70.

TINE (2017) *TINE Rådgivning*. Statistikksamling 2017 [TINE Annual Satistics 2017]. Accessed 12 October 2018, Available at: URL (in Norwegian, with English Tables). https://medlem.tine.no/fagprat/helse/_attachment/427411?_ts=16167bec8ab.

Turner S.D. (2014) *qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots*. *bioRxiv* 10.1101/005165. doi: 10.1101/005165

VanRaden P.M., Wiggans G.R. (1991) Derivation, calculation, and use of national animal model information. *Journal of Dairy Science* **74**, 2737–46.

Wang M., Hancock T.P., Chamberlain A.J., Vander Jagt C.J., Pryce J.E., Cocks B.G., Goddard M.E., Hayes B.J. (2018) Putative bovine topological association domains and CTCF binding motifs can reduce the search space for causative regulatory variants of complex traits. *BMC Genomics* **19**, 395.

Wu X., Lund M.S., Sahana G., Guldbrandtsen B., Sun D., Zhang Q., Su G. (2015) Association analysis for udder health based on SNP-panel and sequence data in Danish Holsteins. *Genetics Selection Evolution* **47**, 50.

Yonezawa T., Haga S., Kobayashi Y., Katoh K., Obara Y. (2008) Unsaturated fatty acids promote proliferation via ERK1/2 and Akt pathway in bovine mammary epithelial cells. *Biochemical Biophysical Research Communications* **367**, 729–35.

Zimin A.V., Delcher A.L., Florea L. *et al.* (2009) A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biology* **10**, R42.

Ødegard J., Heringstad B., Klemetsdal G. (2004) Short communication: bivariate genetic analysis of clinical mastitis and somatic cell count in Norwegian dairy cattle. *Journal of Dairy Science* **87**, 3515–7.

Ødegård J., Klemetsdal G., Heringstad B. (2003) Variance components and genetic trend for somatic cell count in Norwegian Cattle. *Livestock Production Science* **79**, 135–44.

Østeras O., Solbu H., Refsdal A.O., Roalkvam T., Filseth O., Minsaas A. (2007) Results and evaluation of thirty years of health recordings in the Norwegian dairy cattle population. *Journal of Dairy Science* **90**, 4483–97.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1 Top ingenuity canonical pathways.

Table S2 Gene ID detected based on topologically associated domains (TAD) positions.

Table S3 Top networks with score.

Appendix S1. Coordinates of converted TADs and percentage of the UMD3.1.1 genome covered by converted TADs.

List S1 List of abbreviations.

Supporting information.

Additional supporting information related to this article may be found online in the Supporting Information section at the end of the article at https://onlinelibrary.wiley.com/doi/full/10.1111/age.12886

Table S1. Top ingenuity canonical pathways.

Table S2. Gene ID detected based on topologically associated domains (TAD) positions.

Table S3. Top networks with score.

Appendix S1. Coordinates of converted TADs and percentage of the UMD3.1.1 genome converted by converted TADs.

List S1. List of abbreviations.

Paper III

Expression Analysis of Candidate Genes for Chronic Subclinical Mastitis in Norwegian Red

E. Kirsanova[*, 1], P. Boysen[†], G. M. Johansen[†], B. Heringstad[‡], A. Lewandowska-Sabat[*, #], I. Olsaker[*].

[*] Department of Basic Sciences and Aquatic Medicine, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Oslo, Norway

[†] Department of Food Safety and Infection Biology, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Oslo, Norway

[‡] Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, Ås, Norway

[#] Research Support Office, Norwegian University of Life Sciences, Ås, Norway

[1] Corresponding author: elena.kirsanova@nmbu.no

Submitted to Journal of Dairy Science, 7 January 2020

ABSTRACT

Chronic subclinical mastitis is a long-term inflammation in the udder, characterized with high somatic cell count (SCC) in milk. Previously, several novel alternative subclinical mastitis (SCM) traits for Norwegian Red (NR) cattle have been defined in order to improve breeding strategies against chronic SCM. Quantitative trait loci and candidate genes affecting chronic SCM in NR have been identified. The aim of this study was to analyse the expression profiles of fourteen selected candidate genes (*RAD17*, *ACOT2*, *ACOT4*, *FOS*, *CXCL1*, *CXCL8*, *CCNB1*, *CDK7, TGFB3, SEL1L*, *STAT4*, *C6*, *GLI2*, *SLC18A2*). Twenty healthy NR cows with official

genomic estimated breeding values (GEBV) for lactation average somatic cell score (LSCS) were selected. Where ten cows with high GEBV for LSCS (cows with low probability to have high SCC in milk during lactation) and ten cows with low GEBV for LSCS (cows with high probability of having high SCC in milk). RNA from unstimulated peripheral blood mononuclear cells (PBMC) from these were isolated. Two out of the fourteen analysed genes showed significantly different results between groups. The group with high GEBV for LSCS displayed significantly higher expression of the *CXCL1* gene than the low GEBV group. Grouping by lactation stage revealed significant differential expression of the *FOS* gene, with higher expression in early lactation (two – three months after calving) compared to late lactation (seven – eight months after calving). In addition, flow cytometry was performed on the PBMC samples to analyse if number and type of isolated cells influence the gene expression in the groups. The results in the current study provide identified genes that can be considered as possible candidate genes for chronic SCM in NR cows.

Keywords: *RNA expression-profile, peripheral blood mononuclear cells, Dairy Cattle, Somatic Cell Count*

INTRODUCTION

Chronic subclinical mastitis (SCM) is a complex trait, characterized by increased somatic cell count (SCC) in milk (Harmon, 1994). SCM is a frequent disease that leads to economic loss due to reduced milk production and culling of animals. It is often characterized by inflammatory response in the mammary gland mostly caused by environmental or pathogenic microorganisms (Oviedo-Boyso et al., 2007). One of the bacterial species known to cause subclinical bovine mastitis is *Streptococcus agalactiae* (*S. agalactiae*). In order to identify genes associated with susceptibility to mastitis in dairy cattle several studies on expression profiles have been

performed during the last decades (*e.g.* Fonseca et al., 2009; Lewandowska-Sabat et al., 2018). On Norwegian Red (NR) cattle microRNA and mRNA expression profiles of bovine monocyte-derived macrophages stimulated *in vitro* with different strains of *S. agalactiae*, have been analysed (Lewandowska-Sabat et al., 2018; Lewandowska-Sabat et al., 2019). Significantly up-regulated transcript levels of *tumor necrosis factor* and *tumor necrosis factor alpha* (*TNF*, *TNFα*), *interleukin-1 beta* (*IL-1β*), *interleukin-6* (*IL-6*), *interleukin-8* (*IL-8*) and *interleukin-10* (*IL-10*) were found in macrophages stimulated with ST12 and ST103 strains, and *transforming growth factor beta 1* (*TGFβ1*) was identified as down-regulated in macrophages stimulated with ST12 strain. The pro-inflammatory cytokines such as *TNFα*, *IL-1β*, *IL-6* and *IL-8* produced by macrophages are required to kill pathogens entering the mammary glands (Stout & Suttles, 2004). *IL-10* and *TGFβ1* are anti-inflammatory cytokines required to inhibit further activation of pro-inflammatory cytokines and preventing chronic conditions. However, through alternative activation of macrophages (Martinez et al., 2009) intracellular pathogens can survive and be able to escape the host immune response mechanism thus leading to long term chronical infections (Gordon & Martinez, 2010).

The peripheral blood mononuclear cells (PBMC) consist of lymphocytes and monocytes, where the latter can differentiate into macrophages. Monocytes and macrophages are the first line of the innate immune response cells that react against invading pathogens and in the bovine udder macrophages are present in the mammary gland (Sarikaya et al., 2006). The macrophages have ability to perform phagocytosis and produce pro-inflammatory and anti-inflammatory cytokines that have an essential role in defence against invading bacteria and other pathogens. The PBMC has been used to study the pathogenic mechanisms in disease processes and expression profiles of genes in cattle (MacHugh et al., 2009). Therefore, analysis of PBMC constitute a valid and well documented method for analysis of gene expression profiles in infected and healthy cows.

Elevated SCC in milk is an indicator of mastitis, and SCC is commonly used in the genetic evaluation of udder health. Genomic estimated breeding values (GEBV) for lactation average somatic cell score (LSCS) are included in the routine genetic evaluation of NR and in the total merit index used for selection (www.norwegianred.com). Recent research has shown significant differentially expressed genes between groups of NR cows with high and low GEBV for LSCS, respectively (Lewandowska-Sabat et al., 2019). Such as, in the group of animals with high GEBV for LSCS *myosin IF* (*MYO1F*) was found to be up-regulated and *peptidyl arginine deiminase 4* (*PADI4*) was down-regulated in bovine monocyte-derived macrophages (bMDMs) stimulated with *S. agalactiae* strain ST103 or strain ST12 respectively.

In a previous study (Kirsanova et al., 2019a) we have defined novel alternative chronic SCM traits for NR by analysing genotype data from NR bulls with corresponding testday SCC records of their daughters. Moreover, genes associated with the novel defined traits were identified (Kirsanova et al., 2019b). The aim of the present study was to perform further functional analysis of the identified genes. Expression profiles of 14 selected candidate genes for chronic SCM were investigated in unstimulated samples of PBMC from groups of NR cows with high and low GEBV for LSCS, respectively. In parallel, the PBMC samples were assessed for cellular composition by flow cytometry, in order to identify differences in number and types of cells as potential sources for differential gene expression.

MATERIAL AND METHODS

Animals and peripheral blood mononuclear cells isolation

Twenty healthy NR cows (presented in Table 1) from the research herd at Ås gård, Norwegian University of Life Sciences, were selected based on parity (one to six), stage of lactation (from one to eight months after calving) and official GEBV for LSCS. Ten cows with low GEBV for LSCS (< 100), and ten cows with high GEBV for LSCS (> 100) were selected. High GEBV for

LSCS is favourable, and cows with LSCS index above 100 have a higher probability than average to keep a low SCC in milk during the lactation period. Sampling of blood were performed by certified personnel and the study was approved by the Norwegian Animal Research Authority (Norwegian Food Safety Authority, FOTS id 16486). Ten ml blood from each individual were collected from the tail in EDTA tubes. PBMC were isolated from the blood samples using lymphoprep (Axis-Shield, Norway) and density gradient centrifugation (30 min, 3,200xRPM). After isolation, aliquots of the cells were either used for flow cytometry or snap frozen in liquid nitrogen and stored in -86°C for later extraction of total RNA.

Flow cytometry

Flow cytometric analysis was performed on fresh PBMC cell suspensions. Cells were first stained with LIVE/DEAD® Fixable yellow dead cell stain kit (Life Technologies), for 10 min on ice. The washed cells were then incubated for 20 min on ice with mixes of primary monoclonal antibodies as listed in Supplementary Table S1.1. Subsequently washed cells were incubated for 20 min on ice with mixes of secondary isotype-specific antibodies as listed in Supplementary Table S1.2. Finally, cells were washed and fixed for 10 min in room temperature using FACS Lysing solution (BD Biosciences), washed and resuspended with staining buffer until analysis within the same or next few days. All reagents were used at previously determined optimal concentrations. Stained samples were run in a 3-laser Gallios flow cytometer (Beckman Coulter). Compensation matrixes were resolved using single stain controls. Data was analyzed using Kaluza 2.1 software (Beckman Coulter). Gating strategies (Supplementary Figure) were set using secondary antibody-only controls and gross leukocyte subsets were defined as follows: Monocyte gate = high Side scatter (SSC), showing monocytes = $CD172a^{high}/CD26^{-}$ and dendritic-like cells (DC-like) = $CD172a^{low}/CD26^{+}$ (Talker et al., 2018; Hussen et al., 2013). Lymphocyte gate = low SSC, showing B-cells = sum of $CD21^{+}$ and $BB2^{+}$, natural killer (NK)

cells = NKp46^{+}/CD3^{-} (Storset et al., 2004), T-cells = CD3^{+} divided into γδT-cells = TCR1^{+} and αβT-cells = TCR1^{-}, the latter further divided into CD8^{+} and CD4^{+}, and finally TCR^{-}/CD4^{+}/CD25^{+} comprising activated and regulatory αβT-cells.

RNA extraction and reverse transcription-quantitative PCR analysis.

Total RNA were extracted from PBMC using the MirVANA isolation kit (Ambion, Austin, TX, USA) in accordance with the manufacturer`s instructions and subsequently treated with amplification grade DNase I (Invitrogen, Carlsbad, USA) to ensure removal of any traces of genomic DNA. RNA concentration was measured using NanoDrop 1000 (Thermo Fisher Scientific, Wilmington, USA). For reverse transcription 500 ng RNA per sample was used as template for cDNA synthesis employing the Tetro cDNA synthesis kit (Nordic BioSite, Norway) according to the manufacturer's protocol. For one qPCR reaction cDNA equivalent to 5 ng of total RNA was used. qPCR was performed using Express SYBR GreenER SuperMix with premixed ROX (Invitrogen) according to the manufacturer`s protocol, and run in a 7900HT Fast Real-Time PCR System (Applied Biosystems) with standard program: 50˚C for 2 min, 95˚C for 2 min followed by 40 cycles with first 95˚C for 15 sec and then 60˚C for 1 min, with final melting curve analysis included. The qPCR was set up as triple reactions for each sample with each specific primer pair. The efficiencies of all primer pairs were tested by template dilution series using pooled PBMC cDNA from three cows with low and three cows high official GEBV index for LSCS, and were found to be 100 % (±10). Additionally, no RT control for each sample were included to check for genomic DNA contaminations and negative controls with no added template (no template control) was run for each primer pair.

Results from the qPCR reactions were analysed in RQ Manager 1.2 (Applied Biosystems) with accepted standard deviation of ≤ 0.3 per triplicate. The ΔCt method was used to calculate gene expression compared to two reference genes, i.e. $\Delta Ct = Ct_{(target\ gene)} - Ct_{(reference\ gene)}$ and

normalized gene expression was defined as $2^{(-\Delta Ct)}$. The geNORM 6 gene bovine reference gene selection kit (Primerdesign Ltd, Southampton, UK) and the software qbase+ 3.2 (www.qbaseplus.com) was used in order to identify the most stable reference genes for PBMC derived from bovine blood. The optimal normalization factor i.e. average expression stability value (M) lower than 0.3, was calculated as the geometric mean of the reference targets *eukaryotic translation initiation factor 2B subunit beta* (*EIF2B2*) and *ribosomal protein L12* (*RPL12*) hence they were used as reference genes. Primers for each of the other genes were designed using the Primer-BLAST tool (Ye et al., 2012) avoiding Single Nucleotide Polymorphisms (SNPs) in the primer sequences as far as known in the NR genome, no G on the 5`end and max melting temperature difference of 2ºC between the two primers. All primers are listed in Supplementary Table S2. All the designed primer pairs were tested on liver samples of NR by PCR followed by separation in 2% agarose gels using 100 bp DNA ladder standard for verification of correct amplicon size. Moreover, all of the PCR products were sequenced and correct gene sequence verified by BLAST against the bovine genome.

Selection of candidate genes

Fourteen candidate genes (*RAD17* - checkpoint clamp loader component, *ACOT2* - acyl-CoA thioesterase 2, *ACOT4* - acyl-CoA thioesterase 4, *FOS* - Fos proto-oncogene, *CXCL1,* also known as *GRO1*: chemokine (C-X-C motif) ligand 1, *CXCL8,* also known as *IL8* - C-X-C motif chemokine ligand 8, *CCNB1* - cyclin B1, *CDK7* - cyclin-dependent kinase 7, *TGFB3* - transforming growth factor beta 3, *SEL1L* - ERAD E3 ligase adaptor subunit, *STAT4* - signal transducer and activator of transcription 4, *C6* – complement C6, *GLI2* - GLI family zinc finger 2, AP-1 transcription factor subunit, *SLC18A2* - solute carrier family 18 member A2) presented in Table 2 and 3 were selected based on the results from a genome-wide association study of alternative chronic SCM traits (Kirsanova et al., 2019b). The most frequent genes involved in

several canonical pathways associated with several defined traits or unique genes were analysed. The genes were located within converted bovine topologically associated domains (TAD; as converted from mouse and human) containing SNP significantly associated with several novel defined SCM traits described previously (Kirsanova et al., 2019a).

Statistical analyses

GraphPad Prism version 8.1.1 (GraphPad Software, La Jolla California USA, www.graphpad.com) was used for statistical analyses of the flow cytometry and gene expression results. Multiple t-test (one unpaired t-test per row, n=11) was performed by using the approach recommended by GraphPad software, as two-stage step-up method of Benjamini, Krieger and Yekutieli (Benjamini et al., 2006) with FDR (Q) = 5 % used to define significant results. For more power of computation it was assumed that all rows are samples from populations with the same scatter (SD). The samples were grouped by high (>100) and low (<100) official GEBV for LSCS (Table 1). Further, samples were grouped by stage of lactation defined as early period (one – two month after calving, n=12) and late period (seven – eight month after calving, n=8).

RESULTS

Gene expression results for the GEBV for LSCS groups is displayed in Figure 1 In addition gene expression normalized against the two reference genes *EIF2B2* and *RPL12* calculated as $2^{(-\Delta Ct)}$ is presented in Supplementary File S1.1 with GraphPad Prism 8 analyses in Supplementary File S1.2. Significant differential expression for the *CXCL1* gene (p-value = 0.0019, q-value = 0.02) were identified between the groups of cows with low or high GEBV for LSCS, with mean expression values of 1.19 and 1.90 respectively (Figure 1; Supplementary File S1.1 and S1.2). The gene *FOS* showed significant differential expression based on p-value

(0.02), but not on q-value (0.13), for these two groups, mean expression was 2.35 in the low and 1.84 in the high GEBV for LSCS group. Grouping samples by stage of lactation also revealed differences. In the early period of lactation (one or two months after calving) the *FOS* gene displayed higher expression than later in the lactation (seven - eight months after calving) with expression 2.36 and 1.69 respectively (Supplementary File S1.2). The gene *TGFB3* displayed a standard deviation (SD) higher than 0.3 for the triple reactions in sample one and eight. The reactions were repeated using new primer mix and new cDNA samples without any changes in the results. This may be due to undiscovered polymorphisms in the primer regions making primer binding unstable. Sample one and eight were excluded from further analysis, but no differential expression of this gene were observed among the remaining individuals (n=18). For the genes *SCL18A*, *C6* and *GLI2* no detectable expression was observed in the samples.

In order to assess if differences in PBMC composition might influence differential gene expression, the PBMC samples were analysed for lymphocyte subsets. The results are presented in Figure 2 and in Supplementary File S2.1 (lymphocyte composition) and S2.2 (GraphPad Prism 8 analyses). Supplementary Figure present the flow cytometric gating strategy showing representative plots of bovine PBMCs. No significant differences for any of the leukocyte cells were found neither between the high and low GEBV for LSCS groups (Figure 2) nor when grouping the samples for early and late lactation (Supplementary file S2.2).

