ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi:10.1111/j.1365-2052.2010.02165.x

Quantitative trait loci for clinical mastitis on chromosomes 2, 6, 14 and 20 in Norwegian Red cattle

M. Sodeland*, M. P. Kent*, H. G. Olsen*, M. A. Opsal*, M. Svendsen†, E. Sehested†, B. J. Hayes*,‡ and S. Lien*

*Department of Animal and Aquacultural Sciences, Centre for Integrative Genetics, Norwegian University of Life Sciences, N-1432 Aas, Norway. †Geno Breeding and AI organization, Norwegian University of Life Sciences, Box 5003, N-1432 Aas, Norway. ‡Biosciences Research Division, Department of Primary Industries Victoria, Melbourne 3083, Vic., Australia.

Summary

Mastitis is the most frequent and costly disease in dairy production and solutions leading to a reduction in the incidence of mastitis are highly demanded. Here a genome-wide association study was performed to identify polymorphisms affecting susceptibility to mastitis. Genotypes for 17 349 SNPs distributed across the 29 bovine autosomal chromosomes from a total of 2589 sires with 1 389 776 daughters with records on clinical mastitis were included in the analysis. Records of occurrence of clinical mastitis were divided into seven time periods in the first three lactations in order to identify quantitative trait loci affecting mastitis susceptibility in particular phases of lactation. The most convincing results from the association mapping were followed up and validated by a combined linkage disequilibrium and linkage analysis. The study revealed quantitative trait loci affecting occurrence of clinical mastitis in the periparturient period on chromosomes 2, 6 and 20 and a quantitative trait locus affecting occurrence of clinical mastitis in late lactation on chromosome 14. None of the quantitative trait loci for clinical mastitis detected in the study seemed to affect lactation average of somatic cell score. The SNPs highly associated with clinical mastitis lie near both the gene encoding interleukin 8 on chromosome 6 and the genes encoding the two interleukin 8 receptors on chromosome 2.

Keywords clinical mastitis, genome-wide association studies, Linkage disequilibrium linkage analysis, Norwegian Red cattle, Somatic cell score.

Background

Mastitis has a substantial impact on the dairy industry. The disease affects approximately 25% of Norwegian dairy cows each year (Østeras *et al.* 2006), and estimated annual losses of dairy farmers caused by mastitis are $2 billion in the US and £300 million in the UK (Viguier *et al.* 2009). Infections imply great costs to farmers, animal suffering, and use of antibiotics. In Norwegian Red cattle (NRF), *Staphylococcus aureus* is identified in the inflamed udder quarter of 55% of cows with clinical mastitis (CM) (Østeras *et al.* 2006). Bacterial infections are usually cleared by host defence mechanisms or antibiotic treatment within a few days. If host defence and antibiotic treatment are not successful, the outcome may be chronic infection, mammary gland tissue damage or death (Waller 2000; Sordillo 2005; Strandberg *et al.* 2005; Lahouassa *et al.* 2007).

Immunological defences of the mammary gland include anatomical features, cells, soluble molecules, and receptors. The efficiency of this defence changes through the stages of lactation. Susceptibility to mastitis increases during the periparturient period (−15 to 30 days postpartum) and two-thirds of mastitis incidents occur in the first two months of lactation (Syvajarvi *et al.* 1986; Waller 2000; Sordillo 2005). During the periparturient period the mammary gland goes through a transition to initiate milk production. The transition requires both hormonal changes and higher energy demand. Anatomical, cellular and soluble defences against infection are all altered or impaired during this

Address for correspondence

M. Sodeland, Department of Animal and Aquacultural Sciences, Centre for Integrative Genetics, Norwegian University of Life Sciences, N-1432 Aas, Norway
E-mail: marte.sodeland@umb.no

Accepted for publication 27 September 2010

period (Waller 2000; Sordillo 2005; Østeras *et al.* 2006). In NRF patterns of strong genetic correlation for CM in the periparturient period, as well for CM in late lactation, have been found between the first three lactations per cow (Svendsen & Heringstad 2006a). As defence mechanisms change through stages of lactation, it is reasonable to treat occurrence of CM in the different stages of lactation as different traits when attempting to map and characterize quantitative trait loci (QTL) affecting susceptibility to CM. Previous linkage mapping studies have reported QTL affecting CM on a number of *Bos taurus* autosomes (BTAs). Klungland *et al.* (2001) reported a QTL for CM in NRF on BTAs 3, 4, 6, 14 and 27. Moreover, Sahana *et al.* (2008) reported a QTL for CM on BTA9 at 73.9 Mb in three Nordic cattle breeds and Schulman *et al.* (2004) reported QTL for mastitis on BTAs 14 and 18 in Finnish Ayrshire.

Several authors have used high somatic cell score (SCS) in milk as an indication of clinical mastitis in QTL mapping (Bennewitz *et al.* 2003; Kuhn *et al.* 2003; Leyva-Baca *et al.* 2007). A wide range of values for genetic correlation between CM and SCS in cattle populations has been reported, with an average of about 0.7 (Mrode & Swanson 1996). In a NRF study of genetic correlations between SCS and occurrence of CM in different stages of lactation, higher correlations were found between SCS and occurrence of CM in late lactation than between SCS and occurrence of CM in the periparturient period (Svendsen & Heringstad 2006b). While a high value of SCS is an indicator of disease, a low value might not be such a good indicator of udder status (Heringstad *et al.* 2000).

