Proceedings of the World Congress on Genetics Applied to Livestock Production, 11.167

Genome-wide Association Study for Alternative Subclinical Mastitis Traits in Norwegian Red

E. Kirsanova[1], B. Heringstad[2, 3], A. Lewandowska-Sabat[1], I. Olsaker[1]

[1] Department of Basic Sciences and Aquatic Medicine, Faculty of Veterinary Medicine, Norwegian University of Life Sciences, P.O. Box 8146 Dep, NO-0033 Oslo, Norway.
[2] Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432 Aas, Norway.
[3] Geno Breeding and AI Association, 2317 Hamar, Norway.

Summary

The aim of this study was to identify single nucleotide polymorphism (SNP) markers associated with subclinical mastitis (SCM) in the Norwegian Red (NR) population. Alternative SCM traits were defined based on testday somatic cell count (SCC) records, using thresholds of 50, 100, 150, 200, 250, 300, 350, and 400 x10^3 cells/ml. Genome-Wide Association Study (GWAS) was performed for each trait. Genotype information of 35,605 SNP from 3,795 NR sires with daughter records for SCM and lactation mean somatic cell score (LSCS) were included. In total, 210 significant SNP were identified on chromosome-wide significance level, per trait the number ranged from 17 to 29. Chromosome 26 had the highest number of significant SNP for SCM. Some of the SNP were associated with several traits and other were unique for one or two traits. Four unique SNP were detected on chromosome 3, 9, 15, 22 for SCM50, two on 3 and 19 for SCM100, one on 21 and 12 for SCM150 and SCM200, respectively. Two unique SNP were detected for SCM350 on chromosome 26 and two for SCM400 on chromosome 26. LSCS had two unique SNP on chromosome 4 and 19. The LSCS trait had the highest number of common significant SNP with SCM50 (18) and lowest with SCM350 and SCM400 (5).

Keywords: Genome-wide association study, Subclinical mastitis, Norwegian Red, Somatic Cell Count

Introduction

Mastitis is a complex trait, varying from mild subclinical to severe clinical cases, often with increasing of somatic cell count (SCC). In the Norwegian Dairy Herd Recording System (NDHRS) both clinical mastitis (CM) and SCC has been recorded since the late 1970s and both traits are included in the breeding goals for Norwegian Red (NR).

Using Genome-Wide Association Study (GWAS) on single nucleotide polymorphism (SNP) data for NR, Sodeland et al. (2011) reported significant Quantitative Trait Loci (QTL) affecting CM on bovine chromosome 2, 4, 6, 9, 17 and 20 and affecting lactation average Somatic Cell Score (LSCS) on bovine chromosome 12, 19 and 26. The genetic correlation between CM and SCC in cattle varies, with an average of 0.7 (Mrode & Swandson 1996),

while the genetic correlation between subclinical mastitis (SCM) and SCC was reported to vary from 0.83 to 0.99 in NR (Svendsen & Heringstad 2006).

SCC is often used as an indicator of clinical and subclinical mastitis. Estimated genetic correlation between SCC and CM or SCM vary from 0.55 to 0.93 and 0.98, respectively (de Haas et al. 2008). Alternative SCC traits have been suggested by several authors (Svendsen & Heringstad 2006; de Haas et al. 2008), however, no GWAS for such traits for NR was performed so far. Hence, this paper report the results of GWAS identifying associations between SNP and alternative SCM traits in NR.

Material and methods

Animals and phenotypic data

Data from the NDHRS was used, including a total of 7,300,847 testday SCC records from the period 1979 to 2016 for 3,543,764 NR cows. The corresponding pedigree file included 4,126,672 individuals. Testday records between 5 and 305 days in milk from cows with two or more SCC records were included in the analyses. SCM was defined as a binary trait for each of the first 3 lactations, defined as 1 if two SCC records in a row were above the fixed threshold during a 80 day period, 0 otherwise. Eight SCM traits based on different SCC thresholds (50, 100, 150, 200, 250, 300, 350, 400 $x10^3$ cells/ml) were analysed. Testday SCC was $\log_e$ transformed to Somatic Cell Score (SCS) and lactation mean SCS (LSCS) was calculated as the arithmetic mean per lactation. The mean frequency for the SCM traits from first to third lactation based on 2,347,706, 1,673,956 and 1,031,092 phenotypic records, respectively, is presented in Table 1.