DISCUSSION

In the present study we aimed to investigate selected candidate genes for chronic SCM traits in order to identify potential differences between groups of NR cows with high and low GEBV for LSCS. The *CXCL1* gene was observed to have higher level of expression in the group of cows with high GEBV for LSCS, that is, cows with expected low probability to get high SCC

level in milk in comparison to cows with low GEBV for LSCS. This gene was found in association with a long-term chronic subclinical mastitis trait with SCC > 200,000 cells/mL during three test-days period (Table 2; Kirsanova et al., 2019b) and reported previously as up-regulated after infection of monocyte cells with mastitis causing bacteria (Lewandowska-Sabat et al., 2012; Han et al., 2019). In PBMC, the main sources of *CXCL1* are described to be monocytes and dendritic cells (DCs) (Eberlein et al., 2010; Hussen et al., 2013). The differential *CXCL1* expression was unlikely to be caused by elevated numbers of monocytes, DCs, or any other cellular fractions in the PBMC, as all major cell subsets were similar in numbers as measured by flow cytometry. This indicates a stronger constitutive cellular expression of *CXCL1* in the high GEBV for LSCS group. *CXCL1* is a gene involved in massive influx of neutrophils to inflamed or infected tissue (Rainard et al., 2008; Sipka et al., 2014). Higher number of neutrophils has been reported in association with cortisol (stress-related hormone) concentration (Kulberg et al., 2002). Correspondingly, higher number of neutrophils for cows with high milk production compared to average production have been reported for Holstein cows (Kimura et al.. 1999) and in case of stimulation with lipopolysaccharide to activate the immune system (Dickson et al., 2019). In the current study, neutrophils were not quantified. The cows were healthy and had low SCC in milk at the time of sampling (Table 1). However, due to the higher *CXCL1* gene expression, immune cells may react more quickly on invading pathogens in the udder, clearing the infection and thus avoiding development of clinical or chronic subclinical mastitis with critical levels of SCC. Another gene, *FOS,* which was found to be significantly differentially expressed in our study has previously been reported by several studies as significantly higher expressed in cells stimulated with different bacterial species (e.g. *Staphylococcus aureus*, *Streptococcus uberis* and *S. agalactiae*) causing bovine mastitis (e.g. Lutzow et al., 2008; Naeem et al., 2012; Lewandowska-Sabat et al., 2013; Kosciuczuk et al., 2017; Lewandowska-Sabat et al., 2019). In addition, the *FOS* gene was found in association

with a long-term chronic SCM200_3 trait with SCC > 200,000 cells/mL during three test-days period, similarly to the *CXCL1* gene (Table 2; Kirsanova et al. 2019a). In the current study *FOS* displayed higher expression in PBMC samples from cows with low GEBV for LSCS, but significant only based on p-value (0.02) results and not significant for q-value (0.13), which may indicate false results. Further, the *FOS* gene was significantly higher expressed in the early lactation period compared to late (Supplementary File S1.2). Within the PBMC population *FOS* is restricted to myeloid cells (Sariban et al., 1988), and as discussed for *CXCL1*, cell numbers cannot explain the differential *FOS* gene expression, since monocytes were similar between the LSCS groups as well as between the lactation stage groups. This gene has an important role in regulation of the innate immune response through coordination of the induction of genes encoding inflammatory mediators (reviewed by Guha & Mackman, 2001). Moreover, *FOS* plays a main role in proliferation, differentiation and survival of cells (Bahrami & Drabløs, 2016) and is one of the immediate-early genes, that can be activated and react within minutes after stimulation (Healy et al., 2013). Interestingly, the *FOS* gene showed lower expression in the high GEBV for LSCS group, which may indicate that a lower basic expression level leads to fewer number of cells moving from blood in to the infected tissue. This corresponds to the previously reported association of the *FOS* gene with SCM traits with SCC threshold above 150,000 – 400,000 cells/mL (Kirsanova et al., 2019b). The association might be explained by higher expression of the *FOS* gene that in turn leads to higher number of cells in milk. The present information on the expression of the *CXCL1* and *FOS* genes might be useful for development of better strategies to improve immune response and decrease high SCC in milk, eventually reducing subclinical infection in the mammary gland.

However, further investigation of candidate genes for chronic SCM is needed. Analysis of cows during periods of higher stress, such as close to parturition or during an udder infection, might

reveal more important differences related to the genetic and molecular mechanisms involved in defence against infections.

CONCLUSIONS

Analyses of 14 candidate genes for chronic SCM in samples of PBMC from healthy NR cows revealed significant differential expression among the studied groups, with no influence from the number and type of lymphocyte subsets in the PBMC samples. The *CXCL1* gene displayed higher and the *FOS* gene lower expression in the group of cows with high GEBV for LSCS compared to the group with low GEBV for LSCS. Based on these results the expression of *CXCL1* and *FOS* candidate genes for chronic SCM in NR can be used as additional information for selection in order to improve breeding for better udder health.

ACKNOWLEDGEMENTS

Geno Breeding and AI Association and the Norwegian Dairy Herd Recording System are acknowledged for providing access to data. This work is part of the “Multimast” project (no. 233778) funded by the Research Council of Norway, Tine and Geno and coordinated by Norwegian University of Life Sciences (NMBU). We thank the staff at the research herd at Ås gård, Norwegian University of Life Sciences for help with collection of blood samples and head engineer Aleksandra Bodura Göksu for help with the gene expression analyses.

REFERENCES

Bahrami, S. and F. Drablos. 2016. Gene regulation in the immediate-early response process. Adv Biol Regul 62:37-49. doi: https://doi.org/10.1016/j.jbior.2016.05.001

Benjamini, Y., A. M. Krieger, and D. Yekutieli. 2006. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93(3):491-507. doi: https://doi.org/DOI 10.1093/biomet/93.3.491

Dickson, M. J., S. K. Kvidera, E. A. Horst, C. E. Wiley, E. J. Mayorga, J. Ydstie, G. A. Perry, L. H. Baumgard, and A. F. Keating. 2019. Impacts of chronic and increasing lipopolysaccharide exposure on production and reproductive parameters in lactating Holstein dairy cows. J Dairy Sci 102(4):3569-3583. doi: https://doi.org/10.3168/jds.2018-15631

Eberlein, J., T. T. Nguyen, F. Victorino, L. Golden-Mason, H. R. Rosen, and D. Homann. 2010. Comprehensive assessment of chemokine expression profiles by flow cytometry. J Clin Invest 120(3):907-923. doi: https://doi.org/10.1172/JCI40645

Fonseca, I., P. V. Silva, C. C. Lange, M. F. Guimaraes, M. M. Weller, K. R. Sousa, P. S. Lopes, J. D. Guimaraes, and S. E. Guimaraes. 2009. Expression profile of genes associated with mastitis in dairy cattle. Genet Mol Biol 32(4):776-781. doi: https://doi.org/10.1590/S1415-47572009005000074

Gordon, S. and F. O. Martinez. 2010. Alternative activation of macrophages: mechanism and functions. Immunity 32(5):593-604. doi: https://doi.org/10.1016/j.immuni.2010.05.007

Guha, M. and N. Mackman. 2001. LPS induction of gene expression in human monocytes. Cell Signal 13(2):85-94.

Han, H. 2019. Identification of several key genes by microarray data analysis of bovine mammary gland epithelial cells challenged with Escherichia coli and Staphylococcus aureus. Gene 683:123-132. doi: https://doi.org/10.1016/j.gene.2018.10.004

Harmon, R. J. 1994. Symposium - Mastitis and Genetic Evaluation for Somatic-Cell Count - Physiology of Mastitis and Factors Affecting Somatic-Cell Counts. Journal of Dairy Science 77(7):2103-2112. doi: https://doi.org/DOI 10.3168/jds.S0022-0302(94)77153-8

Healy, S., P. Khan, and J. R. Davie. 2013. Immediate early response genes and cell transformation. Pharmacol Ther 137(1):64-77. doi: https://doi.org/10.1016/j.pharmthera.2012.09.001

Hussen, J., A. Duvel, O. Sandra, D. Smith, I. M. Sheldon, P. Zieger, and H. J. Schuberth. 2013. Phenotypic and functional heterogeneity of bovine blood monocytes. PLoS One 8(8):e71502. doi: https://doi.org/10.1371/journal.pone.0071502

Kimura, K., J. P. Goff, and M. E. Kehrli, Jr. 1999. Effects of the presence of the mammary gland on expression of neutrophil adhesion molecules and myeloperoxidase activity in periparturient dairy cows. J Dairy Sci 82(11):2385-2392. doi: https://doi.org/10.3168/jds.S0022-0302(99)75489-5

Kirsanova, E., B. Heringstad, A. Lewandowska-Sabat, and I. Olsaker. 2019a. Alternative subclinical mastitis traits for genetic evaluation in dairy cattle. J Dairy Sci 102(6):5323-5329. doi: https://doi.org/10.3168/jds.2018-16104

Kirsanova, E., B. Heringstad, A. Lewandowska-Sabat, and I. Olsaker. 2019b. Identification of Candidate Genes Affecting Chronic Subclinical Mastitis in Norwegian Red cattle: Combining Genome Wide Association Study, Topologically Associated Domains and Pathway Enrichment Analysis. Anim Genet. doi: https://doi.org/10.1111/age.12886

Kosciuczuk, E. M., P. Lisowski, J. Jarczak, A. Majewska, M. Rzewuska, L. Zwierzchowski, and E. Bagnicka. 2017. Transcriptome profiling of Staphylococci-infected cow mammary gland parenchyma. BMC Vet Res 13(1):161. doi: https://doi.org/10.1186/s12917-017-1088-2

Kulberg, S., A. K. Storset, B. Heringstad, and H. J. Larsen. 2002. Reduced levels of total leukocytes and neutrophils in Norwegian cattle selected for decreased mastitis incidence. J Dairy Sci 85(12):3470-3475. doi: https://doi.org/10.3168/jds.S0022-0302(02)74435-4

Lewandowska-Sabat, A. M., G. M. Boman, A. Downing, R. Talbot, A. K. Storset, and I. Olsaker. 2013. The early phase transcriptome of bovine monocyte-derived macrophages infected with Staphylococcus aureus in vitro. BMC Genomics 14:891. doi: https://doi.org/10.1186/1471-2164-14-891

Lewandowska-Sabat, A. M., S. F. Hansen, T. R. Solberg, O. Osteras, B. Heringstad, P. Boysen, and I. Olsaker. 2018. MicroRNA expression profiles of bovine monocyte-derived macrophages infected in vitro with two strains of Streptococcus agalactiae. BMC Genomics 19(1):241. doi: https://doi.org/10.1186/s12864-018-4591-3

Lewandowska-Sabat, A. M., E. Kirsanova, C. Klopp, T. R. Solberg, B. Heringstad, O. Østerås, P. Boysen, and I. Olsaker. 2019. Transcription Profiling of Monocyte-Derived Macrophages Infected In Vitro With Two Strains of Streptococcus agalactiae Reveals Candidate Pathways Affecting Subclinical Mastitis in Cattle. Frontiers in Genetics 10. doi: https://doi.org/10.3389/fgene.2019.00689

Lutzow, Y. C., L. Donaldson, C. P. Gray, T. Vuocolo, R. D. Pearson, A. Reverter, K. A. Byrne, P. A. Sheehy, R. Windon, and R. L. Tellam. 2008. Identification of immune genes and proteins involved in the response of bovine mammary tissue to

Staphylococcus aureus infection. BMC Vet Res 4:18. doi: https://doi.org/10.1186/1746-6148-4-18

MacHugh, D. E., E. Gormley, S. D. Park, J. A. Browne, M. Taraktsoglou, C. O'Farrelly, and K. G. Meade. 2009. Gene expression profiling of the host response to Mycobacterium bovis infection in cattle. Transbound Emerg Dis 56(6-7):204-214. doi: https://doi.org/10.1111/j.1865-1682.2009.01082.x

Martinez, F. O., L. Helming, and S. Gordon. 2009. Alternative activation of macrophages: an immunologic functional perspective. Annu Rev Immunol 27:451-483. doi: https://doi.org/10.1146/annurev.immunol.021908.132532

Naeem, A., K. Zhong, S. J. Moisa, J. K. Drackley, K. M. Moyes, and J. J. Loor. 2012. Bioinformatics analysis of microRNA and putative target genes in bovine mammary tissue infected with Streptococcus uberis. J Dairy Sci 95(11):6397-6408. doi: https://doi.org/10.3168/jds.2011-5173

Oviedo-Boyso, J., J. J. Valdez-Alarcon, M. Cajero-Juarez, A. Ochoa-Zarzosa, J. E. Lopez-Meza, A. Bravo-Patino, and V. M. Baizabal-Aguirre. 2007. Innate immune response of bovine mammary gland to pathogenic bacteria responsible for mastitis. J Infect 54(4):399-409. doi: https://doi.org/10.1016/j.jinf.2006.06.010

Rainard, P., A. Fromageau, P. Cunha, and F. B. Gilbert. 2008. Staphylococcus aureus lipoteichoic acid triggers inflammation in the lactating bovine mammary gland. Veterinary Research 39(5):52. doi: https://doi.org/10.1051/vetres:2008034

Sariban, E., T. Mitchell, A. Rambaldi, and D. W. Kufe. 1988. c-sis but not c-fos gene expression is lineage specific in human myeloid cells. Blood 71(2):488-493.

Sarikaya, H., G. Schlamberger, H. H. Meyer, and R. M. Bruckmaier. 2006. Leukocyte populations and mRNA expression of inflammatory factors in quarter milk fractions at

different somatic cell score levels in dairy cows. J Dairy Sci 89(7):2479-2486. doi: https://doi.org/10.3168/jds.S0022-0302(06)72322-0

Sipka, A., S. Klaessig, G. E. Duhamel, J. Swinkels, P. Rainard, and Y. Schukken. 2014. Impact of intramammary treatment on gene expression profiles in bovine Escherichia coli mastitis. PLoS One 9(1):e85579. doi: https://doi.org/10.1371/journal.pone.0085579

Storset, A. K., S. Kulberg, I. Berg, P. Boysen, J. C. Hope, and E. Dissen. 2004. NKp46 defines a subset of bovine leukocytes with natural killer cell characteristics. Eur J Immunol 34(3):669-676. doi: https://doi.org/10.1002/eji.200324504

Stout, R. D. and J. Suttles. 2004. Functional plasticity of macrophages: reversible adaptation to changing microenvironments. J Leukoc Biol 76(3):509-513. doi: https://doi.org/10.1189/jlb.0504272

Talker, S. C., A. Baumann, G. T. Barut, I. Keller, R. Bruggmann, and A. Summerfield. 2018. Precise Delineation and Transcriptional Characterization of Bovine Blood Dendritic-Cell and Monocyte Subsets. Front Immunol 9:2505. doi: https://doi.org/10.3389/fimmu.2018.02505

Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. L. Madden. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13:134. doi: https://doi.org/10.1186/1471-2105-13-134

APPENDIX

Supplementary **Table S1.** Antibodies used for flow cytometry analyses.

Supplementary **Table S2.** List of primers used for quantitative PCR (qPCR).

Supplementary **File S1**. Normalized expression of candidate genes for chronic subclinical mastitis (SCM) of twenty Norwegian Red cows (S1.1), including GraphPad Prism8 analyses (S1.2)

Supplementary **File S2.** Number of leukocyte cells in % for twenty Norwegian Red cows (S2.1), including GraphPad Prism8 analyses (S2.2).

Supplementary **Figure.** Flow cytometric gating strategy showing representative plots of bovine PBMCs.

TABLES

Table 1. Twenty healthy Norwegian Red cows selected based on genomic estimated breeding values (GEBV) for lactation average somatic cell score (LSCS) presented with LSCS index, parity, stage of lactation and SCC per November 2018 (the date of collection of the blood samples).

Sample N	LSCS index[1]	parity	stage of lactation[2]	SCC x1,000 cells/mL[3]
s1	118.4	6	Early	90
s2	99.5	6	Early	120
s3	93.6	3	Late	70
s4	120.5	3	Late	20
s5	114.5	2	Late	10
s6	89.7	2	Late	70
s7	97.0	2	Early	20
s8	110.6	2	Early	10
s9	97.9	2	Early	20
s10	107.1	1	Early	10
s11	114.0	1	Early	40
s12	106.4	2	Late	110
s13	114.3	1	Early	20
s14	99.1	1	Late	10
s15	114.4	1	Late	40
s16	98.9	1	Late	50
s17	110.5	1	Early	10
s18	87.1	1	Early	10
s19	95.8	1	Early	10
s20	89.8	1	Early	40

LSCS index[1] – official genomic estimated breeding value (GEBV) for lactation average somatic cell score (LSCS).

stage of lactation[2] – is defined as early period (one – two month after calving, n=12) and late period (seven – eight month after calving, n=8).

SCC x 1,000 cells/mL[3] – Somatic cell count per millilitre milk.

436 **Table 2.** Candidate genes for alternatively defined traits for chronic subclinical mastitis (SCM) and lactation average somatic cell score (LSCS)
437 selected based on the results from Kirsanova et al. (2019a), with NCBI Assembly ID[1], description of function[1] and associated alternatively
438 defined SCM traits[2].

Gene ID	Assembly ID[1]	Function[1]	Association with SCM traits[2]
RAD17	513600	cellular response to DNA damage, DNA replication	SCM50, SCM100, SCM150, SCM200, LSCS
ACOT2	785383	regulate intracellular levels of acyl-CoAs, free fatty acids and CoASH	SCM150, SCM200, SCM250, SCM300, SCM350, SCM400, D400
ACOT4	511431	regulate intracellular levels of acyl-CoAs, free fatty acids and CoASH	SCM150, SCM200, SCM250, SCM300, SCM350, SCM400, D400
FOS	280795	regulation of innate immune response, proliferation, differentiation and survival in cells	SCM150, SCM200, SCM250, SCM300, SCM350, SCM400, SCM200_3, D400
CXCL1	281212	has chemotactic activity for neutrophils, may play a role in inflammation	SCM200_3
CXCL8	280828	attracts neutrophils, basophils, and T-cells, involved in neutrophil activation	SCM200_3
CCNB1	327679	activation of mitosis	SCM50, SCM100, SCM150, SCM200, LSCS
CDK7	515462	cell cycle regulation	SCM50, SCM100, SCM150, SCM200, LSCS
TGFB3	538957	cellular activities, fibrosis and immune responses	SCM150, SCM200, SCM250, SCM300, SCM350, SCM400, D400
SEL1L	614959	protein degradation	SCM400_3

STAT4	515988	regulating the differentiation of T helper cells and responses to IL1	SCM400_3
C6	507749	innate immune response through membrane attack complex	D50
GLI2	510255	a role during embryogenesis	SCM400_3
SLC18A2	282471	correct activity of the monoaminergic systems through synaptic vesicles	SCM400_3

[1] ID /function – gene ID and gene function obtained from the NCBI database (https://www.ncbi.nlm.nih.gov/) and the GeneCards human gene database (http://www.genecards.org).

[2] traits - SCM50, -100, -150, -200, -250, -300, -350, -400 – subclinical mastitis above the threshold in two test-days at 50,000, 100,000, 150,000, 200,000, 250,000, 300,000, 350,000, 400,000 cells/ml respectively; LSCS – lactation-average somatic cell score during 1 – 3 lactations; SCM200_3, SCM400_3 – subclinical mastitis traits above the threshold 200,000 and 400,000 in three test-days; D50, D400 – number of days before the first case with SCM50 and SCM400 respectively.