In NRF occurrence of CM is also correlated with protein yield in milk (Heringstad *et al.* 2005). This correlation could be due to linkage between QTLs influencing milk production traits and QTLs influencing CM or due to the presence of pleiotropic effects. Pleiotropic effects of increased milk production on occurrence of CM might be due to increased energy demand or increased strain on anatomical features of the mammary gland.

Detection of QTLs in NRF and other cattle populations is facilitated by the bovine genome sequence (Elsik *et al.* 2009; Liu *et al.* 2009) and the availability of large scale bovine SNP-arrays (Gibbs *et al.* 2009; Matukumalli *et al.* 2009). NRF is an ideal model breed for genome-wide association studies (GWAS) for CM, as very large numbers of veterinary reported clinical mastitis (VRCM) records are available for this breed (Østeras *et al.* 2007).

Materials and methods

Animals and experimental design

NRF is a mixed breed formed not only from local Norwegian breeds, Ayrshire and Swedish Red and White (SRB), but also with some influence of Holstein. Norway has a nationwide recording system for health data from dairy cattle. The national Norwegian Dairy Herd Recording System has included VRCM since 1975 and SCS since 1978 (Østeras *et al.* 2007). Here a design with genotypes from NRF sires and phenotypic records from their daughters for VRCM or SCS was used. Such a design benefits from the large number of records per sire, which gives a marked decrease in variance due to environmental effects compared with other designs (Weller *et al.* 1990). Records of VRCM were retrieved as a binary trait for a total of 1 389 776 daughters of 2086 paternal half-sib sires from 109 families. Number of daughters per sire ranged from 45 to 5793, with a median value of 285. VRCM records were divided into occurrence of CM in each of seven categorical time periods, treating occurrences of CM in the different time periods as different traits. First lactation was divided into three time periods, whereas second and third lactation were divided into two time periods each. Time periods are described in Table 1. Daughter-yield-deviations (DYD) were calculated for each sire for CM and SCS based on daughter records for each sire. DYD for CM were calculated for each of the seven time periods, as CM in the different time periods were treated as different traits in our analyses. Records on SCS were retrieved as lactation averages for 2 791 524 daughter lactations of 2118 paternal half-sib sires from 109 families. Number of daughter lactations per sire ranged from 11 to 22 516 with a median value of 558.

Genotyping, linkage maps and phase inference

Two-thousand-four-hundred and eighty paternal half-sib sires and 109 founding sires were genotyped with the bovine Affymetrix 25K MIP array (Gibbs *et al.* 2009; Matukumalli *et al.* 2009). A SNP filtering process, considering discordants (<2.5%), minor allele frequency (>0.025) and genotyped percentage (>75%) for each SNP was performed, which reduced the number of usable assays to 17 349. These SNPs were used to construct male linkage maps for the 29 BTAs that in turn were used to assist assembly of the bovine genome sequence (Liu *et al.* 2009). Locally developed software was further used to determine phased chromosomes and impute missing genotypes. This software used information generated from modified versions of CRI-MAP 2.4 (Green *et al.* 1990) and PHASE (Stephens *et al.* 2001).

Genome-wide association study

GWAS for CM and SCS were performed for 17 349 SNPs to estimate marker effects. Genotype and phenotype information for 2086 sires for CM and for 2118 sires for SCS were included in the analyses. Data for both CM and SCS were divided into two datasets, which were analysed independently. This division was done by listing the families by grandsire identity number, subscribing every second family to subset 1 and the remaining families to subset 2, for each of the two traits. GWAS was performed on the two datasets

Table 1 Records on occurrence of CM in granddaughters were divided into seven time periods (CM1, CM2, CM3, CM4, CM5, CM6 and CM7) as described below. Time period (TP), lactation, days postpartum, number of records (*N*) and standard deviation (SD) for DYD of the trait are given.

TP	Lactation	Days postpartum	*N*	SD (DYD)
CM1	1	−15 to 30	1389776	0.02922895
CM2	1	31 to 120	1375776	0.01474941
CM3	1	121 to 305	1283469	0.01943932
CM4	2	−15 to 30	989525	0.02871001
CM5	2	31 to 305	885345	0.04020668
CM6	3	−15 to 30	632262	0.03633787
CM7	3	31 to 305	543408	0.04852926

CM, clinical mastitis; DYD, Daughter-yield-deviations.

and on the combined dataset for all 17 349 SNPs for both CM and SCS. The mixed model was:

$$P_i = Xg_j + Ya_i + Zm_k + e_{ijk}$$

Here phenotypic value P is daughter-yield-deviations of sire i, weighted by number of daughters; g is fixed effect of grandsire j, a is random effect of sire i, where co-variance structure between sires is determined from pedigree relationships, m is random effect of genetic marker k and e is an error term. The polygenic (a) component was fitted to remove the effect of population stratification, for example due to large half-sib families. Macleod *et al.* (2010) demonstrated that including an effect of sire based on pedigree relationships reduces the number of false positives due to population stratification in a genome scan.