Table 1. Estimated heritability, and the mean frequency for subclinical mastitis (SCM) traits (50, 100, 150, 200, 250, 300, 350, 400 $x10^3$ cells/ml) overall and for lactation 1-3.

Trait	Heritability	Frequency			
		1st	2nd	3rd	Total
SCM50	0.12	42.1	57.6	64.6	51.8
SCM100	0.12	24.6	36.4	42.8	32.2
SCM150	0.10	16.3	25.1	30.4	22.1
SCM200	0.08	11.5	18.3	22.6	16.0
SCM250	0.07	8.5	13.8	17.3	12.1
SCM300	0.07	6.5	10.7	13.7	9.4
SCM350	0.05	5.0	8.5	11.0	7.4
SCM400	0.04	4.0	6.9	9.0	5.6
LSCS	0.19				

Variance components for each trait were estimated by DMU (Madsen & Jensen 2013) using linear animal models. SCM and LSCS were assumed to be the same trait across lactations and analysed using repeatability models. Daughter-Yield-Deviation (DYD) were calculated as average yield deviation (YD) among daughters for each sire. The number of YD per sire ranged from 62 to 28,097 with a median value of 1,404. YD is the cow's phenotype corrected for fixed effects (calving year/month, lactations number, age at calving, days open), random

effects (herd year, permanent environment), and 0.5 of dams estimated breeding value.

Genotypes

An imputed SNP dataset for NR was available. Genotyping of NR has been performed using the Affymetrix 25K SNP array (Affymetrix, Santa Clara), the Illumina BovineSNP50 BeadChip (54K) (Illumina Inc., San Diego, CA; Matukumalli et al. 2009) and 55K Affymetrix SNP array (Affymetrix, Santa Clara). Imputation of missing genotypes and estimation of genomic relationships matrix were performed by Geno, the breeding organisation for NR (www.geno.no). The SNP positions were based on the Bos taurus genome UMD 3.1 assembly (Zimin et al. 2009). For the quality control of SNP data, a minimum call rate of 95 %, an individual call rate of 85 % and a minor allele frequency of 1 % were used. A total of 35,605 SNP located on 29 chromosomes passed the control criteria and were used in the analyses.
The significant SNP were defined at the chromosome-wide significance level according to Sahana et al. (2010). The 5 % chromosome-wide significance threshold ranged from genome-wide corrected p-value of 2.35×10^{-5} on chromosome 1 to 7.79×10^{-5} on chromosome 27. Following the chromosomewide significance level and a suggestive association the SNP with p-value $< 5 \times 10^{-4}$ were defined as significant.

Genome-Wide Association Study

GWAS analyses were performed based on genotype information from 3,795 NR sires with DYD values for the 8 SCM traits and LSCS. The GRAMMAR-Gamma (Svischeva et al. 2012) function in the R v.3.3.2 package GenABEL (www.r-project.org; Aulchenko et al. 2007) was implemented to estimate the residuals using pre-corrected phenotypes as described above, for evaluation of each SNP marker effect.

Results and discussion

The GWAS analyses of the SCM and LSCS traits revealed a total of 210 SNP associated on chromosome-wide significant level to one or more of the nine traits on 20 chromosomes. Results are shown in Table 2 and Figure 1.

Table 2. Number of SNP with significant associations with subclinical mastitis (SCM; 50, 100, 150, 200, 250, 300, 350, 400 x10^3 cells/ml) and lactation mean somatic cell score (LSCS) on chromosome-wide significance level.

		Bovine chromosome																			
Trait	Sum	1	3	4	6	8	9	10	11	12	13	15	17	18	19	20	21	22	23	24	26
SCM50	23		4		1	1	1				1	3	1		2	3		2	1		3
SCM100	29		4	1	1	1		1	2		1	1	1		4	3		1	1		7
SCM150	24	1	2	1	1	1		1			1	1	1		3	1	1	1	2		6
SCM200	25	1	3	4	1	1		1		1	1	1	1	1	3	1		1	2		2
SCM250	21	1	1	3	1	1		1			1	1	1	1	4	1			1		3
SCM300	24	1	1	4	1	1		1				1		1	2	1		1	1		8
SCM350	22			3		1		1	1			3	1	1	4			1	2	1	3