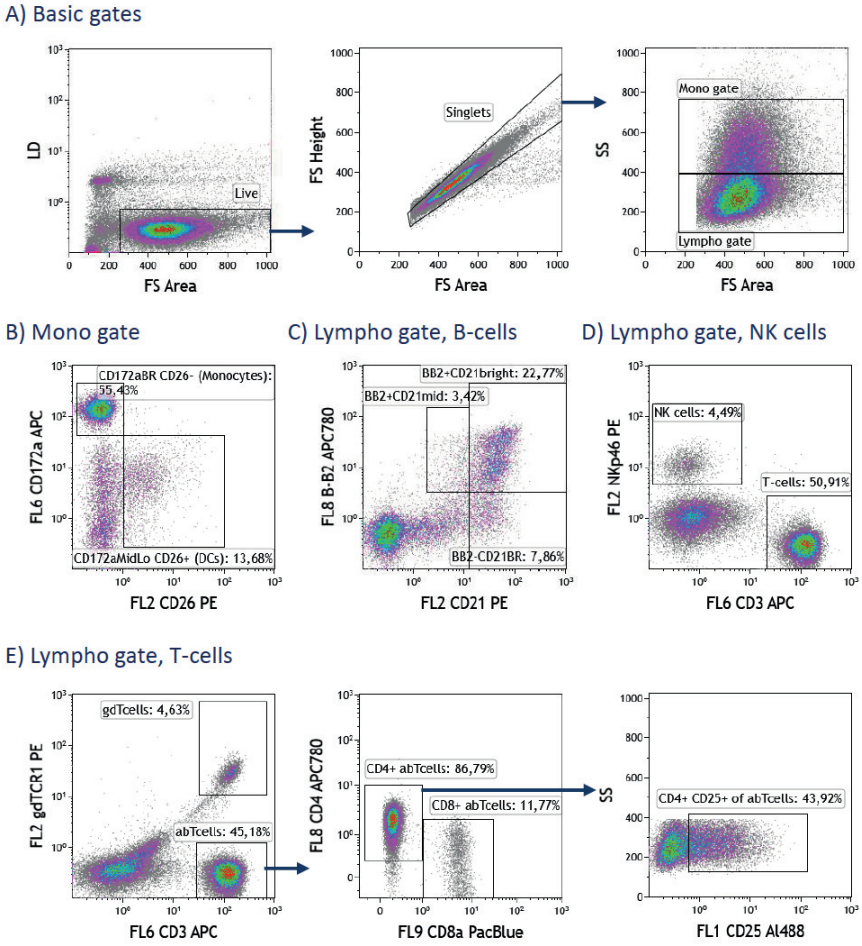
446 **Table 3.** Analysed candidate genes for alternative chronic subclinical mastitis (SCM) traits (Kirsanova et al., 2019a), with position in the bovine
447 genome UMD3.1.1 (UCSC Genome Browser assembly ID: bosTau8), GWAS significant single nucleotide polymorphism (SNP; Kirsanova et al.,
448 2019b), p-value for SNP and topologically associated domain (TAD) region, where the gene was identified.

Gene ID	BTA	Gene position	SNP position	p-value	TAD region
RAD17	20	10266268 – 10300200	10493654	1.10E-06	10159056 – 11360403
ACOT2	10	85358183 – 85365711	87958006	2.90E-06	84896469 – 89937333
ACOT4	10	85375286 – 85380537	87958006	2.90E-06	84896469 – 89937333
FOS	10	86883739 – 86887170	87958006	2.90E-06	84896469 – 89937333
CXCL1	6	90822721 – 90824859	90184750	4.00E-06	89308822 – 90824904
CXCL8	6	90559882 – 90563647	90184750	4.00E-06	89308822 – 90824904
CCNB1	20	10431452 – 10441126	10435385	1.60E-06	10407200 – 11354566
CDK7	20	10356010 – 10384171	10493654	1.10E-06	10159056 – 11360403
TGFB3	10	88242178 – 88266347	87958006	2.90E-06	84896469 – 89937333
SEL1L	10	93900621 – 93965136	94462757	4.30E-06	93958242 – 100721748
STAT4	2	79939329 – 80043984	75996043	2.60E-06	71761492 – 80477117

449

450

FIGURES



Supplementary **Figure.** Flow cytometric gating strategy showing representative plots of bovine PBMCs. A) Exclusion of dead cells, single cell gate and definition of regions for myeloid cells (Mono gate) and lymphocytes (Lympho gate). B) Monocytes defined as CD172abright/CD26- in Mono gate, dendritic-like cells (DCs) defined as CD172Mid-Lo/CD26+. C) B-cells were defined as the sum of indicated regions in the Lympho gate, i.e. expressing B-B2 antigen and/or CD21. D) NK cells defined as NKp46+/CD3- cells in the

Lympho gate. E) T-cells defined as CD3+ cells in the Lympho gate subdivided into gamma delta T-cells (TCR1+) or alpha beta T-cells (TCR1-). Further subdivision of T-cells into CD8+ or CD4+ T-cells (shown for abT-cells), the latter subset analyzed for CD25+ cells to indicate activated or regulatory CD4+ abT-cells.

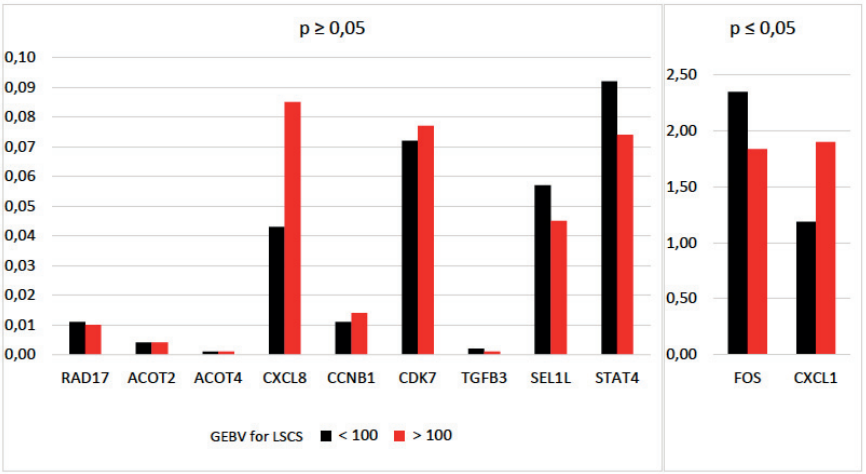
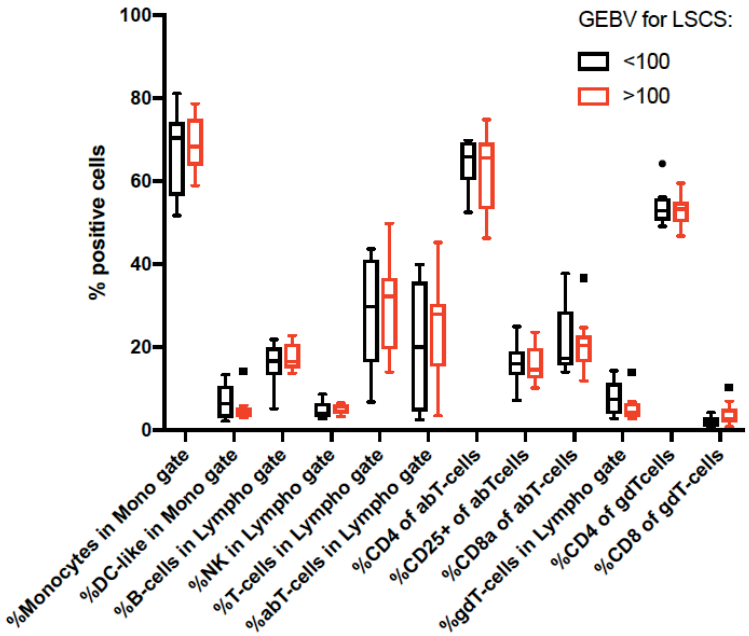


Figure 1. Gene expression measured by reverse transcription-quantitative PCR (RT-qPCR). The expression was normalized against two reference genes (*EIF2B2* and *RPL12*) and calculated as $2^{(-\Delta Ct)}$.



470

471 **Figure 2.** Leukocyte subsets investigated in groups of Norwegian Red cows with low (< 100)

472 and high (> 100) official genomic estimated breeding value (GEBV) for lactation average

473 somatic cell score (LSCS).

474

APPENDIX

Supplementary File S1.1 Expression of candidates genes for chronic subclinical mastitis (SCM) in PBMC from twenty healthy Norwegian Red cows. The expression was normalized against two reference genes (EIF2B2 and RPL12) and is presented as $2^{(-\Delta Ct)}$.

gene ID Sample N	RAD17	ACOT2	ACOT4	FOS	CXCL1	CXCL8	CCNB1	CDK7	TGFB3	SEL1L	STAT4
s1	0,006	0,004	0,001	2,197	2,118	0,047	0,004	0,103	-	0,031	0,052
s2	0,016	0,007	0,001	5,662	1,673	0,079	0,006	0,077	0,002	0,070	0,095
s3	0,006	0,005	0,001	2,652	1,274	0,048	0,006	0,067	0,004	0,068	0,138
s4	0,010	0,004	0,001	2,120	0,955	0,056	0,004	0,043	0,001	0,026	0,095
s5	0,011	0,003	0,001	1,450	1,603	0,135	0,013	0,078	0,001	0,062	0,069
s6	0,005	0,003	0,000	1,276	0,607	0,017	0,007	0,056	0,002	0,060	0,128
s7	0,023	0,005	0,001	2,799	0,997	0,029	0,012	0,092	0,003	0,069	0,110
s8	0,018	0,008	0,001	4,240	1,883	0,183	0,009	0,065	-	0,125	0,094

s9	0,013	0,004	0,001	3,123	1,910	0,077	0,023	0,055	0,002	0,072	0,185
s10	0,010	0,002	0,001	0,738	1,168	0,059	0,016	0,072	0,001	0,031	0,076
s11	0,006	0,003	0,000	1,523	2,566	0,074	0,033	0,139	0,001	0,032	0,047
s12	0,006	0,002	0,000	2,024	5,431	0,183	0,017	0,083	0,000	0,025	0,064
s13	0,004	0,001	0,000	1,383	1,033	0,036	0,012	0,061	0,001	0,023	0,099
s14	0,015	0,006	0,000	1,296	0,486	0,015	0,011	0,076	0,002	0,062	0,079
s15	0,018	0,007	0,001	1,343	1,721	0,046	0,017	0,068	0,002	0,051	0,076
s16	0,006	0,003	0,001	1,335	1,412	0,047	0,010	0,072	0,000	0,035	0,060
s17	0,006	0,002	0,002	1,333	0,516	0,031	0,013	0,061	0,000	0,041	0,063
s18	0,011	0,004	0,001	1,259	1,648	0,065	0,011	0,067	0,001	0,061	0,036
s19	0,004	0,001	0,001	0,438	0,310	0,017	0,009	0,066	0,001	0,019	0,052
s20	0,009	0,004	0,001	3,643	1,573	0,038	0,017	0,086	0,001	0,050	0,036

Supplementary File S1.2 GraphPad Prism8 analyses of candidate genes (n=11) for chronic subclinical mastitis (SCM) in PBMC from healthy Norwegian Red (NR) cows (n=20) grouped by genomic estimated breeding value (GEBV) for lactation average somatic cell score (LSCS) and by lactation stage. Normalized gene expression is presented as $2^{(-\Delta Ct)}$. Estimated P- and q-value are presented with correspondingly discovery range.

gene ID	Discovery?[1]	Mean of GEBV for LSCS <100[2] (n=10)	Mean of GEBV for LSCS > 100[3] (n=10)	P value	q value
RAD17	No	0,01	0,01	0,9954	0,95
ACOT2	No	0,00	0,00	0,9979	0,95
ACOT4	No	0,00	0,00	>0,999999	0,95
FOS	No	2,35	1,84	*0,0241*	0,13
CXCL1	*Yes*	*1,19*	*1,90*	*0,0019*	*0,02*
CXCL8	No	0,04	0,09	0,8533	0,95
CCNB1	No	0,01	0,01	0,9908	0,95
CDK7	No	0,07	0,08	0,9792	0,95
TGFB3	No	0,00	0,00	0,9969	0,95

SEL1L	No	0,06	0,04	0,9580	0,95
STAT4	No	0,09	0,07	0,9351	0,95
gene ID	Discovery?[1]	Mean of Early lactation[4] (n=12)	Mean of Late lactation[5] (n=8)	P value	q value
RAD17	No	0,01	0,01	0,9970	0,95
ACOT2	No	0,00	0,00	0,9987	0,95
ACOT4	No	0,00	0,00	0,9990	0,95
FOS	*Yes*	*2,36*	*1,69*	*0,0043*	*0,05*
CXCL1	No	1,45	1,69	0,3124	0,95
CXCL8	No	0,06	0,07	0,9757	0,95
CCNB1	No	0,01	0,01	0,9893	0,95
CDK7	No	0,08	0,07	0,9632	0,95
TGFB3	No	0,00	0,00	0,9993	0,95
SEL1L	No	0,05	0,05	0,9885	0,95
STAT4	No	0,08	0,09	0,9663	0,95

[1] Discovery? – defined significant results, False Discovery Rate (Q) = 5 %

[2] Mean of GEBV for LSCS < 100 – genomic estimated breeding value (GEBV) for lactation average somatic cell score (LSCS) < 100, i.e. cows with high probability of having high somatic cell count during lactation

[3] Mean of GEBV for LSCS > 100 – genomic estimated breeding value (GEBV) for lactation average somatic cell score (LSCS) > 100, i.e. cows with low probability of having high somatic cell count during lactation

[4] Mean of Early lactation – early period of lactation within one – two month after calving

[5] Mean of Late lactation – late period of lactation within seven – eight month after calving

Supplementary File S2.1 Percent of leukocyte cells in peripheral blood mononuclear cell (PBMC) samples (n=20) from healthy Norwegian Red cows.

Sample N	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15	s16	s17	s18	s19	s20
Lymphocyte subset, %																				
Monocytes in Mono gate	78,6	72,9	81,1	74,8	63,0	73,0	56,8	65,7	69,6	58,9	64,2	71,8	67,9	55,8	74,8	77,2	68,7	71,3	51,7	63,1
DC-like in Mono gate	5,0	5,7	7,8	3,3	14,2	13,0	7,1	5,9	9,8	5,2	4,4	3,7	3,4	13,4	3,0	3,5	4,6	2,2	2,4	3,3
B-cells in Lympho gate	13,7	15,0	16,7	16,1	16,5	16,5	21,8	21,5	13,5	15,0	17,1	22,8	16,5	20,3	20,4	19,6	15,1	18,8	13,0	5,2
NK in Lympho gate	4,7	4,3	4,4	4,2	4,4	3,1	8,7	6,6	-	3,9	5,9	6,2	6,0	2,8	3,3	4,1	4,5	3,6	3,9	8,5
T-cells in Lympho gate	49,8	38,9	43,7	37,7	30,5	40,6	25,2	20,4	42,1	34,4	36,2	14,0	31,4	28,3	33,0	31,3	17,4	6,8	17,5	13,2
abT-cells in Lympho gate	45,2	36,2	39,9	30,8	26,7	29,8	20,6	16,9	35,5	30,2	29,8	11,2	28,4	19,4	27,6	17,0	3,5	2,5	4,4	4,8
CD4 of abT-cells	74,8	69,2	66,3	51,5	68,6	69,8	64,8	65,5	66,4	54,0	61,3	65,9	70,7	61,3	65,8	69,6	46,2	65,4	52,5	57,2
CD25+ of abTcells	23,6	25,0	17,6	12,6	21,5	20,6	15,3	14,9	16,8	15,1	14,3	19,0	12,6	7,2	13,1	10,2	10,2	14,5	14,6	18,3
CD8a of abT-cells	11,9	18,2	16,4	24,7	16,8	14,0	15,8	18,6	15,5	22,0	21,5	19,8	15,4	18,7	21,0	16,3	36,6	26,8	37,8	34,0
gdT-cells in Lympho gate	4,6	2,8	3,8	6,9	3,8	10,7	4,7	3,5	6,6	4,2	6,4	2,8	3,0	8,9	5,4	14,4	13,9	4,2	13,1	8,4
CD4 of gdTcells	59,5	50,9	55,2	48,8	53,4	51,0	49,7	46,8	49,1	50,6	54,1	53,2	54,3	52,6	51,5	53,2	56,6	55,7	56,1	64,2
CD8 of gdT-cells	2,5	4,3	2,3	2,2	3,0	2,5	2,7	7,0	1,5	10,3	2,6	2,9	4,4	0,7	1,3	1,2	0,9	3,9	0,9	1,1

Supplementary File S2.2 GraphPad Prism8 analyses on lymphocyte subset in peripheral blood mononuclear cells (PBMC) from healthy Norwegian Red (NR) cows (n=20) grouped by genomic estimated breeding value (GEBV) for lactation average somatic cell score (LSCS) and by lactation stage. Estimated P- and q-value are presented with corresponding discovery range.

Lymphocyte subset	Discovery?[1]	Mean of GEBV for LSCS < 100[2] Given in % (n=10)	Mean % of GEBV for LSCS > 100[3] Given in % (n=10)	P value	q value
Monocytes in Mono gate	No	67,25	68,84	0,64	>0,99
DC-like in Mono gate	No	6,82	5,27	0,65	>0,99
B-cells in Lympho gate	No	16,04	17,47	0,67	>0,99
NK in Lympho gate	No	4,82	4,97	0,97	>0,99
T-cells in Lympho gate	No	28,76	30,48	0,61	>0,99
abT-cells in Lympho gate	No	21,01	25,03	0,24	>0,99
CD4 of abT-cells	No	64,25	62,43	0,59	>0,99
CD25+ of abTcells	No	16,01	15,69	0,92	>0,99
CD8a of abT-cells	No	21,35	20,83	0,88	>0,99
gdT-cells in Lympho gate	No	7,76	5,45	0,49	>0,99
CD4 of gdTcells	No	53,77	52,88	0,79	>0,99

CD8 of gdT-cells	No	2,11	3,71	0,64	>0,99
Lymphocyte subset	Discovery?[1]	Mean of Early lactation[4] Given in % (n=12)	Mean of Late lactation[5] Given in % (n=8)	P value	q value
Monocytes in Mono gate	No	65,78	71,44	0,10	0,72
DC-like in Mono gate	No	4,92	7,74	0,41	0,78
B-cells in Lympho gate	No	15,52	18,61	0,36	0,78
NK in Lympho gate	No	5,51	4,06	0,68	0,84
T-cells in Lympho gate	No	27,78	32,39	0,17	0,72
abT-cells in Lympho gate	No	21,50	25,30	0,26	0,72
CD4 of abT-cells	No	62,33	64,85	0,46	0,78
CD25+ of abTcells	No	16,27	15,23	0,76	0,85
CD8a of abT-cells	No	22,84	18,46	0,20	0,72
gdT-cells in Lympho gate	No	6,28	7,09	0,81	0,85
CD4 of gdTcells	No	53,97	52,36	0,64	0,84
CD8 of gdT-cells	No	3,51	2,01	0,66	0,84

[1] Discovery? – defined significant results, False Discovery Rate (Q) = 5 %

[2] Mean of GEBV for LSCS < 100 – genomic estimated breeding value (GEBV) for lactation average somatic cell score (LSCS) < 100, i.e. cows with high probability of having high somatic cell count during lactation

[3] Mean of GEBV for LSCS > 100 – genomic estimated breeding value (GEBV) for lactation average somatic cell score (LSCS) > 100, i.e. cows with low probability of having high somatic cell count during lactation

[4] Mean of Early lactation – early period of lactation within one – two month after calving

[5] Mean of Late lactation – late period of lactation within seven – eight month after calving