Combined linkage disequilibrium and linkage analysis

Linkage disequilibrium linkage analysis (LDLA) implements historical recombination events in addition to recombination events within genotyped families in order to estimate haplotype effects. Phased chromosomes from all 2589 genotyped sires were included in the analysis along with phenotype information for 2086 sires for CM. The analysis was performed using the GRIDQTL LDLA software in April 2009 (Hernandez-Sanchez *et al.* 2009). Effective male and female population sizes were set to 200 and 10 000 respectively. Number of discrete generations as population foundation was set to 100 (e.g. Meuwissen & Goddard 2001). Identity-by-decent (IBD) probabilities were calculated based on 5-marker haplotypes for positions separated by 1 Mb intervals along each chromosome. The mixed model was:

$$P_i = Xb_{ij} + Zh_k + e_{ijk}$$

Here phenotypic value P is DYD of sire i, b contains fixed effects for sire i and grandsire j, h is random effect of haplotype k and e is an error term. For further details see Meuwissen & Goddard (2001).

SNP effects

For the detected QTLs the effect of alleles and genotypes of the most significant SNP on DYD for CM were calculated. A model including the same fixed effects as those used for the association analyses was used. In addition, haplotype combination was included as a fixed effect. Effects were found by dividing predicted value relative to a mean of zero by DYD standard deviation, giving predicted standard deviations from mean DYD. Standard deviations for CM DYDs in the seven time periods are given in Table 1.

Test for multiple QTL

Multiple QTL analysis was performed to find out if any of the putative QTL regions for CM contained more than one QTL. SNPs showing the highest test score within each QTL region were modelled as fixed effects to see if this influenced the test scores of the other markers in the region. The mixed model was:

$$P_i = Vs + Xg_j + Ya_i + Zm_k + e_{ijk}$$

Here s is fixed effect of the SNP, with the strongest marker-trait association in the QTL region. Remaining terms are described above for the GWAS.

Test score

The likelihood ratio test (LRT) was used for hypothesis testing. LRT scores were calculated as two times the log-likelihood (LogL) ratio. LogL ratio values were obtained with the ASREML software (Gilmour *et al.* 2000) for each SNP, as the difference between the log-likelihood of a model containing the SNP or haplotype effect and the log-likelihood of a model not containing this effect. LRT scores were expected to be distributed as a mixture of two chi-square distributions with 0 and 1 degrees of freedom. For GWAS on the two independent datasets, a LRT score larger than 2.7 (P-value ≤ 0.05) in both datasets was considered a significant marker-trait association, whereas for GWAS on the combined dataset, a LRT score larger than 13.81 (P-value ≤ 0.0001) was considered, significant marker-trait association. A logarithm of odds (LOD) score >3, corresponding to a LRT score >13.81, is an indication of genome-wide significance (Lander & Botstein 1989). The GRIDQTL LDLA software analysis was expected to give higher on-average LRT scores than single-marker association tests (Hernandez-Sanchez *et al.* 2009). An association was considered confirmed by LDLA if a LRT score above 20 from the LDLA analysis was found for a position within 10 Mb of a significant association from GWAS. The LOD drop-off method (Lander & Botstein 1989) was used to find approximate confidence intervals (CIs) from LDLA LRT scores for confirmed QTL regions. The CIs were defined as including positions within

 doi: 10.1111/j.1365-2052.2010.02165.x

4.6 LRT of the maximum on both sides of the LRT peak for LDLA, yielding an approximate 96.8% CI (Mangin *et al.* 1994).

Results

GWAS for clinical mastitis and somatic cell score were conducted on two independent datasets as well as a combined dataset. Significant marker-trait associations in both of the two independent datasets for CM were found for 10 SNPs positioned on chromosomes 2, 4, 6, 9, 17 and 20 and for SCS for four SNPs positioned on chromosomes 12, 19 and 26. Significant marker-trait associations in the combined dataset for CM were found for 26 SNPs positioned on 10 chromosomes and for SCS for 11 SNPs positioned on six chromosomes. A summary of GWAS results are given in Table 2 for CM and in Table 3 for SCS. The analysis revealed no SNPs showing consistently significant marker-trait associations for both CM and lactation average SCS. Correlations between SNP effects on SCS and occurrence of CM in the seven lactational time periods based on all 17 349 SNPs are presented in Table 4. Higher correlations were found between SNP effects on SCS and CM in late lactation than between SNP effects on SCS and CM in the periparturient period. Further, stronger correlations were found between SNP effects on CM in the same phase of lactation than between SNP effects on CM in different phases of lactation. As the trait of interest in this study was susceptibility to mastitis, only putative QTL for CM were investigated further.

LDLA for CM was performed for all chromosome-trait combinations giving significant associations by GWAS in both the independent datasets and for the ten chromosome-trait combinations giving strongest marker-trait associations by GWAS in the combined dataset. LDLA analysis was conducted on BTAs 2, 4, 6 and 20 for CM in time period 1, BTA14 for CM in time period 2, BTA2 for CM in time period 3, BTA6 for CM in time period 4, chromosomes 7, 9 and 17 for CM in time period 5 and BTA6 for CM in time period 6. All LDLA analyses were performed on the total dataset of 2589 paternal half-sib sires from 109 families. The GRIDQTL software was not able to estimate haplotype effects for all positions, presumably due to convergence issues for the remaining positions. For three of the chromosome-trait combinations (BTA6 for time period 4 and 6 and BTA9 for time period 5) the software was only able to estimate haplotype effects for a few positions. Putative QTLs identified by GWAS on the two independent datasets and on the combined dataset on chromosomes 2, 6 and 20 for CM in the periparturient period of first lactation (CM1) were confirmed by LDLA. A putative QTL only identified by GWAS on the combined dataset on BTA14 for CM in late first lactation (CM2) was also confirmed by LDLA. Results of LDLA and GWAS on the combined dataset for these four chromosome-trait combinations are presented in Fig. 1. Highest LRT scores from LDLA were found for positions on BTA2 at 104 Mb, on BTA6 at 95 Mb, on BTA14 at 42 Mb and on BTA20 at 43 Mb. Approximate 98.6% CIs for the four QTL regions based on the LOD drop-off method (Lander & Botstein 1989) included the regions 103.4–104.3 Mb on BTA2, 94.2–95.3 Mb on BTA6, 41.3–42.3 Mb on BTA14 and 41.8–43.2 Mb on BTA20. Highest LRT scores from GWAS on the combined dataset for each of the four QTL were found for SNP BTA-120624 at 103.9 Mb on BTA2, for SNP BTA-119376 at 90.7 Mb on BTA6, for SNP BTA-34923 at 45.2 Mb on BTA14 and for SNP BTA-19985 at 43.3 on BTA20.