SCM400	17						1			1		1					3		1				3						4		1	2
LSCS	25		4		1	1	1					1			1		2		1		1		3		3			1		1			4

SCM50	SCM200	SCM400

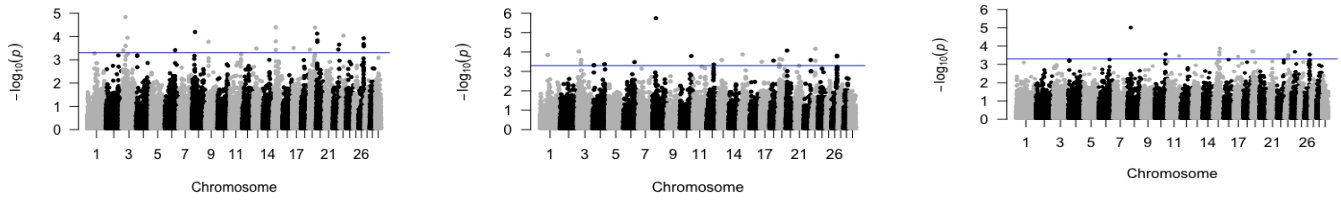

Figure 1. Manhattan plot of genome-wide associations for subclinical mastitis (SCM) traits SCM50, SCM200 and SCM400. Blue line shows defind chromosome-wide significance threshold on $-\log_{10(p)} = 3.30103$ ($p = 5 \times 10^{-4}$) of the genome-wide corrected p-values.

The number of significant associations per trait varied from 17 (SCM400) to 29 (SCM100) and chromosome 26 had the highest number of significant SNP for SCM, followed by chromosome 19. The SCM traits with the lower thresholds, SCM50 – SCM200, had higher number significant SNP compared to the higher thresholds SCM250 – SCM400. Some of the SNP were associated with several traits and other were unique for one or two SCM traits: unique SNP were detected for SCM50 on chromosome 9, 3, 22 and 15, for SCM100 on chromosome 19, 3 and 11, for SCM150 on chromosome 21, for SCM200 on chromosome 12, for SCM300 and SCM350 on chromosome 26, and for LSCS on chromosome 4 and 19. In general, results for LSCS confirm the previous studies (Sodeland et al. 2011; Klungland et al. 2001; Meredith et al. 2012; Meredith et al. 2013). However, in this study supplementary significant SNP on chromosome 11 and 18 were detected. The results for LSCS were most similar to SCM50 with 18 significant SNP in common and having the least common number of significant SNP (5 each) with SCM350 and SCM400. The most significant SNP for traits SCM150, SCM200, SCM250, SCM300, SCM350 and SCM400 were found on chromosome 8, while for SCM50, SCM100 and LSCS on chromosome 3, 20 and 15, respectively. The association with different chromosomal regions indicates that the various SCM traits and LSCS represent partly different traits. SCM is a complex trait involving many genes with no major QTL explaining a large proportion of the detected variation.

Conclusion

In the present GWA study, 210 significant SNP positioned on 20 chromosomes identified for SCM and LSCS traits, at 5 % chromosome-wide significant level. The most significant SNP was located on chromosome 15 (LSCS), 3 (SCM50), 20 (SCM100), and 8 (SCM150 – SCM400).

Acknowledgments

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.

List of References

Aulchenko, YS. et al. Bioinformatics. 2007;23(10):1294-6.
De Haas, Y. et al. J Dairy Sci. 2008;91(6):2501-11.
Klungland, H. et al. Mamm Genome. 2001;12(11):837-42.
Madsen, P. & Jensen, J. Version 6, release 5.2. University of Aarhus. 2013.
Matukumalli, LK. et al. PLoS One. 2009;4(4):e5350.
Meredith, BK. et al. Front Genet. 2013;4:229.
Meredith, BK. et al. BMC Genet. 2012;13:21.
Mrode, R. & Swanson, G. Animal Breeding Abstracts (United Kingdom). 1996.
Sahana, G. et al. Anim Genet. 2010;41(6):579-88.
Sodeland, M. et al. Anim Genet. 2011;42(5):457-65.
Svendsen, M. & Heringstad, B. Interbull Bulletin. 2006(35):12.
Svischeva, G. et al. Nature Genetics. 2012;44:1166-1170. doi:10.1038/ng.2410
Zimin AV. et al. Genome Biol. 2009;10(4):R42.