Table S1. Antibodies used for flow cytometry

1. Primary antibodies

Molecule recognized	Mono-/Polyclonal	Clone	Isotype	Conjugation	Typical expression pattern	Manufacturer	Raised against	Panel 1; mono-/granulocytes	Panel 2; myeloid cells	Panel 3; B-cells	Panel 4; NK cells	Panel 5; T-cells
B-B2	M	BAQ44a	mouse IgM	unconjugated	B-cells	VMRD/Kingfisher Biotech	Bovine			x		
CD11c	M	BAQ153A	mouse IgM	unconjugated	Lymphocytes	VMRD/Kingfisher Biotech	Bovine		x			
CD14	M	Tük4	mouse IgG2a	Pacific Blue	Monocytes	Bio-Rad/AbD Serotec	Human	x	x			
CD16	M	KD1	mouse IgG2a	FITC	Monocytes	Bio-Rad/AbD Serotec	Human	x				
CD163	M	Edhu-1	mouse IgG1	unconjugated	Monocyte subsets	Bio-Rad/AbD Serotec	Human	x				
CD172a	M	DH59B	mouse IgG1	unconjugated	Myeloid cells	Bio-Rad/AbD Serotec	Bovine		x			
CD2	M	B26A4	mouse IgM	unconjugated	T-cells and NK-cell subsets	VMRD/Kingfisher Biotech	Bovine				x	
CD21	M	CC51	mouse IgG2b	unconjugated	B-cells	Bio-Rad/AbD Serotec	Bovine			x		
CD25	M	LCTB2a	mouse IgG3	unconjugated	Regulatory T-cells, activated T- and NK cells	VMRD/Kingfisher Biotech	Bovine					x
CD26	M	CACT114A	mouse IgG2b	unconjugated	Lymphocytes	VMRD/Kingfisher Biotech	Bovine		x			
CD3	M	MM1A	mouse IgG1	unconjugated	All T cells	VMRD/Kingfisher Biotech	Bovine				x	x
NKp46/CD335/NCR1	M	AKS6	mouse IgG2b	unconjugated	Natural killer cells	In-house	Bovine				x	
CD4	M	GC50A1	mouse IgM	unconjugated	T cell subset	VMRD/Kingfisher Biotech	Bovine					x
CD8a	M	7C2b	mouse Ig2a	unconjugated	T cell subset	VMRD/Kingfisher Biotech	Bovine				x	
Granulocytes	M	CH138A	mouse IgM	unconjugated	Granulocytes	VMRD/Kingfisher Biotech	Bovine	x				x
IgM	M	BIG73A	mouse IgG1	unconjugated	B-cells	VMRD/Kingfisher Biotech	Bovine			x		
MHC class II	M	IL-A21	mouse IgG2a	FITC	Myeloid cells, B-cells	Bio-Rad/AbD Serotec	Bovine		x	x		
MHC class II	M	TH21A	mouse IgG2b	unconjugated	Myeloid cells, B-cells	VMRD/Kingfisher Biotech	Bovine	x				
TCR1 (gamma-delta receptor)	M	GB21a	mouse IgG2b	unconjugated	T cells, gamma-delta subset	VMRD/Kingfisher Biotech	Bovine					x

2. Secondary antibodies

Goat anti-mouse IgG2a	P			DyLight 405		Jackson Immunoresearch	Mouse					x
Goat anti-mouse Ig2a	P			Alexa 488		Life Technologies	Mouse				x	
Goat anti-mouse Ig3	P			Alexa 488		Life Technologies	Mouse					x
Goat anti-mouse Ig2b	P			PE		Southern Biotechnologies	Mouse	x	x	x	x	x
Goat anti-mouse IgG1	P			APC		Southern Biotechnologies	Mouse	x	x	x	x	x
Goat anti-mouse IgM	M	II/41		APC-efluor780		Affymetrix/eBioscience	Mouse	x	x	x	x	x

Supplementary Table S2. List of designed primers used for reverse transcription-quantitative PCR (RT-qPCR).

Gene no.	Gene symbol, accession no.	Sequence (5`→ 3`)	Product length (bp)
1	RAD17, *checkpoint clamp loader component*, NM_001103244.1	CAGATTGGTCGGTGTCCTCTA GTGCCACAGGGTTGAAACTAA	133
2	ACOT2, *acyl-CoA thioesterase 2*, NM_001101938.1	GATTCTCCAGCCGTCTTCAGTTGAT CTCAGCTTTAGAGATCCTGCCAACT	143
3	ACOT4, *acyl-CoA thioesterase 4*, NM_001098941.2	ATGCTTCAACACAGCCAGATAAAG TGGAAACTGTGGCTGAGATGTT	118
4	FOS, *Fos proto-oncogene, AP-1 transcription factor subunit*, NM_182786.2	CGTCAATGCGCAGGACTACT GGAGACTAGGGTGGGCTGTA	121
5	CXCL1, *chemokine (C-X-C motif) ligand 1*, NM_175700.2	AAGATGGTGTGTTTTAAGTGTTCTTCA CTACTGTCTCTAGCCTTTACTATAGTGGAT	70
6	CXCL8, *C-X-C motif chemokine ligand 8*, NM_173925.2	AACGAGGTCTGCCTAAACCC TGCTTCTCAGCTCTCTTCACAA	77
7	CCNB1, *cyclin B1*, NM_001045872.1	AAGACGGAGCGGATCCAAAC ACTGCTTGCTCTTCCTCAAGT	85
8	CDK7, *cyclin dependent kinase 7*, NM_001075715.1	TTGCAGCAGGAGATGACTTG TTCAAAGCCTGTGTGGCTGTA	88

9	TGFB3, *transforming growth factor beta 3*, NM_001101183.1	TGGCTCTTGCGGAGAGAATC TCATCGCTGTCCACACCTTT	140
10	SCL18A,	CTCCTCACCAACCCGTTCAT CGGGAGAAGGCGAACATAACT	113
11	C6, *complement C6*, NM_001045979.1	AACCCTTTTCACTGTGAGTACCT AGGGCCTCCAGATCCTTAAAC	108
12	SEL1L, *SEL1L ERAD E3 ligase adaptor subunit*, NM_001206908.1	CTTGTGGCCAATCATGTTGCT TCGAGCATTCCGCTGTTCAT	113
13	GLI2, *GLI family zinc finger 2*, NM_001192250.2	GACGCCAATCAGAACAAGCA GAGGACAGCTGGCAGGTTAC	155
14	STAT4, *signal transducer and activator of transcription 4*, NM_001083692.2	AATGAGGGCTGTCACATGGT AGGTAATGAGCAGGTCTCCAA	105

Paper IV

frontiers in Genetics

ORIGINAL RESEARCH
published: 26 July 2019
doi: 10.3389/fgene.2019.00689

Transcription Profiling of Monocyte-Derived Macrophages Infected *In Vitro* With Two Strains of *Streptococcus agalactiae* Reveals Candidate Pathways Affecting Subclinical Mastitis in Cattle

Anna Monika Lewandowska-Sabat[1], Elena Kirsanova[1], Christophe Klopp[2], Trygve Roger Solberg[3], Bjørg Heringstad[3,4], Olav Østerås[5], Preben Boysen[6] and Ingrid Olsaker[1]*

[1] Department of Basic Sciences and Aquatic Medicine, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Oslo, Norway, [2] Sigenae, UBIA INRA, Castanet-Tolosan, France, [3] Geno Breeding and A.I. Association, Hamar, Norway, [4] Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, Ås, Norway [5] Norwegian Cattle Health Services and TINE Extension Services, Ås, Norway, [6] Department of Food Safety and Infection Biology, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Oslo, Norway

OPEN ACCESS

Edited by:
Haja N. Kadarmideen, Technical University of Denmark, Denmark

Reviewed by:
Mairead Lesley Bermingham, University of Edinburgh, United Kingdom
Michelle Martinez-Montemayor, Central University of the Caribbean, Puerto Rico

***Correspondence:**
Anna Monika Lewandowska-Sabat
anna.sabat@nmbu.no

Specialty section:
This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics

Received: *07 January 2019*
Accepted: *02 July 2019*
Published: *26 July 2019*

Citation:
Lewandowska-Sabat AM, Kirsanova E, Klopp C, Solberg TR, Heringstad B, Østerås O, Boysen P and Olsaker I (2019) Transcription Profiling of Monocyte-Derived Macrophages Infected In Vitro With Two Strains of Streptococcus agalactiae Reveals Candidate Pathways Affecting Subclinical Mastitis in Cattle. Front. Genet. 10:689. doi: 10.3389/fgene.2019.00689

Macrophages are key cells of innate immune response and serve as the first line of defense against bacteria. Transcription profiling of bacteria-infected macrophages could provide important insights on the pathogenicity and host defense mechanisms during infection. We have examined transcription profiles of bovine monocyte-derived macrophages (bMDMs) isolated from the blood of 12 animals and infected *in vitro* with two strains of *Streptococcus agalactiae*. Illumina sequencing of RNA from 36 bMDMs cultures exposed *in vitro* to either one of two sequence types of *S. agalactiae* (ST103 or ST12) for 6 h and unchallenged controls was performed. Analyses of over 1,656 million high-quality paired-end sequence reads revealed 5,936 and 6,443 differentially expressed genes ($p < 0.05$) in bMDMs infected with ST103 and ST12, respectively, versus unchallenged controls. Moreover, 588 genes differentially expressed between bMDMs infected with ST103 versus ST12 were identified. Ingenuity pathway analysis of the differentially up-regulated genes in the bMDMs infected with ST103 revealed significant enrichment for granulocyte adhesion and diapedesis, while significant enrichment for the phagosome formation pathway was found among down-regulated genes. Moreover, Ingenuity pathway analysis of the differentially up-regulated genes in the bMDMs infected with ST12 showed significant enrichment for type 1/type 2 T helper cell activation, while the complement activation pathway was overrepresented in the down-regulated genes. Our study identified pathogen-induced regulation of key genes and pathways involved in the immune response of macrophages against infection but also likely involved in bacterial evasion of the host immune system. These results may contribute to better understanding of the mechanisms underlying subclinical infection such as bovine streptococcal mastitis.

Keywords: cattle, macrophages, pathway analysis, RNA sequencing, *Streptococcus agalactiae*, subclinical mastitis

INTRODUCTION

Monocytes and macrophages are critical cells associated with innate immunity, regulation of inflammation, and host defense against invading pathogens. Macrophages are capable of phagocytosis, a major immune mechanism used to remove pathogens, and local recruitment and action of macrophages in the mammary gland is an essential immunological defense mechanism against infection such as mastitis (Politis et al., 1991; Barber and Yang, 1998; Riollet et al., 2001; Alluwaimi et al., 2003). Intramammary infections will trigger macrophages to produce not only pro-inflammatory cytokines required to eliminate pathogens but also anti-inflammatory factors essential for immune regulation of the inflammation and preventing chronic conditions (Gunther et al., 2016).

Mastitis is a multi-factorial disease, affected by load and virulence of infecting pathogens and host genotype. Some cows develop chronic subclinical mastitis with high somatic cell count (SCC) in milk. Such animals may shed the bacteria and contribute to infection spreading to other cows and herds. However, the mechanisms that underlie the pathogenesis of subclinical mastitis are not well understood.

In the last decade, genomic selection has been implemented in animal breeding worldwide. Large genetic differences exist in cows' susceptibility to mastitis (Heringstad et al., 2007). Routine genetic evaluation of udder health of Norwegian Red is based on information both on veterinary treatment of clinical mastitis and on SCC. SCC is an indicator of mastitis widely used in genetic evaluation of udder health (Miglior et al., 2017). Genomic breeding values (GEBV) for SCC can be used to identify cows with different susceptibility to chronic subclinical mastitis.

The first objective of this study was to examine the genome-wide transcription profiles of primary bovine monocyte-derived macrophages (bMDMs) infected *in vitro* with two *Streptococcus agalactiae* strains, one of the leading causative agents of subclinical bovine mastitis worldwide. Strain ST103 is the most prevalent bovine *S. agalactiae* isolate found in farms with considerable environmental contamination, while ST12, a strain usually associated with colonization of pregnant women, has been found in cattle herds with no positive environmental samples (Manning et al., 2009; Jorgensen et al., 2016). The second objective was to examine differences in transcript profiles of bMDMs isolated from two groups of animals differing in their genomic (low vs high GEBV for SCC) and phenotypic (high vs low SCC) characteristics.

To the best of our knowledge, this study is the first report of genome-wide transcriptome profiling of bMDMs infected with *S. agalactiae*. The identification of key differentially expressed genes and pathways in infected bMDMs provides a basis for improved understanding of central mechanisms of host defense during subclinical infections such as mastitis.

MATERIAL AND METHODS

Animals and Cell Isolation

Twelve healthy Norwegian Red cows aged 4–8 years were used in the experiment. The cows came from three Norwegian herds, and both animals with low (total $n = 6$) and high (total $n = 6$) GEBV were selected from each farm to avoid potential farm induced bias. Among these 12 animals, three with high and nine with low SCC phenotype were found (for details, see **Supplementary Table 1**).

Five hundred milliliters of blood was collected from the neck of each animal ($n = 12$) in sterile glass bottles with sodium citrate as anticoagulant. Peripheral blood mononuclear cells were extracted as described earlier (Lewandowska-Sabat et al., 2013). Briefly, density gradient centrifugation (2,210 × g, 30 min) on lymphoprep (Axis-Shield, Norway) was used, and CD14+ cells were isolated by positive selection of monocyte differentiation antigen CD14 using anti-human CD14 MACS MicroBeads (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany), according to the manufacturer's instructions. Purity of selected cells was verified by flow cytometry by staining positively selected cells with PE-conjugated anti-mouse IgG2 (Southern Biotech, Birmingham, AL, USA), analyzing in a Gallios flow cytometer (Beckton Dickinson), and found to be in the range of 95–98%. The CD14+ cells were subsequently grown in six-well dishes at a density of 1.5×10^6 cells per well in RPMI medium supplemented with 10% FCS (Invitrogen, Carlsbad, USA). Cells were left over night at 37°C in an atmosphere with 5% CO_2. The phenotypic morphology of cells, i.e., differentiation of monocytes into an early-stage adherent macrophage phenotype, was visualized and confirmed by phase contrast microscopy.

Bacterial Infection

Two previously described *S. agalactiae* strains (ST103 and ST12) derived from Norwegian dairy herds were obtained from The Norwegian Veterinary Institute (Jorgensen et al., 2016). These bovine adapted strains were isolated from milk samples. Bacteria were collected from blood agar plates and grown in "Todd Hewitt broth" (Sigma-Aldrich) until mid-log phase as described previously (Lewandowska-Sabat et al., 2018a). Briefly, growth was measured by optical density at 600 nm. The mid-log phase cultures were aliquoted and frozen at −70°C as 20% glycerol stocks. The final number of colony-forming units (CFU) was determined by serial dilutions and plating on blood agar plates. Bacteria used in this study all came from aliquots of the same batch.

For each individual animal ($n = 12$), the wells with bMDMs were grouped into three classes with an equal number of wells and cells per class as possible, constituting in total 36 different samples. Two classes were infected with either ST103 or ST12 using a multiplicity of infection of 1 (1 bacterium per cell, on average). The third cell class was left uninfected (negative control). After 1 h of exposure, 1% of penicillin/streptomycin (60 pg/ml penicillin and 100 μg/ml streptomycin) were added to prevent growth of the remaining extracellular bacteria. The controls and the infected cells were treated equally. Inhibition of bacterial growth by antibiotics was verified by microscopy. Incubation was continued for five more hours for both bacteria-infected cells and negative controls for a total of 6-h incubation (i.e., 6-h infection in the samples exposed to bacteria). Medium was aspirated, and the cells were collected using a cell scraper. Cells were centrifuged (400 × g, 5 min), and the pellet was

washed with cold PBS buffer, snap frozen in liquid nitrogen, and stored at −70°C.

RNA Extraction

Thirty-six different samples of cells, i.e., three classes (control, ST103-infected, and ST12-infected), from each of the 12 animals, were used to extract RNA. Total RNA was isolated from control and infected cells using the MirVANA isolation kit (Ambion, Austin, TX) following the manufacturer's instructions. All RNA samples were treated with amplification grade DNase I (Invitrogen) to remove any traces of genomic DNA. RNA concentration and quality was measured using NanoDrop 1000 (Thermo Fisher Scientific, Wilmington, USA) and 2100 BioAnalyzer with Agilent RNA 6000 Nano kit (Agilent Technologies, Palo Alto, USA), respectively. The RNA integrity numbers (RIN), concentrations, and optical density A260/A280 ratios are listed in **Supplementary Table 2**.

RNA Sequencing and Data Analyses

The 36 RNA samples (control = 12, ST12 = 12, ST103 = 12; **Supplementary Table 3**) extracted above were subsequently submitted for RNA deep sequencing. RNA-seq libraries were prepared from 500 ng of total RNA using TruSeq Stranded mRNA prep kit with polyA enrichment (Illumina), according to the manufacturer's protocol. Libraries were sequenced (150 bp paired-end, 250–300 M reads per end) on five lanes with an Illumina HiSeq 3000 machine (Illumina). Preparation of RNA-seq libraries and sequencing were performed by the Norwegian Sequencing Centre (Oslo, Norway; http://www.sequencing.uio.no/).

The read quality of the RNA-seq libraries was evaluated using FastQC software ver. 0.11.2 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Cutadapt ver. 1.8.3 (http://www.cutadapt/) was then used to trim adaptor sequences. Reads were mapped against the bovine genome assembly v.3.1. (UMD3.1) using the STAR aligner (v.2.3.1y) (Dobin et al., 2013).

The "rsem prepare reference" script of RSEM package was used to generate reference transcript sequences by using gene annotation file (GTF format) and the full genome sequence (FASTA format) of Bos taurus v.3.1. The calculation of relative transcript abundances in each sample was performed by the "rsem calculate expression" script of the RSEM package (Li and Dewey, 2011).

Differential Expression Analysis

In order to identify statistically significant differentially expressed (DE) transcripts between negative control and treatment samples, DESeq2 ver. 1.4.5. (Love et al., 2014) and edgeR ver.3.20.9 (Robinson and Oshlack, 2010) were applied in R ver.3.4.1 (http://www.R-project.org). The gene raw counts obtained from RSEM were used as an input, and genes with low expression were removed with a minimal set threshold of one count per million in edgeR. The analyses were performed pairwise, i.e., negative control with one treatment group (ST103 or ST12, respectively). In addition, a comparison between the two bacterial treatment groups (ST103 and ST12) was performed.

Furthermore, a pairwise comparison between low ($n = 6$) and high ($n = 6$) GEBV animals was performed for each treatment (e.g., negative control low GEBV vs negative control high GEBV, ST103 low GEBV vs ST103 high GEBV, etc). We have also performed a pairwise comparison between high ($n = 3$) and low ($n = 9$) SCC animals for each treatment. DE genes were considered significant when both the false discovery rate (FDR) was <0.05 in edgeR and the Benjamini and Hochberg corrected P-value (adjusted) was <0.05 in DESeq2 (i.e., only genes that were identified by both methods as significantly DE are presented in the results and analyzed further). In order to determine how well samples cluster together based on the similarity of their overall RNA expression profiles, hierarchical clustering of the samples was performed in R using "pheatmap" package ver.1.0.8. The heatmap of all samples is presented as **Supplementary Figure 1**.

Pathway Analysis

Significant DE genes identified both by edgeR and DESeq2, and with a log fold change (logFC) either ≤−1.5 for down-regulated or ≥1.5 for up-regulated as estimated by edgeR were used in constructing Venn diagrams (http://jvenn.toulouse.inra.fr/app/index.html) and for Ingenuity Pathway Analysis (IPA; http://www.ingenuity.com). The list of genes found to be DE between negative controls and ST103- or ST12-infected bMDMs, and between ST103- and ST12-infeceted bMDMs, was used as input in IPA in order to identify canonical pathways, biological functions, and networks overrepresented in the datasets. The down- and up-regulated genes were analyzed separately. Benjamini and Hochberg corrected *P*-value ≤ 0.05 was defined as significant in IPA.