For each of these four QTLs, effects of alleles and genotypes for the most significant SNP by GWAS for CM are shown in Table 5. Results are in DYD standard deviations from the mean DYD for CM. Difference in effect on DYD between alleles was largest for SNP BTA-119376 on BTA6 for CM in the periparturient period of first lactation. This SNP also gave the highest test score from GWAS on the combined dataset for CM among all 17 349 SNPs. For SNP BTA-34923 on BTA14 and SNP BTA-19985 on BTA20, low minor allele frequencies (MAF) contributed to large standard deviation relative to effect sizes, resulting in more uncertain effect estimates.

All the four QTL regions contained several SNPs significantly associated with occurrence of CM (see Table 2). Therefore, tests for multiple QTL were performed on the combined dataset to find out whether any of these regions contained more than one QTL. For all four QTL regions, LRT scores of the remaining SNPs were lowered to non-significant levels (<13.81) when the SNP with the highest LRT score in each of the regions was included in the model as a fixed effect. Highest test scores found in each region by the multiple QTL test were for SNP BTA-122697 on BTA2 at 112.8 Mb with LRT score 5.16, SNP BTA-86975 on BTA6 at 88.4 Mb with LRT score 4.82, SNP rs29018717 on BTA14 at 61.3 Mb with LRT score 6.66 and SNP BTA-50239 on BTA20 at 35.5 Mb with LRT score 9.04.

Discussion

GWAS performed on two independent datasets identified 10 SNPs positioned on chromosomes 2, 4, 6, 9, 17 and 20 that were significantly associated with occurrence of CM (LRT >2.7 in both datasets). GWAS on the combined dataset identified 26 SNPs positioned on 10 chromosomes significantly associated with occurrence of CM (LRT >13.81). Three QTLs on BTAs 2, 6 and 20 for CM in the periparturient period and a QTL on BTA14 for CM in late lactation were confirmed by LDLA. The QTL on BTA14, confirmed by LDLA, was detected by GWAS on the combined dataset but not by GWAS on the two independent datasets. This could be an indication that the requirement for associations to be significant in two independent datasets (LRT >2.7) is a more conservative test than the

Table 2 Significant marker-trait associations from GWAS on two independent datasets and on the combined dataset for CM on the 29 *Bos taurus* autosomes (BTAs). BTA, SNP, position (Mb), LRT score for datset 1 (D1), dataset 2 (D2) and the combined dataset (CD) and time period (TP) are given.

BTA	SNP	Position (bp)	D1 LRT	D2 LRT	CD LRT	TP
2	BTA-47902	68 074 185	4.92	3.26	15.50	CM3
2	BTA-120621	103 854 622	3.29	3.05	12.72	CM1
2	BTA-120624	103 892 096	5.03	3.26	15.52	CM1
2	rs29025784	112 396 633	4.06	4.10	14.62	CM1
4	rs29020694	90 418 076	5.65	3.23	20.14	CM1
6	BTA-119376	90 670 190	3.28	12.74	32.36	CM1
6	BTA-119376	90 670 190	2.76	3.16	15.46	CM6
6	BTA-77136	94 544 954			18.80	CM4
6	BTA-109071	95 256 811			14.18	CM6
6	BTA-77356	96 189 520	3.08	5.18	15.60	CM1
7	BTA-78563	22 841 729			16.92	CM5
7	BTA-99486	26 347 331			16.80	CM5
9	BTA-84619	88 416 414	2.95	2.71		CM5
10	BTA-79349	92 455 424			14.20	CM6
13	rs29022774	62 928 524			13.90	CM4
13	rs29022775	62 928 533			15.90	CM4
14	rs29012803	17 289 087			14.90	CM5
14	BTA-34796	40 769 096			15.36	CM2
14	BTA-34923	45 153 040			18.34	CM2
14	BTA-111421	47 425 522			15.96	CM2
16	BTA-38543	33 335 777			14.64	CM1
17	BTA-103789	34 861 876	3.02	2.76	12.84	CM5
20	BTA-25160	31 659 731			15.16	CM2
20	BTA-50239	35 530 051	2.96	3.84	16.02	CM1
20	BTA-50236	35 861 339			17.10	CM1
20	rs29021255	38 261 470			14.88	CM3
20	BTA-19985	43 267 496			20.76	CM1
20	BTA-22852	50 780 080			14.36	CM5
29	rs29027496	45 602 144			14.36	CM1

CM, clinical mastitis; GWAS, genome-wide association studies; LRT, likelihood ratio test.