RESULTS

Summary Statistics for the RNA-seq Data

For each library, 31–69 million paired-end reads were generated during the sequencing run. Preliminary quality control of the resulting 36 fastq files revealed that all libraries passed the quality criteria with a Phred score > 32. The sequencing resulted in 1,656 million high-quality paired-end reads, and 1,648 million sequences (99.5%) remained after adaptor trimming. A mean of 86.3% of the reads per RNA-Seq sample aligned to unique locations in the Bos taurus v.3.1 genome. The total number of read pairs per sample, the number and percentage of reads passing filters, and the percentage of uniquely mapped reads for each sample are presented in **Supplementary Table 2**. The datasets analyzed during the current study have been deposited at EMBL-EBI (https://www.ebi.ac.uk/ena/) under study accession number PRJEB24166 (Lewandowska-Sabat et al., 2018b).

Differentially Expressed Genes in Macrophages Challenged With *S. agalactiae*

The difference between the negative control and the ST103- or ST12-infected samples was confirmed by hierarchical clustering (**Supplementary Figure 1**). All control samples clustered together and separately from both the bacterial strain-treated samples.

Analyses of differential expression revealed 5,936 significant DE genes in bMDMs challenged with ST103 and 6443 DE genes

in bMDMs challenged with ST12, compared to the respective unchallenged negative controls (FDR < 0.05 and $p < 0.05$). Moreover, 588 significant DE genes were detected when bMDMs challenged with ST12 were compared to bMDMs challenged with ST103 (FDR < 0.05 and $p < 0.05$; **Supplementary Table 4**).

The analyses of differential expression between low and high GEBV animals revealed one significant DE gene in bMDMs challenged with ST103, one significant DE gene in bMDMs challenged with ST12, and no significant genes in unchallenged negative controls. Furthermore, the analyses of differential expression between low and high SCC animals revealed six significant DE genes in bMDMs challenged with ST103, three significant DE genes in bMDMs challenged with ST12, and two significant DE genes in unchallenged controls (**Supplementary Table 5**).

Enriched Pathways of Differentially Expressed Genes

In order to identify the most biologically relevant DE genes, those with logFC either ≤−1.5 for down-regulated or ≥1.5 for up-regulated genes were identified and used in IPA (**Supplementary Table 4**). Venn diagrams for up-regulated and down-regulated genes identified above were constructed (**Figure 1**; **Supplementary Table 6**). The top significantly enriched canonical pathways identified by IPA are presented in **Table 1**.

Analysis of genes that were up-regulated in response to ST103 infection ($n = 809$) revealed that one of the top biological networks with 27 focus molecules was associated with cellular development, cellular growth and proliferation, hematological system development, and function (**Figure 2A**). Furthermore, *Interleukin 1* (*IL-1*) was identified as one of the top regulators affecting the viability of leukocytes (**Figure 3A**).

Analysis of ST103 down-regulated genes ($n = 882$) showed that one of the top biological networks with 30 focus molecules was associated with cell morphology, cellular assembly and organization, cellular function, and maintenance (**Figure 2B**). Moreover, top regulators affecting movement of leukocytes were identified among ST103 down-regulated genes (**Figure 3B**).

Analysis of genes that were up-regulated in response to ST12 infection ($n = 934$) revealed that the top biological networks with 32 focus molecules was associated with cancer, organismal injury and abnormalities, and lipid metabolism (**Figure 4A**). Furthermore, *angiotensinogen* (*AGT*) was identified as one of the top regulators affecting recruitment of leukocytes (**Supplementary Figure 2**).

Analysis of ST12 down-regulated genes ($n = 845$) showed that one of the top biological networks with 19 focus molecules was associated with cell-to-cell signaling and interaction, cellular movement, hematological system development, and function (**Figure 4B**). Moreover, top regulators affecting movement of antigen-presenting cells were identified for ST12 down-regulated genes (**Supplementary Figure 3**).

Analysis of genes up-regulated in ST12 compared to ST103 ($n = 42$) revealed that the top biological network with 13 focus molecules was associated with cell-mediated immune response, cellular development, cellular function, and maintenance (**Figure 5A**). Moreover, the top regulators involved in apoptosis of lymphocytes and inflammatory response were identified for these up-regulated genes (**Supplementary Figure 4**).

Analysis of genes down-regulated in ST12 compared to ST103 ($n = 18$) revealed that the top biological network with eight focus molecules was associated with hematological system development and function, immune cell trafficking, and inflammatory response (**Figure 5B**). There were no top regulators identified for these genes.

All detailed information from IPA are presented in **Supplementary Table 7**.

DISCUSSION

Transcription profiles of macrophages infected with several types of bacteria and bacterial and viral ligands have been studied earlier (Lewandowska-Sabat et al., 2013; Nalpas et al., 2013; Casey et al., 2015; Mavromatis et al., 2015; Hop et al., 2017; Roy et al., 2018; Toca et al., 2019). To the best of our knowledge, this is the first study of the early-phase transcriptome of bMDMs infected *in vitro* with two strains of *S. agalactiae*. *S. agalactiae* is not a main causative agent of subclinical mastitis in Norway, but its occurrence has increased in modern freestalls and automatic milking systems the last years. The agent is also a main cause of life-threatening infections in human neonates, pregnant females, and elderly adults (Henneke and Berner, 2006). The two strains used in this study vary in their capability to survive in the environment and transmit within dairy herds. ST103 is the most prevalent and persistent isolate found in bovine herds with substantial environmental contamination, while ST12 was found in cattle herds with no positive environmental samples (Jorgensen et al., 2016). We have analyzed over 1,648 million reads and found several thousand transcripts that were significantly differentially regulated in response to infection (**Supplementary Table 4**). In order to identify the most biologically relevant genes, the logFC threshold was set to ≥1.5 for up-regulated and ≤−1.5 for down-regulated transcripts (**Figure 1**), and pathways analyses of these transcripts were performed.

We observed that ST103 induced a different transcript profile during the early stage of macrophage infection compared to ST12 (**Figure 1** and **Table 1**). The observed differences may be explained by differences in virulence factors and/or the different niches and modes of transmission of these strains. It has been demonstrated that specific sequence types of bacteria often associate with specific virulence factors, which, in turn, are associated with the outcome of infection in the host (Tassi et al., 2015; Budd et al., 2016). The virulence factors in strain ST103 are associated with growth ability in milk, biofilm formation, and adhesion to bovine mammary epithelial cells as well as higher cytotoxicity compared to most of the other bovine strains (Pang et al., 2017). On the other hand, the characterization of ST12 revealed that this strain can survive within macrophages (Korir et al., 2017).

Analyses of global transcriptomics and gene networks have been performed earlier in livestock in order to reveal candidate genes and pathways (Kogelman et al., 2014; Drag et al., 2017). One of the top canonical pathways for the up-regulated transcripts

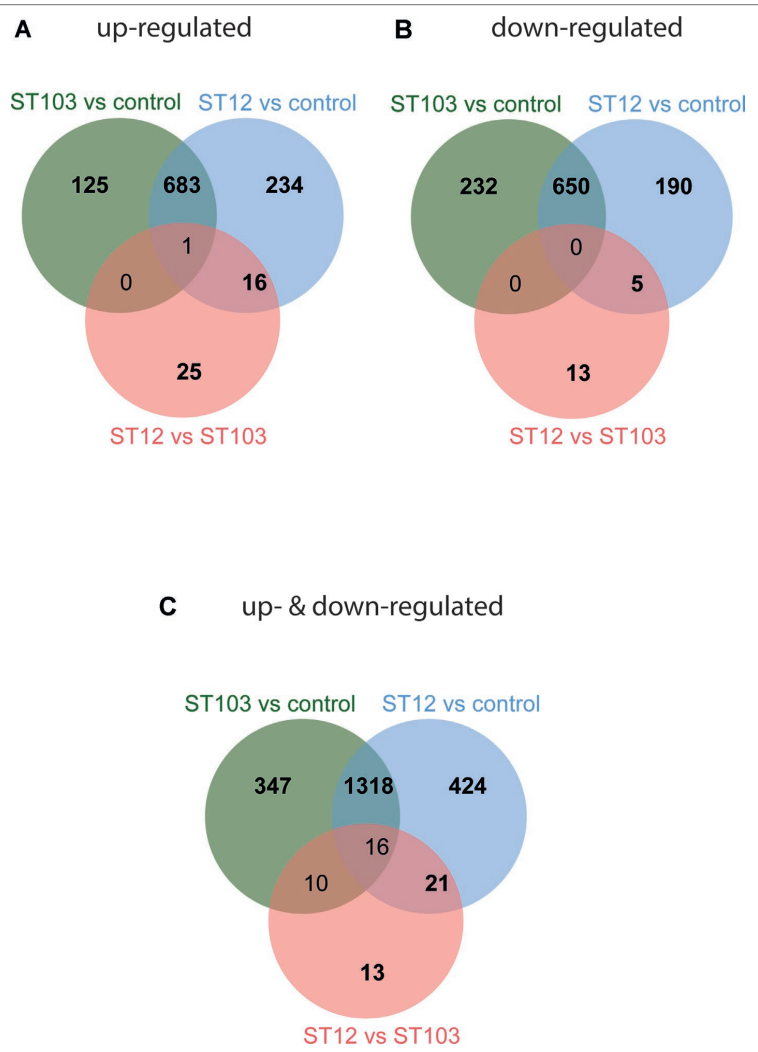


FIGURE 1 | Venn diagrams of differentially expressed (DE) genes in bovine monocyte-derived macrophages (bMDMs) challenged with *Streptococcus agalactiae* strain ST103 or strain ST12, compared to the respective unchallenged negative controls, and in bMDMs challenged with strain ST12 compared to strain ST103 (FDR < 0.05). **(A)** Genes that were only up-regulated: log fold change (logFC) ≥ 1.5. **(B)** Genes that were only down-regulated: logFC ≤ –1.5. **(C)** Genes that were both up- and down-regulated: logFC ≥ 1.5 and logFC ≤ –1.5.

in ST103-infected macrophages was granulocyte adhesion and diapedesis (**Table 1**). During host defense against infection, the polymorphonuclear neutrophils and macrophages are rapidly recruited to sites of bacterial invasion. This multi-step process involves an action of several chemokines and their ligands mediating the initiation of contact between granulocytes and endothelial cells. Integrins are cell adhesion molecules involved in this process, and many bacterial species, such as group A streptococci (Ozeri et al., 2001), *S. aureus* (Agerer et al., 2005), and *E. coli* (Plancon et al., 2003), use host integrins for adhering to and invading host cells. Furthermore, a recent study has identified virulence factors in strain ST103 of *S. agalactiae* associated with ability to adhere to bovine mammary epithelial cells (Pang et al., 2017). Interestingly, a recent study from our group showed that target genes of microRNAs that were down-regulated in ST103-infected bovine macrophages were associated

TABLE 1 | Top significant pathways overrepresented among differentially expressed (DE) genes in response to *in vitro* exposure of bovine monocyte-derived macrophages (bMDM) to *Streptococcus agalactiae* strain ST103 or strain ST12, as compared to negative controls.

		Pathway name	*P*-value	Genes
ST103 vs control	Up-regulated (809)	Differential Regulation of Cytokine Production in Macrophages and T Helper Cells by IL-17A and IL-17F	3,44E–11	IL9, CSF2, CCL5, TNF, CCL2, IL17A, IL10, IL12B, CSF3, CCL4
		Granulocyte Adhesion and Diapedesis	5,46E–10	GNAI1, CCL8, CLDN4, CCL20, HRH1, CCL3L3, SDC1, IL1RN, TNF, ITGB3, CCL4, SDC3, IL33, CCL5, CXCL10, EZR, CXCL2, HSPB1, CCL22, CCL2, CSF3, SDC4, ICAM1, IL36G, PECAM1
		Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F	7,78E–10	IL9, CSF2, CCL5, TNF, CCL2, IL17A, IL10, IL12B, CSF3, CCL4
	Down-regulated (882)	IL-8 Signaling	9,28E–08	PLCB2, GNG7, PLD1, ARRB2, IRAK3, VCAM1, NCF2, FOS, PIK3CD, RND3, VEGFC, PRKCE, LIMK2, ITGB5, ANGPT2, FLT4, PRKD3, PLD4, MYL9, PIK3CG, PIK3R2, CCND1, CXCR2, PLD2, CYBB
		Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	3,49E–06	PRKCE, OAS1, IL12A, TLR7, PRKD3, NOD1, OAS2, C3AR1, C1QA, PIK3CG, C1QC, IL18, PIK3CD, TLR3, PIK3R2, C5AR1, CLEC6A, CLEC7A
		Phagosome Formation	2,76E–05	PRKCE, TLR7, PRKD3, PLCB2, MRC2, MSR1, C3AR1, PIK3CG, INPP5D, PIK3CD, TLR3, PIK3R2, C5AR1, ITGA4, RND3, CLEC7A
ST12 vs control	Up-regulated (934)	Th1 and Th2 Activation Pathway	9,08E–12	PRKCQ, JAK3, IL12RB2, GAB1, CD274, IL27, DLL4, IL9, STAT1, CD3D, STAT5A, IL12B, CRLF2, NFIL3, IL33, IL27RA, S1PR1, CD3E, STAT4, PIK3R3, JAG1, LTA, IL2RA, HLA-DOB, NFATC1, TNFSF4, DLL1, IL10, ICAM1, TBX21
		Differential Regulation of Cytokine Production in Macrophages and T Helper Cells by IL-17A and IL-17F	1,52E–10	CSF2, IL9, CCL5, TNF, CCL2, IL17A, IL12B, IL10, CSF3, CCL4
		Role of Hypercytokinemia/Hyperchemokinemia in the Pathogenesis of Influenza	2,88E–09	CCL5, CXCL10, IL37, CXCL8, IL9, IL1RN, TNF, CCL2, IL17A, IL12B, CCL4, IL36G, IL33
	Down-regulated (845)	IL-8 Signaling	5,55E–05	ITGB5, LIMK2, ANGPT2, FLT4, PLCB2, GNG7, MYL9, PLD4, PLD1, ARRB2, VCAM1, NCF2, FOS, PIK3CD, PIK3R2, CCND1, CXCR2, PLD2, RND3
		Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	7,97E–05	OAS1, TLR7, C1QB, NOD1, OAS2, C3AR1, C1QA, C1QC, IL18, PIK3CD, TLR3, C5AR1, PIK3R2, CLEC6A, CLEC7A
		Complement System	2,36E–04	C1QC, C5AR1, C1QB, C7, CFD, C3AR1, C1QA
ST12 vs ST103	Up-regulated (42)	IL-12 Signaling and Production in Macrophages	1.30E–04	FOS, REL, PIK3R3, ALOX15
		Prolactin Signaling	4.36E–04	FOS, PIK3R3, TCF7
		Rac Signaling	1.18E–03	PIK3R3, CYFIP2, ABI2
	Down-regulated (18)	Thiosulfate Disproportionation III (Rhodanese)	1.78E–03	TST
		Th2 Pathway	3.46E–03	TNFSF4, IL2RA
		Granulocyte Adhesion and Diapedesis	4.89E–03	CCL8, CLDN4

The number in parentheses represents the number of DE genes used for Ingenuity Pathway Analyses (IPA).

with the integrin signaling pathway (Lewandowska-Sabat et al., 2018a). These findings, together with the present results, may indicate that granulocyte adhesion is crucial for invasion of macrophages or mammary epithelial cells by *S. agalactiae* strain ST103. Furthermore, one of the most central molecules in the network associated with the response to strain ST103 was *interleukin-9* (*IL-9;* **Figure 2A**). This cytokine stimulates cell proliferation and prevents apoptosis, is associated with type 2 immune responses, and has been identified as a susceptibility factor in bacterial infection (Arendse et al., 2005; Wynn, 2015). In recent years, type 2 immune response has emerged as a major effector response either acting in a host protective manner or having pathogenic activity (Wynn, 2015). It has been shown that type 2 cytokine responses suppress the development of protective type 1 immunity to bacterial pathogens and facilitate uncontrolled or persistent infection (Erb et al., 1998; Harris et al., 2007). Although type 2 immune response enables tissue repair, prolonged activation of this pathway may contribute to the development of pathological conditions as it is observed in many chronic fibroproliferative diseases (Wynn and Ramalingam, 2012). We have also identified several molecules associated with cell viability of leukocytes (**Figure 3A**). It has been demonstrated recently that *S. agalactiae* survive intracellularly inducing injury to murine macrophages (Guo et al., 2014) and that an increased expression of anti-apoptotic genes in *Staphylococcus aureus* infected human macrophages contributes to extended phagocyte lifetime allowing intracellular bacterial survival (Koziel et al., 2009). These observations, together with our findings, may suggest that decreased apoptosis and amplified type 2 immune response may likely be an evasion strategy developed by

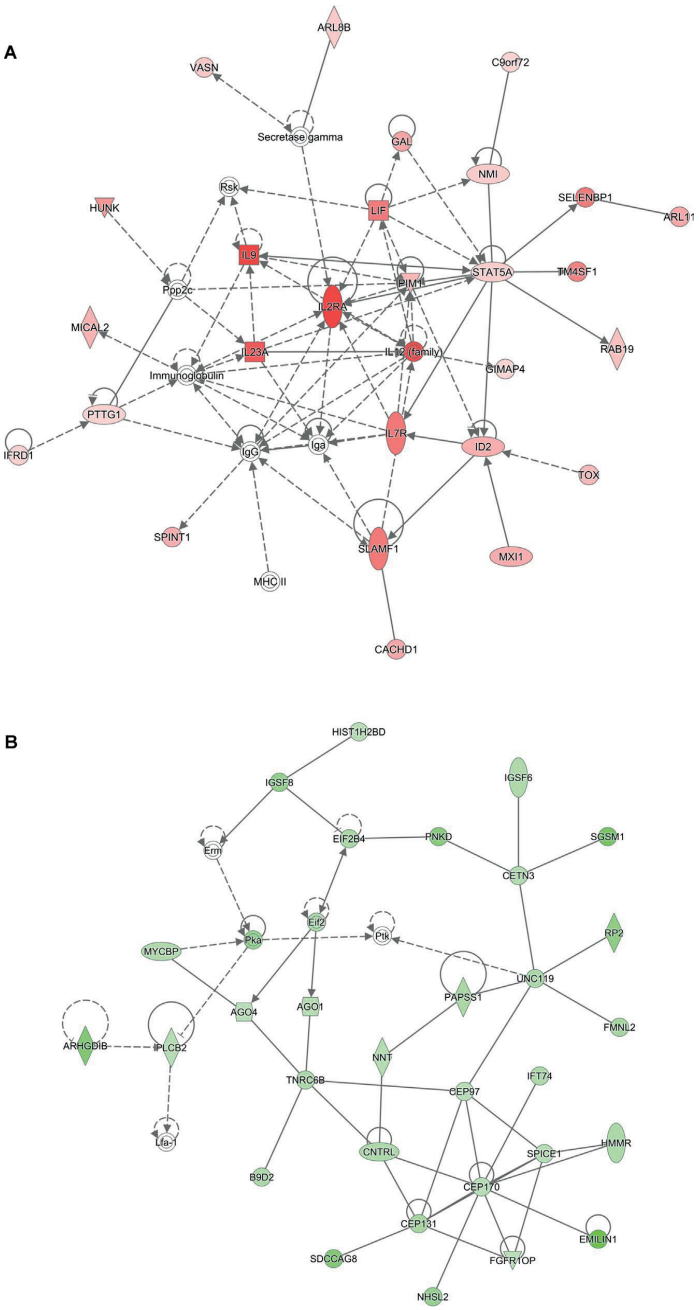


FIGURE 2 | Functional networks overrepresented in the list of DE mRNA in response to infection of bMDMs with *Streptococcus agalactiae* strain ST103 compared to uninfected controls. Ingenuity Pathway Analysis (IPA) identified 27 associated molecules with a network score of 37 for **(A)** up-regulated genes and 30 associated molecules with a network score of 41 for **(B)** down-regulated genes.