higher significance threshold value (LRT >13.81) used for the combined dataset. Inability to detect a QTL in a split-dataset analysis could also be caused by lowered power due to fewer observations in each dataset. Highest test scores from GWAS on the combined dataset for each of the four QTLs were found for SNP BTA-120624 at 103.9 Mb on BTA2, for SNP BTA-119376 at 90.7 Mb on BTA6, for SNP BTA-34923 at 45.2 Mb on BTA14 and for SNP BTA-19985 at 43.3 on BTA20. Approximate 98.6% CIs were found by the LOD drop-off method for all four QTLs based on LDLA test scores. The CI on BTA2 included the SNP with the highest test score from GWAS in this region. The CIs on BTA6, BTA14 and BTA20 did not include the SNPs with the highest test score from GWAS from each of these regions. For LDLA, LRT scores were only estimated for every 1 Mb across each chromosome and unfortunately the GRIDQTL software was not able to estimate haplotype effects for all positions. Construction of CIs based on LDLA analysis might therefore be an inaccurate approach for this study.

Table 3 Significant marker-trait associations from GWAS on two independent datasets and on the combined dataset for SCS on the 29 *Bos taurus* autosomes (BTAs). BTA, SNP, position (Mb) and LRT score for datset 1, dataset 2 and the combined dataset are given.

BTA	SNP	Position (bp)	D1 LRT	D2 LRT	CD LRT
8	BTA-120681	27 391 619			14.00
8	BTA-102648	30 020 795			14.34
12	rs29021760	63 796 167			14.46
12	BTA-28030	66 114 084	2.75	3.16	13.76
17	rs29019473	52 225 317			13.98
17	rs29019471	52 229 956			14.26
19	ss46526232	42 098 180			16.20
19	BTA-45702	50 276 021	4.28	3.73	14.86
19	BTA-45709	50 359 720	5.44	3.64	15.66
20	BTA-95391	11 740 236			14.08
21	rs29016404	33 141 160			25.12
21	BTA-52383	46 299 936			14.06
26	BTA-61841	46 096 255	2.88	2.98	

GWAS, genome-wide association studies; LRT, likelihood ratio test; SCS, somatic cell score.

Multiple QTL analysis did not provide evidence of more than one QTL in any of the four regions. However, a relatively high LRT score of 9.04 was found for a SNP at 3.53 Mb on BTA20 when including the SNP at 43.27 Mb as a fixed effect. This LRT score is below our threshold for GWAS on the combined dataset, but the presence of an additional QTL in this region cannot be completely ruled out.

While the QTL on BTA20 has not previously been reported, the QTLs on BTAs 2, 6 and 14 are supported by other studies. A QTL for SCS in German Holsteins has been reported on BTA2 at 100 cM (Bennewitz *et al.* 2003) close to our most significant SNP for CM on BTA2 at 103.9 Mb. This QTL was not found to affect SCS in our study. For NRF, records for SCS are retrieved as lactation averages and are thus not directly comparable with the CM recordings. This may reduce the power of QTL detection for SCS. Our analyses showed no consistencies between QTL for CM and for SCS. Similar results have been reported for the NRF population in a previous study (Klungland *et al.* 2001).

Correlations in SNP effects based on all 17 348 SNPs were higher between SCS and CM in late lactation than between SCS and CM in the periparturient period, supporting a previous report on genetic correlations between SCS and CM in different stages of lactation (Svendsen & Heringstad 2006b).

QTLs affecting both milk production traits and CM on BTA6 around 90 Mb have previously been reported for NRF (Nilsen *et al.* 2009) in a study including the same population and many of the same individuals as were included in the GWAS described here. This is in the same region as the QTL detected at approximately 90 Mb on BTA6 in this study. QTLs affecting susceptibility to CM on BTA14, close to our most significant SNP on BTA14 at 45.2 Mb, have been reported by both Lund *et al.* (2007) and Schulman *et al.* (2004). More milk production trait QTLs have been

Table 4 Correlations between SNP effects on SCS and CM in the seven lactational time periods (the seven lactational time periods are described in Table 1).

	CM1	CM2	CM3	CM4	CM5	CM6	CM7
SCS	0.04	0.11	0.11	0.08	0.15	0.10	0.12
CM1		0.14	0.08	0.34	0.15	0.22	0.07
CM2			0.30	0.16	0.33	0.11	0.21
CM3				0.10	0.31	0.07	0.17
CM4					0.13	0.33	0.09
CM5						0.10	0.31
CM6							0.09

CM, clinical mastitis; SCS, somatic cell score.

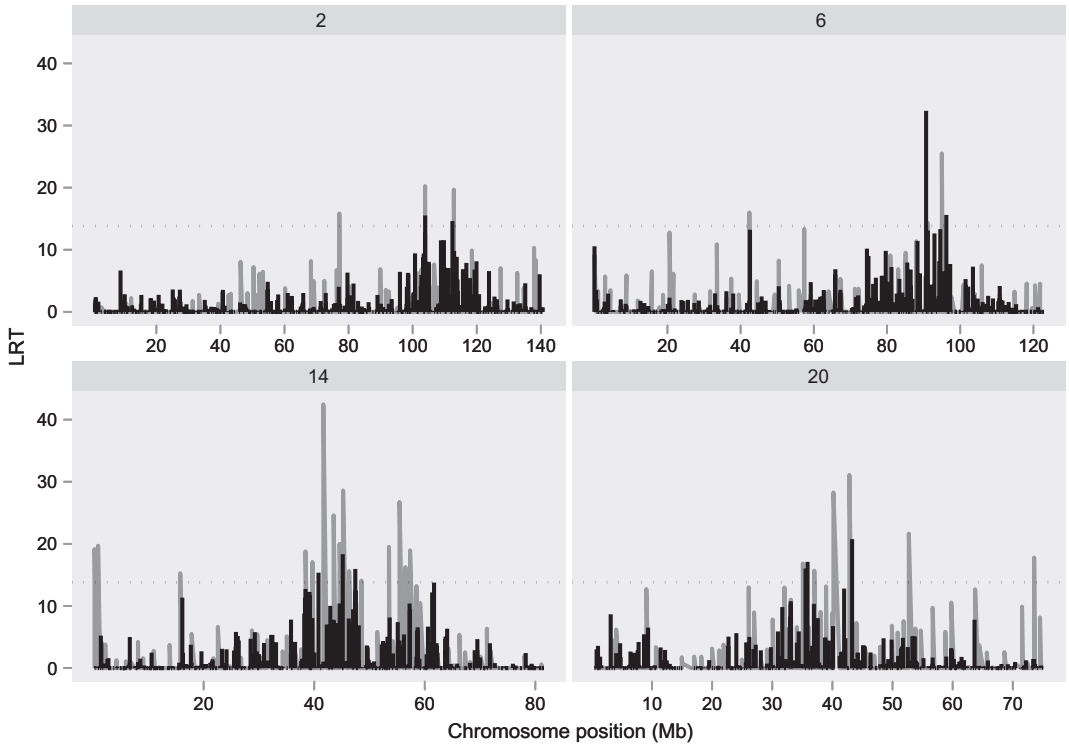

Figure 1 Results from genome-wide association studies (GWAS) and linkage disequilibrium linkage analysis (LDLA) for BTA2 for CM1, BTA6 for CM1 BTA14 for CM2 and BTA20 for CM1. Black bars show likelihood ratio test (LRT) score from GWAS, grey lines show LRT scores from LDLA. Dotted horizontal line indicate LRT threshold for GWAS (LRT >13.81).

 doi: 10.1111/j.1365-2052.2010.02165.x

Table 5 Effects of SNPs on DYD for CM for the most significant SNP in each of the four QTL regions. *Bos taurus* autosome (BTA), SNP, alleles and genotypes, frequencies of alleles and genotypes, effect of alleles and genotypes, standard deviation (SD) of effect and time period (TP) are given. Effects are expressed as DYD standard deviations from the mean value. Standard deviations for DYD for clinical mastitis in the seven time periods are given in Table 1.

BTA	SNP	Alleles	Frequency	Effect	SD (Effect)	TP
2	BTA-120624	A	0.50	−0.09	0.04	CM1
		G	0.50	0.06	0.04	
		AA	0.25	−0.17	0.05	
		AG	0.51	−0.02	0.04	
		GG	0.24	0.15	0.05	
6	BTA-119376	C	0.26	0.20	0.05	CM1
		T	0.74	−0.01	0.03	
		CC	0.07	0.36	0.09	
		CT	0.38	0.09	0.04	
		TT	0.55	−0.11	0.04	
14	BTA-34923	C	0.19	−0.03	0.03	CM2
		T	0.81	−0.04	0.03	
		CC	0.03	−0.05	0.17	
		CT	0.31	0.01	0.04	
		TT	0.66	−0.06	0.03	
20	BTA-19985	C	0.19	−0.01	0.03	CM1
		G	0.81	−0.01	0.03	
		CC	0.04	−0.05	0.12	
		CG	0.31	−0.04	0.05	
		GG	0.66	0.00	0.04	

CM, clinical mastitis; DYD, Daughter-yield-deviations; QTL, quantitative trait loci.

reported for chromosomes 6, 14 and 20 than for the other bovine chromosomes (Khatkar *et al.* 2004).

Genetic susceptibility to CM could have different biological causes at different stages of lactation (Waller 2000; Sordillo 2005; Østeras *et al.* 2006), which could reduce power to detect QTLs affecting susceptibility to CM only in a particular phase of lactation. In this study, phenotypic records on occurrence of CM were divided into seven time periods (Table 1) and thus the occurrence of CM in each stage of lactation was treated as a different trait. There are more records for CM incidents in the periparturient period, as approximately two-thirds of incidents occur in the two-first months after lactation. Consequently, this study had more power to detect QTLs for CM in the periparturient period than for CM in later lactation. The study also had more power to detect QTLs for CM in first lactation than in later lactations. Three of the four QTLs detected here were for CM in the periparturient period of first lactation. Correlations based on all 17 349 SNPs were higher between SNP effects on CM in the same phase of lactation than between SNP effects on CM in different phases of lactation. This is in accordance with a report on genetic correlation between CM in different phases of lactation (Svendsen & Heringstad 2006a) and supports division of records on CM into lactational time periods.