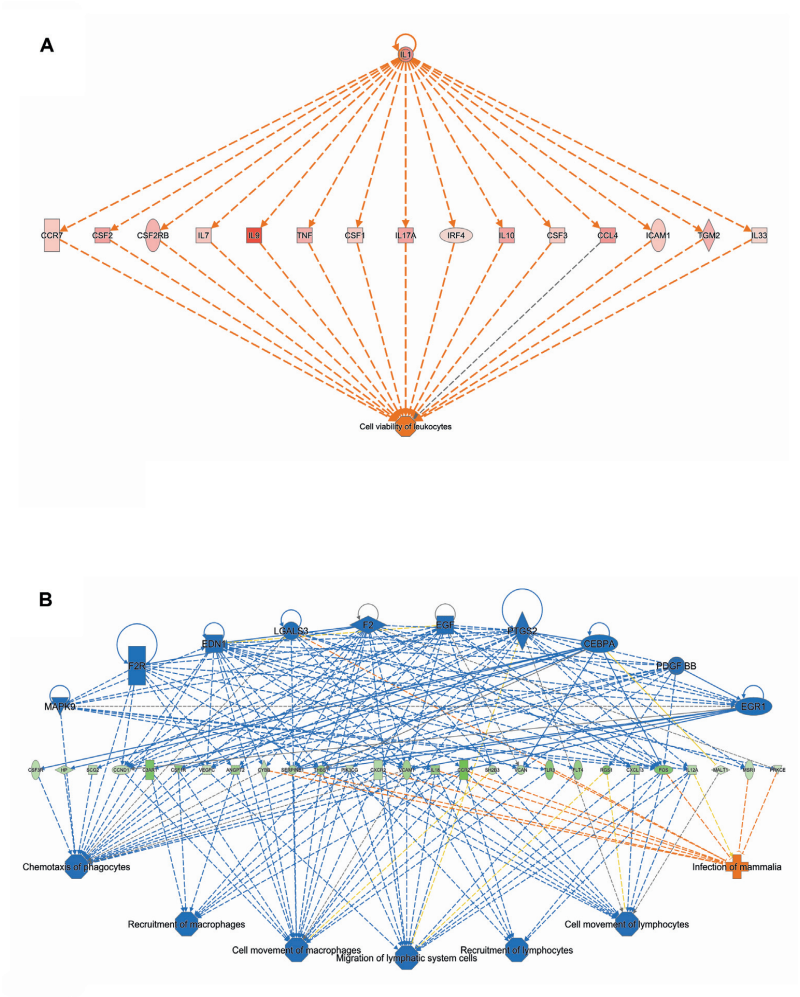


FIGURE 3 | The top regulators identified by IPA **(A)** for genes up-regulated and **(B)** down-regulated in response to infection of bMDMs with *Streptococcus agalactiae* strain ST103 compared to uninfected controls.

S. agalactiae strain ST103 to avoid host defense and promote its intracellular replication and persistence.

One of the top canonical pathways for the down-regulated transcripts in ST103-infected macrophages was phagosome formation (**Table 1**). In addition, cellular assembly and organization was identified as one of the most significant networks associated with the down-regulated transcripts (**Figure 2B**). Phagocytosis is driven by a tightly controlled rearrangement of the actin cytoskeleton to form the phagosome upon recognition of pathogens, and it has been shown that some bacteria, e.g. *Mycobacterium tuberculosis* (Wong et al., 2013), *Listeria monocytogenes* (Portnoy et al., 2002), and *Group B Streptococcus* (Korir et al., 2017), can exploit phagocytosis as an evasion strategy. Furthermore, interleukin-8 (IL-8) signaling pathway (**Table 1**) and movement of leukocytes (**Figure 3B**) were recognized as the most significant pathways and regulator effects, respectively, for the ST103 down-regulated transcripts in bMDMs. It has been shown that monocyte recruitment is critical for immune defense against a broad range of pathogens (Peters et al., 2001; Pamer, 2004; Copin et al., 2007). These results suggest that inhibition of chemotaxis of other immune cells, e.g., neutrophils to the site of infection, may be used by strain ST103 in order to either evade or suppress host

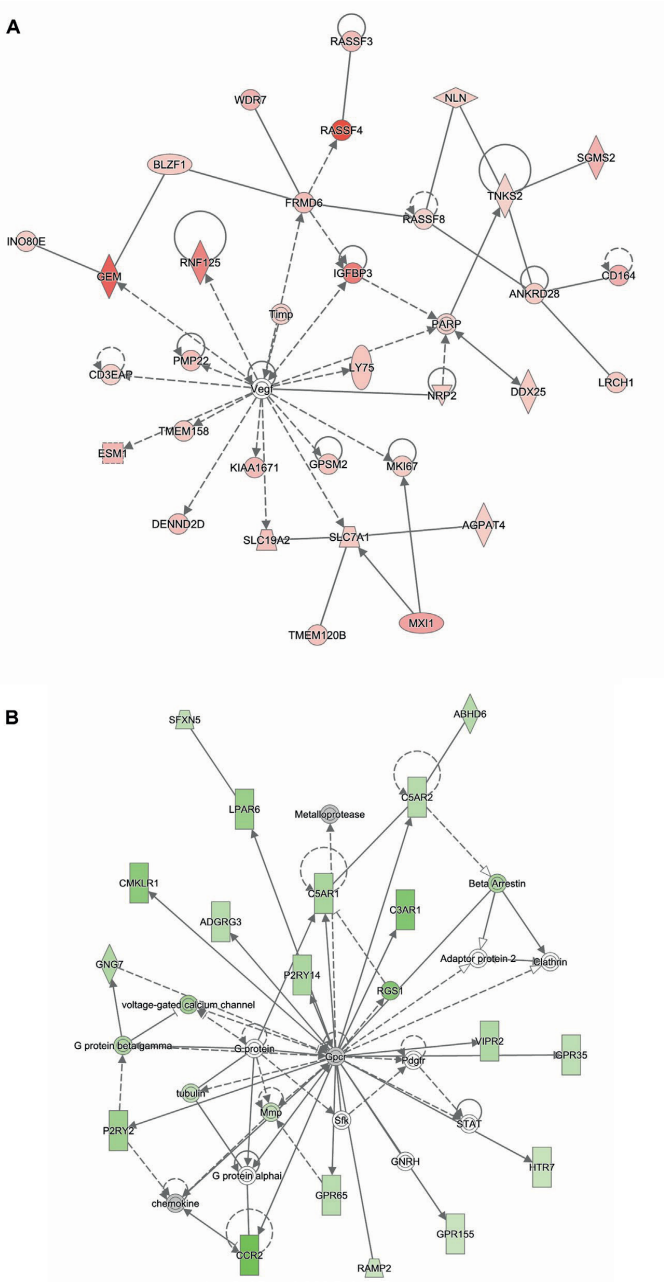


FIGURE 4 | Functional networks overrepresented in the list of DE mRNA in response to infection of bMDMs with *Streptococcus agalactiae* strain ST12 compared to uninfected controls. IPA identified 32 associated molecules with a network score of 46 for **(A)** up-regulated genes and 19 associated molecules with a network score of 21 for **(B)** down-regulated genes.

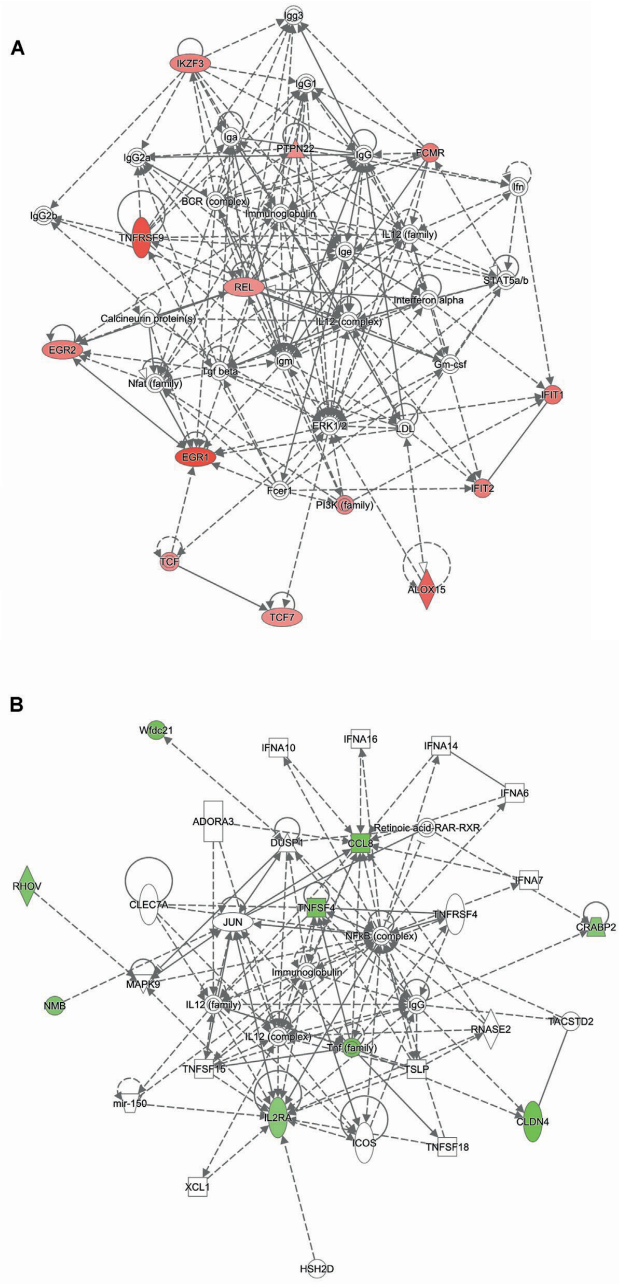


FIGURE 5 | Functional networks overrepresented in the list of DE mRNA in response to infection of bMDMs with *Streptococcus agalactiae* strain ST12 compared to infection with strain ST103. IPA identified 11 associated molecules with a network score of 23 for **(A)** up-regulated genes and eight associated molecules with a network score of 20 for **(B)** down-regulated genes.

immune response. These hypotheses need to be confirmed, however, by functional studies.

The most significant canonical pathway identified among the up-regulated transcripts in ST12-infected macrophages was the type 1/type 2 T helper cell (Th1/Th2) activation pathway (**Table 1**). In addition, many genes associated with the lipid metabolism network were identified (**Figure 4A**). While M1 macrophages produce *IL-12* and constitute the first line of defense against pathogens, M2 macrophages appear as a favorable environment for long-term persistence of many pathogens (reviewed in Muraille et al., 2014). Several studies suggest that some intracellular bacteria such as *Brucella abortus* and *Mycobacterium tuberculosis* responsible for chronic infections actively manipulate signal transducer and activator of transcription 6 (STAT6)–peroxisome proliferator activated receptor gamma/delta (PPARγ/δ) pathways to circumvent M1 macrophage polarization and benefit from a nutrient-rich niche associated with lipid metabolism (Jensen et al., 2011; Xavier et al., 2013; Almeida et al., 2014). In our previous study, we demonstrated that both bacterial strains regulate miRNAs involved in the alternative activation of macrophages (Lewandowska-Sabat et al., 2018a). This may suggest that both strains of *S. agalactiae* may induce alternative macrophage activation; however, ST12 seems to contribute to M2 polarization to a larger extent than ST103. This result, however, needs to be confirmed by more comprehensive functional studies.

Complement system activation was one of the top canonical pathways identified among the down-regulated transcripts in the ST12-infected macrophages (**Table 1**). The complement system is a central component of innate immunity, produced by several cellular sources, including macrophages (Lubbers et al., 2017). Several microorganisms have developed many ways to evade complement actions (reviewed in Lambris et al., 2008). Several of the genes of the complement system were down-regulated in bMDMs in response to ST12 infection (**Table 1** and **Figure 4B**), which indicate that this may be one of the mechanisms developed by ST12 to escape the attack of complement.

Analysis of differential gene response between ST103- and ST12-infected bMDMs revealed that IL-12 signaling and production in macrophages (**Table 1**) and gene network associated with cell-mediated immune response (**Figure 5A**) were significantly up-regulated in ST12 compared to ST103. *Fos proto-oncogene* (*FOS*), *early growth response 1* (*EGR1*), and *early growth response 2* (*EGR2*) are transcription factors known as immediate-early genes (IEGs), i.e., they are able to respond very quickly to regulatory signals, for example, in immune responses (reviewed in Bahrami and Drablos, 2016). It has been demonstrated that stimulation of mouse bone marrow-derived macrophages with lipopolisacharide leads to strong induction of these genes (Ramirez-Carrozzi et al., 2009). *FOS* plays a key role in proliferation, differentiation, and survival of the cell (O'Donnell et al., 2012), and *EGR1* promotes expression of *IL-8*, *interleukin-6* (*IL-6*), and *tumor necrosis factor* (*TNF*) by binding to their promoter region (de Klerk et al., 2017). *FOS*, *EGR1*, and *EGR2* are significantly up-regulated in ST12- compared to ST103-infected bMDMs. These results may suggest that infection of bMDMs by ST12 promotes gene responses crucial for bacterial intracellular survival. Moreover, suppression of these transcription factors in ST103-infected cells may possibly constitute a strategy to evade an early immune response and favor persistence of this *S. agalactiae* strain.

One of the top canonical pathways identified for down-regulated genes in ST12- compared to ST103-infected bMDMs was granulocyte adhesion and diapedesis (**Table 1**). This pathway was also significantly overrepresented among up-regulated genes in ST103-infected bMDMs compared to uninfected controls, which suggests that ST103 modifies expression of genes involved in cell adhesion in order to invade the macrophages.

Myosin IF (*MYO1F*) was found to be significantly up-regulated in ST103-infected bMDMs from animals with low compared to high GEBV for SCC (**Supplementary Table 5**). It has been demonstrated recently that *MYO1F* plays an important role in neutrophil trafficking during inflammation by regulating neutrophil extravasation and migration (Salvermoser et al., 2018). This suggests that low GEBV animals, i.e., those with a higher probability of having high SCC in milk, have increased *MYO1F* expression that facilitates neutrophil migration to the infected udder during mastitis, which, in turn, may explain high SCC in milk. *MYO1F* is a possible functional candidate gene for susceptibility to subclinical mastitis and high SCC in Norwegian Red; however, functional studies are required to verify this hypothesis.

Peptidyl arginine deiminase 4 (*PADI4*) was found to be significantly down-regulated in ST12-infected bMDMs from animals with low compared to high GEBV for SCC (**Supplementary Table 5**). It has been shown that *PADI4* mediates chromatin decondensation in neutrophils in response to inflammatory stimuli (Neeli et al., 2008). This suggests the possibility that *PADI4* might display a role in the susceptibility to subclinical mastitis and that low GEBV animals may have impaired neutrophil responses to infection. However, this hypothesis needs to be further tested.

A few genes were found to be differentially expressed between high and low SCC animals (**Supplementary Table 5**). *Defensin beta 5* (*DEFB5*) was down-regulated in both ST103- and ST12-infected bMDMs, while *C-X-C motif chemokine receptor 2* (*CXCR2*) was down-regulated only in ST103-infected bMDMs from animals with high compared to low SCC. Bovine beta-defensins are antimicrobial peptides providing first-line protection against pathogens, and their importance in host defense against mastitis-causing pathogens has been demonstrated (reviewed in Meade et al., 2014). The lower abundance of *DEFB5* mRNA observed in our study in bMDMs from high compared to low SCC animals may be associated with delayed clearance of bacteria from the udder during infection. *CXCR2* regulates neutrophil recruitment in a number of pathological conditions (Del Rio et al., 2001). It has been recently demonstrated that CXCR2-deficient mice display pro-inflammatory responses and increase in macrophage accumulation at inflamed sites during acute inflammation (Dyer et al., 2017). The observed lower abundance of *CXCR2* mRNA in bMDMs from high compared to low SCC animals can indicate a higher level of inflammatory response in these animals, resulting in a high number of SCC in milk. Additionally, *macrophage receptor with collagenous structure* (*MARCO*) was up-regulated in ST103-infected, ST12-infected, and uninfected control bMDMs from animals with high compared to low SCC. *MARCO* is a scavenger receptor that enables recruitment of mononuclear cells and pro-inflammatory cytokine production in response to

bacterial infection (Areschoug and Gordon, 2009; Dorrington et al., 2013). Interestingly, this gene was found 4.8 Mb from a single nucleotide polymorphism (SNP) significantly associated with high SCC in Norwegian Red (Kirsanova et al. in prep). Higher abundance of *MARCO* mRNA in bMDMs from high compared to low SCC animals may be associated with higher level of pro-inflammatory response and macrophage accumulation in high SCC animals. *MARCO* mRNA level was higher in unstimulated control bMDMs from high compared to low SCC animals. This indicates that transcription of this gene in monocytes is triggered by a pathogen infection and remains at the high level even long after the exposure. This may point toward *MARCO* as a potential biomarker for persistent chronic mastitis.

Whereas high milk SCC is a mastitis-related phenotype, the animals with low GEBV for SCC are likely to develop a high milk SCC phenotype, though it is not always the case. This indicates that the differentially expressed genes between low and high GEBV animals may be putative candidates for susceptibility to subclinical mastitis in Norwegian Red. However, comprehensive functional studies, both *in vitro* and *in vivo*, are necessary in order to verify all of these findings before including information on genetic variants of these genes in the genomic selection schemes. It is worth noting that even though the methods used for estimation of the differential expression (edgeR and DESeq2) perform well and are recommended for these type of analyses (Schurch et al., 2016), it may still be likely that some of these genes are false positives. Furthermore, it is important to mention that the results presented in our study are based on relative low sample size, which can, in turn, affect the outcomes of the analyses.

In conclusion, in our study, several genes and pathways were identified as differentially expressed during an early stage of bMDM infection with two *S. agalactiae* strains. Genes of granulocyte adhesion and phagosome formation pathways were significantly regulated in ST103-infected bMDMs, which may indicate the essential strategy of this strain to avoid host defense and promote its intracellular replication and persistence. Genes of Th1/Th2 activation and complement activation pathways were affected by ST12 infection, suggesting putative mechanisms developed by ST12 to escape the attack of the immune system. Furthermore, significant differences in gene responses between ST103 and ST12 have been identified, such as IL-12 signaling and granulocyte adhesion, which may partly be explained by differences in virulence factors between these strains. However, whether these pathways are evasion strategies developed by *S. agalactiae* to avoid host defense and promote its intracellular survival remains to be determined by functional studies. We have also identified *MARCO* as a potential biomarker for persistent chronic mastitis, and *MYO1F* and *PADI4* as putative candidates for susceptibility to subclinical mastitis in Norwegian Red. These results contribute to understanding of mechanisms involved in the pathogenesis of subclinical mastitis.