The strongest marker-trait association by GWAS on the combined dataset was between SNP BTA-119376 at 90.7 Mb on BTA6 and clinical mastitis in the periparturient period of first lactation (CM1). Genes coding for most of the C-X-C motif chemokines are clustered near this SNP on BTA6. In particular, the gene coding for interleukin 8 (IL8) is positioned on BTA6 at 91.78 Mb. IL8 is a C-X-C motif chemokine important for initial recruitment of circulating neutrophils to the site of infection. In NRF, *Staphylococcus aureus* is the most common cause of CM (Østeras *et al.* 2006). Gram positive bacteria such as *S. aureus* are assumed to initiate expression of IL8 by toll-like receptor 2 activation after the cell wall components lipoteichoic acid or peptidoglycan (Bannerman *et al.* 2004). There have been inconsistent reports on the ability of different cattle breeds to express IL8 upon exposure to *S. aureus* (Bannerman *et al.* 2004; Strandberg *et al.* 2005; Lahouassa *et al.* 2007; Griesbeck-Zilch *et al.* 2008; Yang *et al.* 2008). These inconsistencies could be due to genetic differences between cattle, or due to different strains of *S. aureus* being investigated in the different studies.

Candidate genes associated with immunological defence are also found close to other QTL. On BTA2, genes coding for IL8 receptors chemokine (C-X-C motif) receptor 2 (*CXCR2*) and chemokine (C-X-C motif) receptor 1 (*CXCR1*), are positioned at 110.6 Mb and 110.6 Mb, between SNPs BTA-120624 and rs29025784 at 103.9 Mb and 112.4 Mb. Both these SNPs showed significant association with CM in the periparturient period of first lactation (CM1), with LRT values 15.52 and 14.62. Polymorphisms in *CXCR2* have been associated with the occurrence of subclinical mastitis and ability to recruit neutrophils to the site of infection (Youngerman *et al.* 2004; Rambeaud & Pighetti 2005). Polymorphisms in CXCR1 have been associated with SCS level and expression level of CXCR1 (Leyva-Baca *et al.* 2008a,b). The IL8 receptors have other C-X-C motif chemokines as ligands in addition to IL8. As mentioned above, the QTL on BTA2 was not found to affect lactation average SCS level in this study.

Three genes coding for complement components 6, 7 and 9 (C6, C7 and C9) are positioned at 34.4 Mb, 35.7 Mb and 37.3 Mb on BTA20. By GWAS, four SNPs in the interval 35.5–45.6 Mb on this chromosome showed significant association with CM in the periparturient period. Complement components are an important part of immunological defence, being involved in inflammation, phagocytosis, attack on bacterial membranes and promotion of antibody production. Although the genes mentioned above could potentially have an effect on susceptibility to CM in NRF, the number of other potential candidate genes in these regions is quite high. It may be of particular that our two most significant SNPs on BTA2 are near the genes coding for

 doi: 10.1111/j.1365-2052.2010.02165.x

CXCR1 and CXCR2, whereas our most significant SNP on BTA6 is near the gene coding for IL8.

Acknowledgements

Thanks to H. Henriksen and T. Nome for technical assistance and to GENO Breeding and AI association for providing relationship information and DYD values. This project has been funded by The Research Council of Norway, GENO Breeding and AI association and BoviBank Ltd.

References

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.

Bennewitz J., Reinsch N., Grohs C. *et al.* (2003) Combined analysis of data from two granddaughter designs: a simple strategy for QTL confirmation and increasing experimental power in dairy cattle. *Genetics, Selection, Evolution* **35,** 319–38.

Elsik C.G., Tellam R.L., Worley K.C. *et al.* (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* **324,** 522–8.

Gibbs R.A., Taylor J.F., Van Tassell C.P. *et al.* (2009) Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* **324,** 528–32.

Gilmour A.R., Cullis B.R., Welham S.J. & Thompson R. (2000) *ASREML reference manual*, New South Wales Agriculture, NSW, Australia.

Green P., Falls K. & Crooks S. (1990) *Documentation for CRI-MAP, version 2.4*. Washington University School of Medicine, St. Louis, MO, USA.

Griesbeck-Zilch B., Meyer H.H., Kuhn C.H., Schwerin M. & Wellnitz O. (2008) *Staphylococcus aureus* and *Escherichia coli* cause deviating expression profiles of cytokines and lactoferrin messenger ribonucleic acid in mammary epithelial cells. *Journal of Dairy Science* **91,** 2215–24.

Heringstad B., Klemetsdal G. & Ruane J. (2000) Selection for mastitis resistance in dairy cattle: a review with focus on the situation in the Nordic countries. *Livestock Production Science* **64,** 95–106.

Heringstad B., Chang Y.M., Gianola D. & Klemetsdal G. (2005) Genetic association between susceptibility to clinical mastitis and protein yield in norwegian dairy cattle. *Journal of Dairy Science* **88,** 1509–14.

Hernandez-Sanchez J., Grunchec J.A. & Knott S. (2009) A web application to perform linkage disequilibrium and linkage analyses on a computational grid. *Bioinformatics* **25,** 1377–83.

Khatkar M.S., Thomson P.C., Tammen I. & Raadsma H.W. (2004) Quantitative trait loci mapping in dairy cattle: review and meta-analysis. *Genetics, Selection, Evolution* **36,** 163–90.

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.

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.

Lahouassa H., Moussay E., Rainard P. & Riollet C. (2007) Differential cytokine and chemokine responses of bovine mammary epithelial cells to *Staphylococcus aureus* and *Escherichia coli*. *Cytokine* **38,** 12–21.

Lander E.S. & Botstein D. (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121,** 185–99.

Leyva-Baca I., Schenkel F., Sharma B.S., Jansen G.B. & Karrow N.A. (2007) Identification of single nucleotide polymorphisms in the bovine *CCL2, IL8, CCR2* and *IL8RA* genes and their association with health and production in Canadian Holsteins. *Animal Genetics* **38,** 198–202.