ETHICS STATEMENT

Blood sampling was performed by certified personnel and conducted in accordance with the laws and regulations controlling experiments using live animals in Norway. The study was approved by the Norwegian Animal Research Authority (Norwegian Food Safety Authority; FOTS id: 8661).

AUTHOR CONTRIBUTIONS

AL-S participated in the design of the study, carried out the infection experiments, analyzed the RNA sequencing data, performed the pathway analyses, and drafted the manuscript. EK analyzed the RNA sequencing data and participated in the part of the infection experiment and the manuscript drafting. CK provided software and participated in the RNA sequencing data analyses and manuscript drafting. PB, TS, BH, OØ, and IO participated in the study design, discussion, and interpretation of the results and manuscript drafting. All authors read and approved the final version of the manuscript.

FUNDING

This paper is a part of the "Multimast" project (no. 233778) funded by the Research Council of Norway, Tine and Geno.

ACKNOWLEDGMENTS

We thank Hege Lund and Cathinka Celine Jørgensen for help in collection of the blood samples and flow cytometry analysis, and Tormod Mørk and Bjørg Kvitle from the Norwegian Veterinary Institute for providing the *S. agalactiae* isolates. We also thank the farmers for their collaboration during blood sampling of the animals. We are grateful to Silje Furre Hansen for culturing the bacteria. We thank Anne Storset for help in the planning of the project, design of the experiments, and discussion of the results.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00689/full#supplementary-material.

REFERENCES

Agerer, F., Lux, S., Michel, A., Rohde, M., Ohlsen, K., and Hauck, C. R. (2005). Cellular invasion by Staphylococcus aureus reveals a functional link between focal adhesion kinase and cortactin in integrin-mediated internalisation. *J. Cell. Sci.* 118 (Pt 10), 2189–2200. doi: 10.1242/jcs.02328

Alluwaimi, A. M., Leutenegger, C. M., Farver, T. B., Rossitto, P. V., Smith, W. L., and Cullor, J. S. (2003). The cytokine markers in Staphylococcus aureus mastitis of bovine mammary gland. *J. Vet. Med. B Infect. Dis. Vet. Public Health* 50 (3), 105–111.

Almeida, P. E., Roque, N. R., Magalhaes, K. G., Mattos, K. A., Teixeira, L., Maya-Monteiro, C., et al. (2014). Differential TLR2 downstream signaling regulates lipid metabolism and cytokine production triggered by Mycobacterium bovis

BCG infection. *Biochim. Biophys. Acta.* 1841 (1), 97–107. doi: 10.1016/j.bbalip.2013.10.008

Arendse, B., Van Snick, J., and Brombacher, F. (2005). IL-9 is a susceptibility factor in Leishmania major infection by promoting detrimental Th2/type 2 responses. *J. Immunol.* 174 (4), 2205–2211.

Areschoug, T., and Gordon, S. (2009). Scavenger receptors: role in innate immunity and microbial pathogenesis. *Cell Microbiol.* 11 (8), 1160–1169. doi: 10.1111/j.1462-5822.2009.01326.x

Bahrami, S., and Drablos, F. (2016). Gene regulation in the immediate-early response process. *Adv. Biol. Regul.* 62, 37–49. doi: 10.1016/j.jbior.2016.05.001

Barber, M. R., and Yang, T. J. (1998). Chemotactic activities in nonmastitic and mastitic mammary secretions: presence of interleukin-8 in mastitic but not nonmastitic secretions. *Clin. Diagn. Lab. Immunol.* 5 (1), 82–86.

Budd, K. E., Mitchell, J., and Keane, O. M. (2016). Lineage associated expression of virulence traits in bovine-adapted Staphylococcus aureus. *Vet. Microbiol.* 189, 24–31. doi: 10.1016/j.vetmic.2016.04.013

Casey, M. E., Meade, K. G., Nalpas, N. C., Taraktsoglou, M., Browne, J. A., Killick, K. E., et al. (2015). Analysis of the bovine monocyte-derived macrophage response to Mycobacterium avium subspecies paratuberculosis infection using RNA-seq. *Front. Immunol.* 6, 23. doi: 10.3389/fimmu.2015.00023

Copin, R., De Baetselier, P., Carlier, Y., Letesson, J. J., and Muraille, E. (2007). MyD88-dependent activation of B220-CD11b+LY-6C+ dendritic cells during Brucella melitensis infection. *J. Immunol.* 178 (8), 5182–5191.

de Klerk, N., Saroj, S. D., Wassing, G. M., Maudsdotter, L., and Jonsson, A. B. (2017). The host cell transcription factor EGR1 is induced by bacteria through the EGFR-ERK1/2 pathway. *Front. Cell Infect. Microbiol.* 7, 16. doi: 10.3389/fcimb.2017.00016

Del Rio, L., Bennouna, S., Salinas, J., and Denkers, E. Y. (2001). CXCR2 deficiency confers impaired neutrophil recruitment and increased susceptibility during Toxoplasma gondii infection. *J. Immunol.* 167 (11), 6503–6509.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29 (1), 15–21. doi: 10.1093/bioinformatics/bts635

Dorrington, M. G., Roche, A. M., Chauvin, S. E., Tu, Z., Mossman, K. L., Weiser, J. N., et al. (2013). MARCO is required for TLR2- and Nod2-mediated responses to Streptococcus pneumoniae and clearance of pneumococcal colonization in the murine nasopharynx. *J. Immunol.* 190 (1), 250–258. doi: 10.4049/jimmunol.1202113

Drag, M., Skinkyté-Juskiené, R., Do, D. N., Kogelman, L., and Kadarmideen, H. N. (2017). Differential expression and co-expression gene networks reveal candidate biomarkers of boar taint in non-castrated pigs. *Sci. Rep.* 7 (1), 12205. doi: 10.1038/s41598-017-11928-0

Dyer, D. P., Pallas, K., Ruiz, L. M., Schuette, F., Wilson, G. J., and Graham, G. J. (2017). CXCR2 deficient mice display macrophage-dependent exaggerated acute inflammatory responses. *Sci. Rep.* 7, 42681. doi: 10.1038/srep42681

Erb, K. J., Holloway, J. W., Sobeck, A., Moll, H., and Le Gros, G. (1998). Infection of mice with Mycobacterium bovis-Bacillus Calmette-Guerin (BCG) suppresses allergen-induced airway eosinophilia. *J. Exp. Med.* 187 (4), 561–569.

Gunther, J., Koy, M., Berthold, A., Schuberth, H. J., and Seyfert, H. M. (2016). Comparison of the pathogen species-specific immune response in udder derived cell types and their models. *Vet. Res.* 47, 22. doi: 10.1186/s13567-016-0307-3

Guo, C. M., Chen, R. R., Kalhoro, D. H., Wang, Z. F., Liu, G. J., Lu, C. P., et al. (2014). Identification of genes preferentially expressed by highly virulent piscine Streptococcus agalactiae upon interaction with macrophages. *PLoS One* 9 (2), e87980. doi: 10.1371/journal.pone.0087980

Harris, J., De Haro, S. A., Master, S. S., Keane, J., Roberts, E. A., Delgado, M., et al. (2007). T helper 2 cytokines inhibit autophagic control of intracellular Mycobacterium tuberculosis. *Immunity* 27 (3), 505–517. doi: 10.1016/j.immuni.2007.07.022

Henneke, P., and Berner, R. (2006). Interaction of neonatal phagocytes with group B streptococcus: recognition and response. *Infect. Immun.* 74 (6), 3085–3095. doi: 10.1128/iai.01551-05

Heringstad, B., Klemetsdal, G., and Steine, T. (2007). Selection responses for disease resistance in two selection experiments with Norwegian red cows. *J. Dairy Sci.* 90 (5), 2419–2426. doi: 10.3168/jds.2006-805

Hop, H. T., Arayan, L. T., Reyes, A. W. B., Huy, T. X. N., Min, W., Lee, H. J., et al. (2017). Simultaneous RNA-seq based transcriptional profiling of intracellular Brucella abortus and B. abortus-infected murine macrophages. *Microb. Pathog.* 113, 57–67. doi: 10.1016/j.micpath.2017.10.029

Jensen, K. D., Wang, Y., Wojno, E. D., Shastri, A. J., Hu, K., Cornel, L., et al. (2011). Toxoplasma polymorphic effectors determine macrophage polarization and intestinal inflammation. *Cell Host Microbe* 9 (6), 472–483. doi: 10.1016/j.chom.2011.04.015

Jorgensen, H. J., Nordstoga, A. B., Sviland, S., Zadoks, R. N., Solverod, L., Kvitle, B., et al. (2016). Streptococcus agalactiae in the environment of bovine dairy herds–rewriting the textbooks? *Vet. Microbiol.* 184, 64–72. doi: 10.1016/j.vetmic.2015.12.014

Kogelman, L. J., Cirera, S., Zhernakova, D. V., Fredholm, M., Franke, L., and Kadarmideen, H. N. (2014). Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA Sequencing in a porcine model. *BMC Med. Genomics* 7, 57. doi: 10.1186/1755-8794-7-57

Korir, M. L., Laut, C., Rogers, L. M., Plemmons, J. A., Aronoff, D. M., and Manning, S. D. (2017). Differing mechanisms of surviving phagosomal stress among group B Streptococcus strains of varying genotypes. *Virulence* 8 (6), 924–937. doi: 10.1080/21505594.2016.1252016

Koziel, J., Maciag-Gudowska, A., Mikolajczyk, T., Bzowska, M., Sturdevant, D. E., Whitney, A. R., et al. (2009). Phagocytosis of Staphylococcus aureus by macrophages exerts cytoprotective effects manifested by the upregulation of antiapoptotic factors. *PLoS One* 4 (4), e5210. doi: 10.1371/journal.pone.0005210

Lambris, J. D., Ricklin, D., and Geisbrecht, B. V. (2008). Complement evasion by human pathogens. *Nat. Rev. Microbiol.* 6 (2), 132–142. doi: 10.1038/nrmicro1824

Lewandowska-Sabat, A. M., Boman, G. M., Downing, A., Talbot, R., Storset, A. K., and Olsaker, I. (2013). The early phase transcriptome of bovine monocyte-derived macrophages infected with Staphylococcus aureus *in vitro*. *BMC Genomics* 14, 891. doi: 10.1186/1471-2164-14-891

Lewandowska-Sabat, A. M., Hansen, S. F., Solberg, T. R., Osteras, O., Heringstad, B., Boysen, P., et al. (2018a). MicroRNA expression profiles of bovine monocyte-derived macrophages infected *in vitro* with two strains of Streptococcus agalactiae. *BMC Genomics* 19 (1), 241. doi: 10.1186/s12864-018-4591-3

Lewandowska-Sabat, A. M., Kirsanova, E., Klopp, C., Solberg, T. R., Heringstad, B., Osteras, O., et al. (2018b). Transcriptional profiling of monocyte-derived macrophages infected *in vitro* with two strains of Streptococcus Agalactiae reveals candidate pathways affecting subclinical mastitis in cattle. *EMBL-EBI*. https://www.ebi.ac.uk/ena/data/view/PRJEB24166.

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323. doi: 10.1186/1471-2105-12-323

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi: 10.1186/s13059-014-0550-8

Lubbers, R., van Essen, M. F., van Kooten, C., and Trouw, L. A. (2017). Production of complement components by cells of the immune system. *Clin. Exp. Immunol.* 188 (2), 183–194. doi: 10.1111/cei.12952

Miglior, F., Fleming, A., Malchiodi, F., Brito, L. F., Martin, P., and Baes, C. F. (2017). A 100-Year Review: identification and genetic selection of economically important traits in dairy cattle. *J. Dairy Sci.* 100 (12), 10251–10271. doi: 10.3168/jds.2017-12968

Manning, S. D., Springman, A. C., Lehotzky, E., Lewis, M. A., Whittam, T. S., and Davies, H. D. (2009). Multilocus sequence types associated with neonatal group B streptococcal sepsis and meningitis in Canada. *J. Clin. Microbiol.* 47 (4), 1143–1148. doi: 10.1128/jcm.01424-08

Mavromatis, C. H., Bokil, N. J., Totsika, M., Kakkanat, A., Schaale, K., Cannistraci, C. V., et al. (2015). The co-transcriptome of uropathogenic Escherichia coli-infected mouse macrophages reveals new insights into host-pathogen interactions. *Cell Microbiol.* 17 (5), 730–746. doi: 10.1111/cmi.12397

Meade, K. G., Cormican, P., Narciandi, F., Lloyd, A., and O'Farrelly, C. (2014). Bovine beta-defensin gene family: opportunities to improve animal health? *Physiol. Genomics* 46 (1), 17–28. doi: 10.1152/physiolgenomics.00085.2013

Muraille, E., Leo, O., and Moser, M. (2014). TH1/TH2 paradigm extended: macrophage polarization as an unappreciated pathogen-driven escape mechanism? *Front. Immunol.* 5, 603. doi: 10.3389/fimmu.2014.00603

Nalpas, N. C., Park, S. D., Magee, D. A., Taraktsoglou, M., Browne, J. A., Conlon, K. M., et al. (2013). Whole-transcriptome, high-throughput RNA sequence analysis of the bovine macrophage response to Mycobacterium bovis infection *in vitro*. *BMC Genomics* 14, 230. doi: 10.1186/1471-2164-14-230

Neeli, I., Khan, S. N., and Radic, M. (2008). Histone deimination as a response to inflammatory stimuli in neutrophils. *J. Immunol.* 180 (3), 1895–1902.

O'Donnell, A., Odrowaz, Z., and Sharrocks, A. D. (2012). Immediate-early gene activation by the MAPK pathways: what do and don't we know? *Biochem. Soc. Trans.* 40 (1), 58–66. doi: 10.1042/bst20110636

Ozeri, V., Rosenshine, I., Ben-Ze'Ev, A., Bokoch, G. M., Jou, T. S., and Hanski, E. (2001). De novo formation of focal complex-like structures in host cells by invading Streptococci. *Mol. Microbiol* 41 (3), 561–573.

Pamer, E. G. (2004). Immune responses to Listeria monocytogenes. *Nat. Rev. Immunol.* 4 (10), 812–823. doi: 10.1038/nri1461

Pang, M., Sun, L., He, T., Bao, H., Zhang, L., Zhou, Y., et al. (2017). Molecular and virulence characterization of highly prevalent Streptococcus agalactiae circulated in bovine dairy herds. *Vet. Res.* 48 (1), 65. doi: 10.1186/s13567-017-0461-2

Peters, W., Scott, H. M., Chambers, H. F., Flynn, J. L., Charo, I. F., and Ernst, J. D. (2001). Chemokine receptor 2 serves an early and essential role in resistance to Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. U.S.A.* 98 (14), 7958–7963. doi: 10.1073/pnas.131207398

Plancon, L., Du Merle, L., Le Friec, S., Gounon, P., Jouve, M., Guignot, J., et al. (2003). Recognition of the cellular beta1-chain integrin by the bacterial AfaD invasin is implicated in the internalization of afa-expressing pathogenic Escherichia coli strains. *Cell Microbiol.* 5 (10), 681–693.

Politis, I., McBride, B. W., Burton, J. H., Zhao, X., and Turner, J. D. (1991). Secretion of interleukin-1 by bovine milk macrophages. *Am. J. Vet. Res.* 52 (6), 858–862.

Portnoy, D. A., Auerbuch, V., and Glomski, I. J. (2002). The cell biology of Listeria monocytogenes infection: the intersection of bacterial pathogenesis and cell-mediated immunity. *J. Cell. Biol.* 158 (3), 409–414. doi: 10.1083/jcb.200205009

Ramirez-Carrozzi, V. R., Braas, D., Bhatt, D. M., Cheng, C. S., Hong, C., Doty, K. R., et al. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* 138 (1), 114–128. doi: 10.1016/j.cell.2009.04.020

Riollet, C., Rainard, P., and Poutrel, B. (2001). Cell subpopulations and cytokine expression in cow milk in response to chronic Staphylococcus aureus infection. *J. Dairy Sci.* 84 (5), 1077–1084. doi: 10.3168/jds.S0022-0302(01)74568-7

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11 (3), R25. doi: 10.1186/gb-2010-11-3-r25

Roy, S., Schmeier, S., Kaczkowski, B., Arner, E., Alam, T., Ozturk, M., et al. (2018). Transcriptional landscape of Mycobacterium tuberculosis infection in macrophages. *Sci. Rep.* 8 (1), 6758. doi: 10.1038/s41598-018-24509-6

Salvermoser, M., Pick, R., Weckbach, L. T., Zehrer, A., Lohr, P., Drechsler, M., et al. (2018). Myosin 1f is specifically required for neutrophil migration in 3D environments during acute inflammation. *Blood* 131 (17), 1887–1898. doi: 10.1182/blood-2017-10-811851

Schurch, N. J., Schofield, P., Gierlinski, M., Cole, C., Sherstnev, A., Singh, V., et al. (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *Rna* 22 (6), 839–851. doi: 10.1261/rna.053959.115

Tassi, R., McNeilly, T. N., Sipka, A., and Zadoks, R. N. (2015). Correlation of hypothetical virulence traits of two Streptococcus uberis strains with the clinical manifestation of bovine mastitis. *Vet. Res.* 46, 123. doi: 10.1186/s13567-015-0268-y

Toca, F. N., Dunaway, K., Smaltz, F., Szulc-Dąbrowska, L., Drnevich, J., Mielcarska, M. B., et al. (2019). Bacterial and viral pathogen-associated molecular patterns induce divergent early transcriptomic landscapes in a bovine macrophage cell line. *BMC Genomics* 20 (1), 15. doi: 10.1186/s12864-018-5411-5

Wong, D., Chao, J. D., and Av-Gay, Y. (2013). Mycobacterium tuberculosis-secreted phosphatases: from pathogenesis to targets for TB drug development. *Trends Microbiol.* 21 (2), 100–109. doi: 10.1016/j.tim.2012.09.002

Wynn, T. A. (2015). Type 2 cytokines: mechanisms and therapeutic strategies. *Nat. Rev. Immunol.* 15 (5), 271–282. doi: 10.1038/nri3831

Wynn, T. A., and Ramalingam, T. R. (2012). Mechanisms of fibrosis: therapeutic translation for fibrotic disease. *Nat. Med.* 18 (7), 1028–1040. doi: 10.1038/nm.2807

Xavier, M. N., Winter, M. G., Spees, A. M., den Hartigh, A. B., Nguyen, K., Roux, C. M., et al. (2013). PPARgamma-mediated increase in glucose availability sustains chronic Brucella abortus infection in alternatively activated macrophages. *Cell Host Microbe* 14 (2), 159–170. doi: 10.1016/j.chom.2013.07.009

Conflict of Interest Statement: Authors TS and BH were employed by company Geno SA. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

Supplementary Table 1.

The animals used in the study.

Animal ID	Farm ID	GEBV on SCC	SCC (cells/ml)	Grouping GEBV	Grouping SCC	Farm type
1	1	93	754,000	Low	High	AMS/F
2	1	114	15,000	High	Low	AMS/F
3	2	95	15,000	Low	Low	AMS/F
4	2	97	18,000	Low	Low	AMS/F
5	2	105	18,000	High	Low	AMS/F
6	2	109	18,000	High	Low	AMS/F
7	1	90	779,000	Low	High	AMS/F
8	1	93	667,000	Low	High	AMS/F
9	1	107	18,000	High	Low	AMS/F
10	3	97	13,000	Low	Low	CMS/F
11	3	120	16,000	High	Low	CMS/F
12	3	122	16,000	High	Low	CMS/F

GEBV – genomic estimated breeding value, SCC – somatic cell count, SCC (cells/ml) – geometric mean of three measurements of somatic cell count in milk during the last lactation, AMS – automatic milking system, F – freestall, CMS – conventional milking system.

Supplementary Table 2.

Summary of the results from Illumina sequencing of bovine monocyte-derived macrophages (bMDMs) transcriptome and alignment to the bovine reference genome Bos taurus v3.1. Control – negative control: uninfected bMDMs; ST103 and ST12 – bMDMs infected with the ST103 or the ST12 strain of *Streptococcus agalactiae*, respectively.

Sample name	Total number of reads (paired end)	Total number of reads passing filters (percentage)	Percentage of uniquely mapped reads
1-RNA1-control	46,716,220	46,522,156 (99.6%)	85.73%
2-RNA1-ST12	55,211,689	54,841,861 (99.3%)	86.77%
3-RNA1-ST103	43,583,615	43,416,422 (99.6%)	86.79%
4-RNA2-control	46,453,702	46,130,305 (99.3%)	79.27%
5-RNA2-ST12	57,132,859	56,924,239 (99.6%)	86.89%
6-RNA2-ST103	44,305,652	44,124,250 (99.6%)	86.66%
7-RNA4-control	43,210,326	42,942,817 (99.4%)	87.42%
8-RNA4-ST12	49,028,231	48,819,816 (99.6%)	86.67%
9-RNA4-ST103	54,733,215	54,497,956 (99.6%)	78.89%
10-RNA5-control	59,267,812	58,930,499 (99.4%)	87.53%
11-RNA5-ST12	41,598,991	41,246,491 (99.2%)	89.13%
12-RNA5-ST103	40,005,853	39,512,928 (98.8%)	82.61%
13-RNA6-control	56,282,438	55,966,732 (99.4%)	88.02%
14-RNA6-ST12	31,101,943	30,768,233 (98.9%)	85.72%
15-RNA6-ST103	35,049,353	34,698,255 (99.0%)	87.96%
16-RNA7-control	43,382,792	43,144,051 (99.4%)	86.73%
17-RNA7-ST12	69,204,326	68,743,984 (99.3%)	87.70%
18-RNA7-ST103	67,310,229	66,574,650 (98.9%)	88.14%
19-RNA8-control	41,369,950	41,173,893 (99.5%)	87.98%
20-RNA8-ST12	37,210,508	37,090,345 (99.7%)	84.42%
21-RNA8-ST103	34,940,864	34,839,101 (99.7%)	87.77%
22-RNA9-control	36,572,962	36,439,529 (99.6%)	85.89%
23-RNA9-ST12	44,110,490	43,971,700 (99.7%)	89.15%
24-RNA9-ST103	39,831,179	39,744,268 (99.8%)	89.11%
25-RNA10-control	52,317,464	52,151,716 (99.7%)	88.14%
26-RNA10-ST12	66,448,527	66,008,925 (99.3%)	86.96%
27-RNA10-ST103	49,827,661	49,656,291 (99.7%)	87.93%
28-RNA11-control	39,688,837	39,511,757 (99.6%)	80.72%
29-RNA11-ST12	46,690,018	46,539,724 (99.7%)	83.38%
30-RNA11-ST103	36,320,764	36,104,819 (99.4%)	82.45%

31-RNA12-control	37,743,793	37,634,899 (99.7%)	87.31%
32-RNA12-ST12	44,230,782	44,119,535 (99.7%)	87.09%
33-RNA12-ST103	44,485,208	44,406,157 (99.8%)	88.21%
34-RNA13-control	33,538,407	33,372,974 (99.5%)	86.31%
35-RNA13-ST12	52,134,761	51,948,766 (99.6%)	88.43%
36-RNA13-ST103	35,726,168	35,571,863 (99.6%)	86.91%

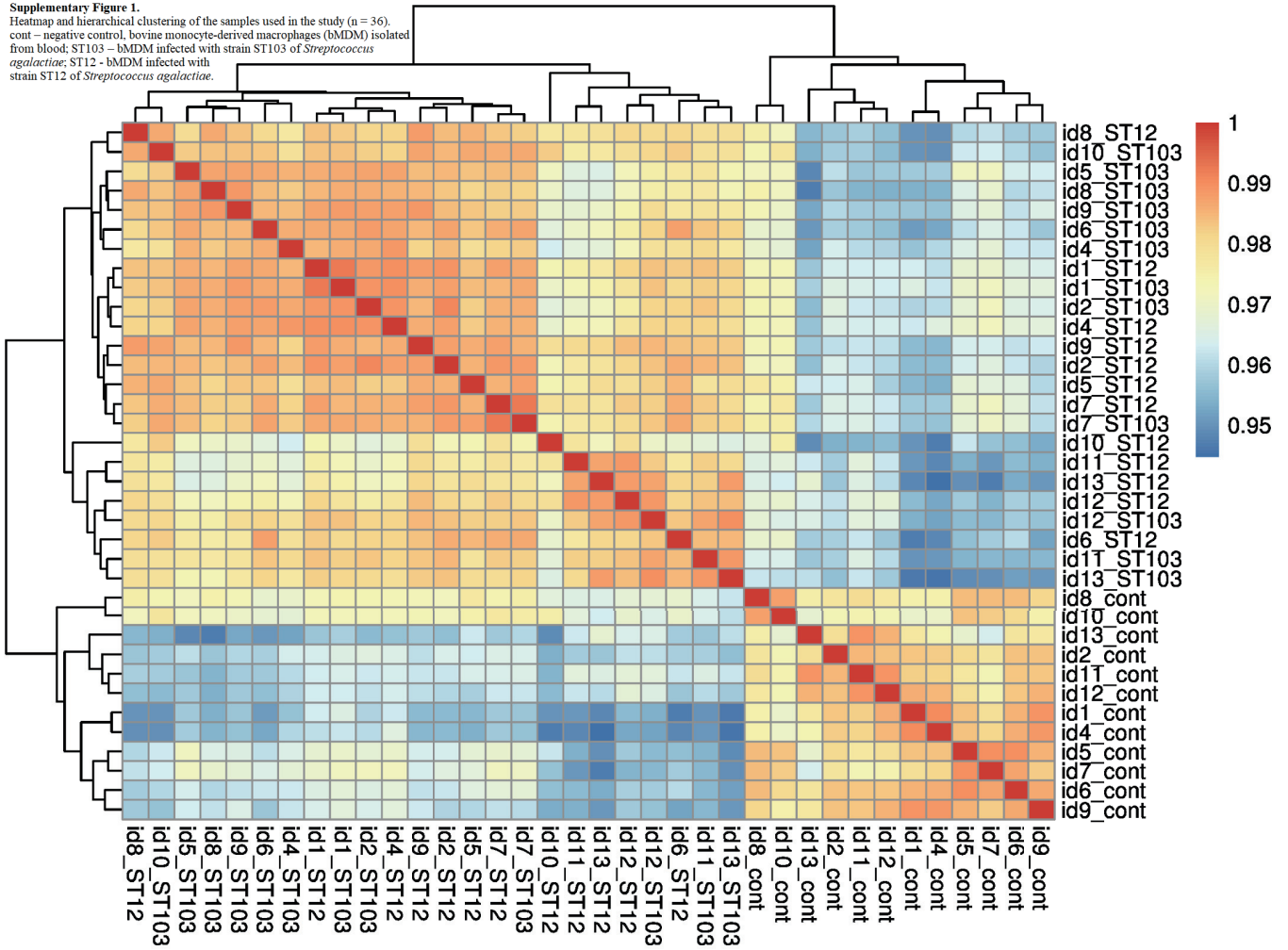
Supplementary Table 3.

The RNA samples used in the study.

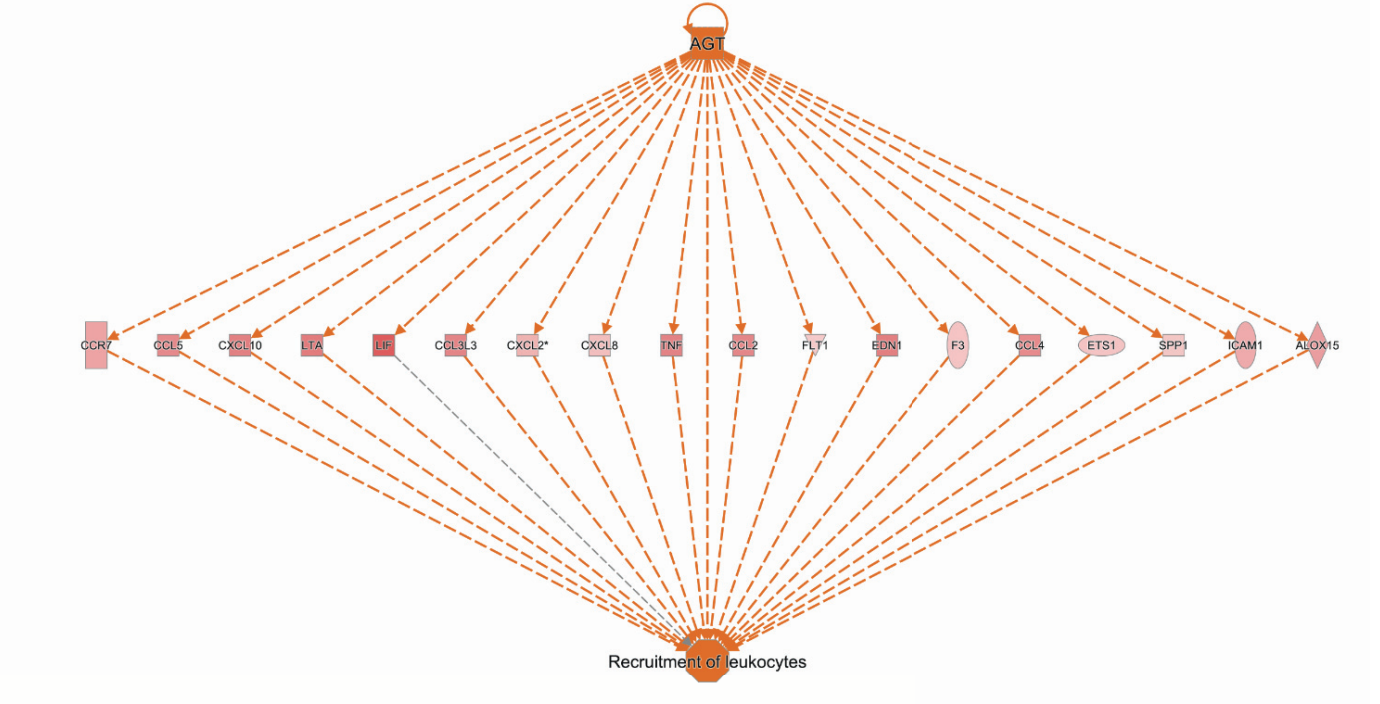
Sample ID	Animal ID	Conc. (ng/µl)	A260/280	RIN
RNA1_control	1	75,8	2,1	8,8
RNA1_ST12	1	74,1	2,1	9,2
RNA1_ST103	1	45,5	2,2	8,7
RNA2_control	2	85,5	2,1	9,0
RNA2_ST12	2	60,4	1,9	N/A
RNA2_ST103	2	32,8	2,0	9,0
RNA4_control	3	111,8	2,0	8,5
RNA4_ST12	3	62,9	2,2	8,4
RNA4_ST103	3	49,2	2,0	9,0
RNA5_control	4	35,5	2,2	9,2
RNA5_ST12	4	38	2,0	8,3
RNA5_ST103	4	25,5	2,8	8,7
RNA6_control	5	61,6	2,1	8,8
RNA6_ST12	5	24,7	2,0	8,6
RNA6_ST103	5	22,6	2,1	8,8
RNA7_control	6	28	2,1	9,0
RNA7_ST12	6	35,5	1,8	8,2
RNA7_ST103	6	23	1,7	9,1
RNA8_control	7	18,9	2,0	8,9
RNA8_ST12	7	18,9	2,0	7,7
RNA8_ST103	7	10,1	2,3	N/A
RNA9_control	8	63,5	2,2	8,6
RNA9_ST12	8	40,5	2,2	N/A
RNA9_ST103	8	57,7	2,2	8,7
RNA10_control	9	16,2	2,0	N/A
RNA10_ST12	9	24,3	2,0	N/A
RNA10_ST103	9	13,4	2,7	N/A
RNA11_control	10	80,4	2,0	9,4
RNA11_ST12	10	35,6	2,2	4,0
RNA11_ST103	10	104,2	2,1	7,2
RNA12_control	11	110,7	2,0	9,9
RNA12_ST12	11	38,9	2,0	N/A
RNA12_ST103	11	72,9	2,0	N/A
RNA13_control	12	79,1	2,1	9,5
RNA13_ST12	12	34,8	2,3	N/A
RNA13_ST103	12	66,7	2,1	N/A

RIN - RNA integrity number, N/A – not available.

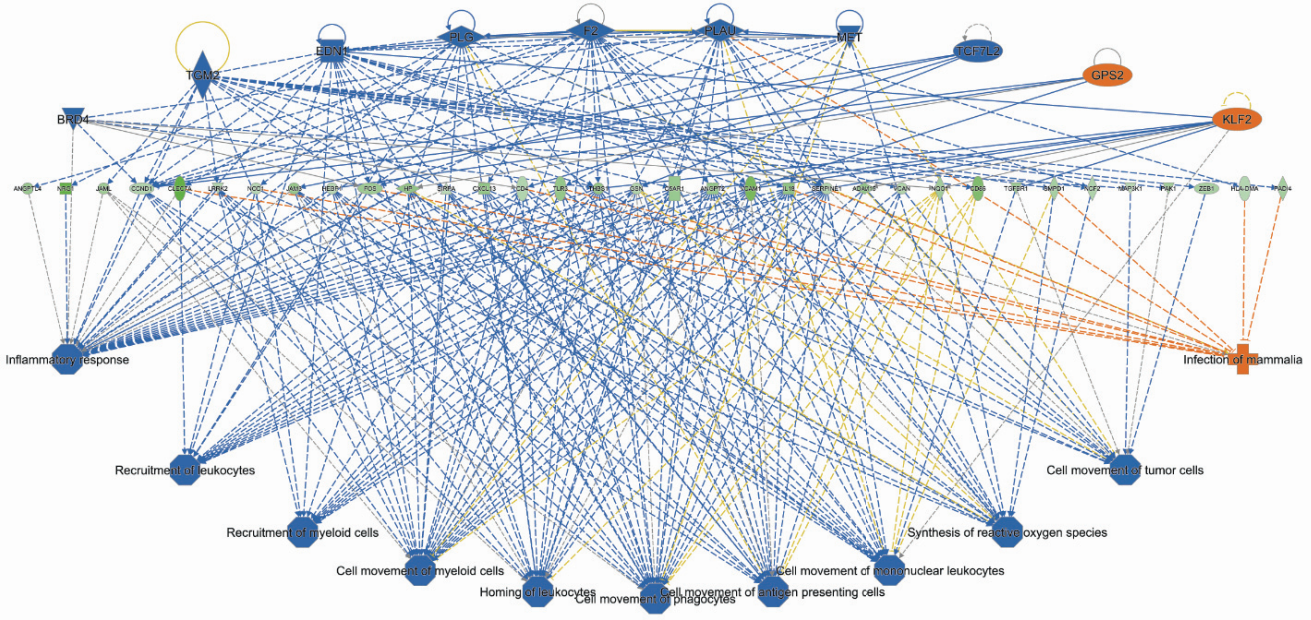
Supplementary Figure 1.
Heatmap and hierarchical clustering of the samples used in the study (n = 36). cont – negative control, bovine monocyte-derived macrophages (bMDM) isolated from blood; ST103 – bMDM infected with strain ST103 of *Streptococcus agalactiae*; ST12 - bMDM infected with strain ST12 of *Streptococcus agalactiae*.

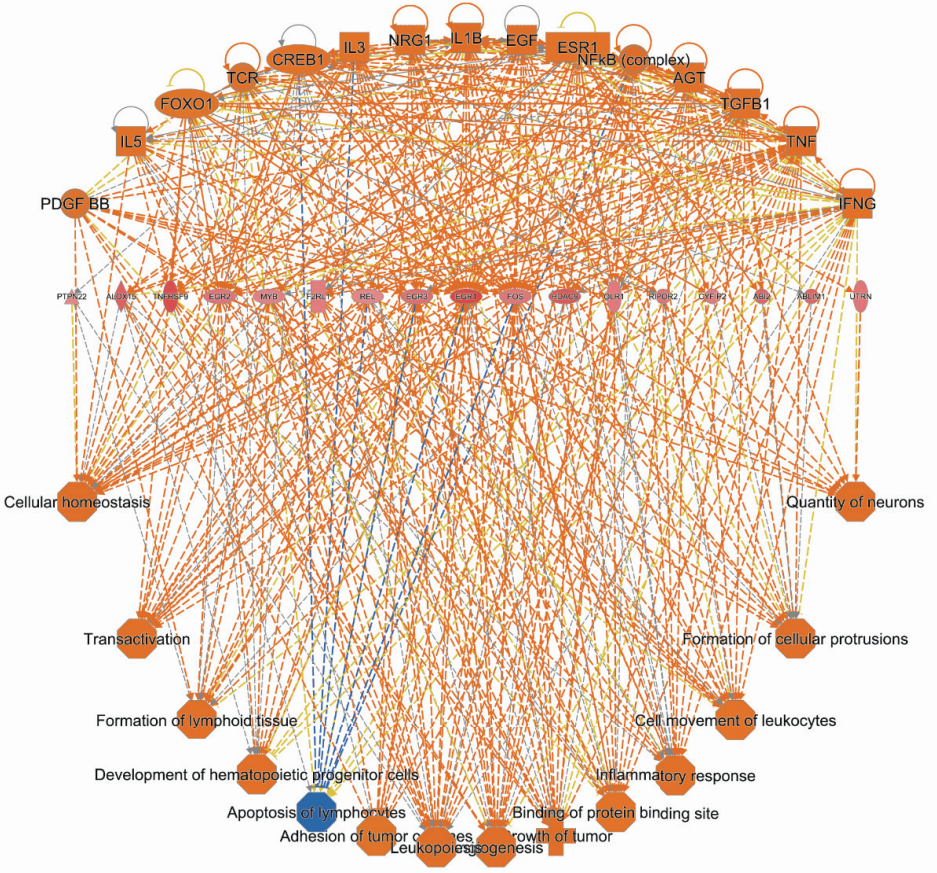


Supplementary Figure 2.
The top regulator identified by Ingenuity Pathway Analysis (IPA) for genes up-regulated in response to infection of bovine monocyte-derived macrophages (bMDM) with *Streptococcus agalactiae* strain ST12.



Supplementary Figure 3.
The top regulators identified by Ingenuity Pathway Analysis (IPA) for genes down-regulated in response to infection of bovine monocyte-derived macrophages (bMDM) with *Streptococcus agalactiae* strain ST12.





Supplementary Figure 4.
The top regulators identified by Ingenuity Pathway Analysis (IPA) for genes up-regulated in response to infection of bovine monocyte-derived macrophages (bMDM) with *Streptococcus agalactiae* strain ST12 compared to ST103.

Additional supplementary material for this article can be found online at:

https://www.frontiersin.org/articles/10.3389/fgene.2019.00689/full#supplementary-material

Supplementary Table 4.

Supplementary Table 5.

Supplementary Table 6.

Supplementary Table 7.

ISBN: 978-82-575-1682-6
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

106801 / ANDVORDGRAFISK.NO