Leyva-Baca I., Pighetti G. & Karrow N.A. (2008a) Genotype-specific *IL8RA* gene expression in bovine neutrophils in response to *Escherichia coli* lipopolysaccharide challenge. *Animal Genetics* **39,** 298–300.

Leyva-Baca I., Schenkel F., Martin J. & Karrow N.A. (2008b) Polymorphisms in the 5′ upstream region of the *CXCR1 chemokine receptor* gene, and their association with somatic cell score in Holstein cattle in Canada. *Journal of Dairy Science* **91,** 407–17.

Liu Y., Qin X., Song X.Z. *et al.* (2009) *Bos taurus* genome assembly. *BMC Genomics* **10,** 180.

Lund M.S., Sahana G., Andersson-Eklund L. *et al.* (2007) Joint analysis of quantitative trait loci for clinical mastitis and somatic cell score on five chromosomes in three Nordic dairy cattle breeds. *Journal of Dairy Science* **90,** 5282–90.

Macleod I.M., Hayes B.J., Savin K.W., Chamberlain A.J., Mcpartlan H.C. & Goddard M.E. (2010) Power of a genome scan to detect and locate quantitative trait loci in cattle using dense single nucleotide polymorphisms. *Journal of Animal Breeding and Genetics* **127,** 133–42.

Mangin B., Goffinet B. & Rebai A. (1994) Constructing confidence intervals for QTL location. *Genetics* **138,** 1301–8.

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.

Meuwissen T.H. & Goddard M.E. (2001) Prediction of identity by descent probabilities from marker-haplotypes. *Genetics, Selection, Evolution* **33,** 605–34.

Mrode R.A. & Swanson G.J.T. (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* **64,** 847–857.

Nilsen H., Olsen H.G., Hayes B., Nome T., Sehested E., Svendsen M., Meuwissen T.H. & Lien S. (2009) Characterization of a QTL region affecting clinical mastitis and protein yield on BTA6. *Animal Genetics* **40,** 701–12.

Østeras O., Kruse H., Sølverød L., Gjestvang J. & Mørk T. (2006) Nordic View Concerning Mastitis Pathogen Resistance. *Proceedings NMC 45th Annual Meeting*. Tampa, Florida.

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.

Rambeaud M. & Pighetti G.M. (2005) Impaired neutrophil migration associated with specific bovine *CXCR2* genotypes. *Infection and Immunity* **73,** 4955–9.

Sahana G., Lund M.S., Andersson-Eklund L. *et al.* (2008) Fine-mapping QTL for mastitis resistance on BTA9 in three Nordic red cattle breeds. *Animal Genetics* **39,** 354–62.

Schulman N.F., Viitala S.M., De Koning D.J., Virta J., Maki-Tanila A. & Vilkki J.H. (2004) Quantitative trait Loci for health traits in Finnish Ayrshire cattle. *Journal of Dairy Science* **87,** 443–9.

Sordillo L.M. (2005) Factors affecting mammary gland immunity and mastitis susceptibility. *Livestock Production Science* **98,** 89–99.

Stephens M., Smith N.J. & Donnelly P. (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* **68,** 978–89.

Strandberg Y., Gray C., Vuocolo T., Donaldson L., Broadway M. & Tellam R. (2005) Lipopolysaccharide and lipoteichoic acid induce different innate immune responses in bovine mammary epithelial cells. *Cytokine* **31,** 72–86.

Svendsen M. & Heringstad B. (2006a) New Genetic Evaluation for Clinical Mastitis in Multiparous Norwegian Red Cows. *Interbull Bulletine* **35,** 8–11.

Svendsen M. & Heringstad B. (2006b) Somatic Cell Count as an Indicator of Subclinical Mastitis. Genetic Parameters and Correlations with Clinical Mastitis. *Interbull Bulletine* **35,** 12–16.

Syvajarvi J., Saloniemi H. & Grohn Y. (1986) An epidemiological and genetic study on registered diseases in Finnish Ayrshire cattle. IV. Clinical mastitis. *Acta veterinaria Scandinavica* **27,** 223–34.

Viguier C., Arora S., Gilmartin N., Welbeck K. & O'kennedy R. (2009) Mastitis detection: current trends and future perspectives. *Trends in Biotechnology* **27,** 486–93.

Waller K.P. (2000) Mammary gland immunology around parturition. Influence of stress, nutrition and genetics. *Advances in experimental medicine and biology* **480,** 231–45.

Weller J.I., Kashi Y. & Soller M. (1990) Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *Journal of Dairy Science* **73,** 2525–37.

Yang W., Zerbe H., Petzl W., Brunner R.M., Gunther J., Draing C., Von Aulock S., Schuberth H.J. & Seyfert H.M. (2008) Bovine TLR2 and TLR4 properly transduce signals from *Staphylococcus aureus* and *E. coli*, but *S. aureus* fails to both activate NF-kappaB in mammary epithelial cells and to quickly induce TNFalpha and interleukin-8 (CXCL8) expression in the udder. *Molecular Immunology* **45,** 1385–97.

Youngerman S.M., Saxton A.M., Oliver S.P. and Pighetti G.M. (2004) Association of *CXCR2* polymorphisms with subclinical and clinical mastitis in dairy cattle. *J Dairy Sci* **87,** 2442–8.