Short Communication: Genes associated with somatic cell count index in Brown Swiss cattle

J. Jeretina[*], D. Jevsinek Skok[*, 1]

[*] Agricultural Institute of Slovenia, Hacquetova ulica 17, SI-1000 Ljubljana, Slovenia

[1] Corresponding author: dasa.jevsinekskok@kis.si

Abstract

Subclinical mastitis (SM) is one of the most common diseases of cows in milk production herds caused by contagious and/or environmental pathogens. Since there are no visible abnormalities in the milk or udder, detection of subclinical mastitis requires special diagnostic tests. Somatic cell count (SCC) is the most common test used to detect changes in milk due to the inflammatory process. Previously we developed somatic cell count index (SCCI), a new method for the accurate prediction of milk yield losses caused by elevated SCC. The aim of this study was to identify new candidate genetic markers for SCCI in the Slovenian population of Brown Swiss cattle (BS). For that purpose, we analyzed samples of BS cows, which were genotyped using SNP microarray ICBF International Dairy and Beef v3 (ICBF, Ireland) for a total of 53,262 single nucleotide polymorphism (SNP) markers. After quality control, the set of 18,136 SNPs were used in association analysis. Our association analysis revealed 130 SNPs associated with SCCI, which were used for haplotype and overlap analysis. Haplotypes generated from the genotyped data for those 130 SNPs revealed 10 haplotype blocks among 22 SNPs. Additionally, all 130 SNPs, mastitis related QTL, and protein-coding genes are shown on bovine genome. Overlap analysis shows that the majority of significantly associated SNPs (70) are intergenic, while 60 SNPs are mapped within, up- or downstream of the protein-coding genes. However, those genes can serve as strong candidate genes for the marker-assisted selection programs in our and possibly other populations of cattle.

Keywords: cattle, subclinical mastitis, SNP, SCCI

Abbreviations:

BS	Brown Swiss
CM	clinical mastitis
SM	subclinical mastitis
IMI	inframammary infections
SCC	somatic cell count
SCCI	somatic cell count index
QG	quality control
LD	linkage disequilibrium
SNP	single nucleotide polymorphism
QTL	quantitative trait loci

Introduction

Mastitis is a very prevalent and costly disease that affects dairy industry caused by contagious and/or environmental pathogens (Abebe et al., 2016). Generally, mastitis is classified as clinical or subclinical, depending on the severity of inflammation of the mammary gland (De Vliegher et al., 2012). Clinical mastitis (CM) is characterized by visible abnormalities in the milk or the udder. On the other hand, subclinical mastitis (SM) is characterized by the inflammation of the mammary gland without visible abnormalities in the milk and udder or other visible clinical signs (Adkins and Middleton, 2018; Martins et al., 2019). SM is considered a hidden threat to healthy cows in herds that can also lead to clinical mastitis. Development of inframammary infections (IMI) leads to a reduction of milk production, decrease of technological quality milk traits, animal welfare decreasing, and longevity. The total cost of SM varies between farms and depends on the number of cows with increased somatic cell count (SCC) which is one of the best indicator of IMI. Increased somatic cell count is a response of immune system to infection and therefore essential for the interpretation of mammary gland health. A SCC of 150,000 or less indicates an "uninfected cow" (Hawkins, 2019). The SCC in milk increases after calving when colostrum is produced decreases after a few days to normal values and tends to rise again towards the end of lactation. Essentially, low SSC indicates good animal health, as SCC originated from inside of udder, so monitoring of their number is an important part of good herd management. Particular low SCC is sometimes a sign of poor immune system which is as undesirable as high SCC. Immune response is best measured by how quickly it responds to infection and how long it lasts. Recently developed somatic cell count index (SCCI) proposed for the accurate prediction of milk yield losses (Jeretina et al., 2017) also takes into account frequency increases, intensity, and duration of SM. Many risk factors related to mastitis, such as the microorganism, environment or management practices, are known (Rahularaj et al.,

2019), as well as genetics causes for mastitis resistance (Rupp and Boichard, 2003). In addition to phenotype-based selection, where estimated breeding values are used, selection can be also based on genotype. Direct selection based on genotype is looking for genes directly related to mastitis (Sender et al., 2013). For a better understanding of the genetic architecture of this *complex trait*, more informative genetic markers to allow faster and more accurate selection of cattle resistant to mastitis are needed. For that purpose, the objective of our study was to find new potential markers for SM in Brown Swiss cows (BS). The workflow of the study is presented in Supplementary Figure S1.

Material and methods

The data were collected during the regular monthly milk recordings, which are performed by the Slovenian Agency for the Control of Milk Production of Cows. As such, no animals were directly involved in this study, so Animal Care and Use Committee approval was not obtained for this study.

This study was conducted on 1,158 genotyped animals of Brown Swiss breed from the Slovenian cattle population, which are collected in the Central Cattle Database at the Agricultural Institute of Slovenia (Jeretina et al., 1997). Cows were genotyped using SNP microarray ICBF International Dairy and Beef v3 (ICBF, Ireland) for a total of 53,262 SNP markers. Polymorphisms were annotated by SNPChimp v3 (Nicolazzi et al., 2015). The Haploview 4.2 program (Barrett et al., 2005) was used to determine the LD of 130 SNPs, based on r^2 measurements (Figure 1**)**. After matching the phenotype to the genotype data, dataset contains data for 816 cows with three or less known standard lactations. Among them, there were 811 first, 691 second, and 510 third lactations. On average the cows have 2.5 calvings. Quality control (QC) of the genotype data was done using PLINK (Purcell et al., 2007). For the details of QCs in this study, we refer to Anderson *et al.* (Anderson et al., 2010). Briefly, SNPs with call rate ≥90% and minor allele frequency >0.05 were retained for

subsequent analysis. Finally, a total of 18,136 SNPs in 816 animals were retained. All the 18,136 SNP markers used for the final association analyses were distributed over the cattle genome as shown in Figure 1. The association analysis, using one *SNP* at a time, was performed with a linear mixed model using "*lme4"* package in R software (Bates et al., 2015). The single SNP regression analysis was performed using the following statistical model:

$$y_{ijklm} = \mu + P_i + G_j + M_k + b(x_{ijklm} - \bar{x}) + sh_l + e_{ijklm}, \quad (1)$$

where y_{ijklm} is dependent variable SCCI, μ represents SCCI average, P_i parity (i = 1-3), G_j genotype (j = 1-3), M_k stall type (k = 1 (free-stall), 2 (tie-stall)), x_{ijklm} age at first calving, *sh* herd-sire random effect and e_{ijklm} random error.

Genomic overlap analysis was performed comparing the location of SNPs associated with SCCI, protein-coding genes, and somatic cell-related QTL (Figure 2). To analyze genomic overlaps, QTL were obtained from Animal QTL Database, release 38 (http://www.animalgenome.org/cgi-bin/QTLdb/index/). The QTL for somatic cells (somatic cell score, somatic cell count) were selected and overlapped with SNPs associated with SCCI. Genes were obtained from Ensembl Biomart Release 94, while gene nomenclature was based on the HUGO Gene Nomenclature Guidelines (http://www.genenames.org). Throughout the text, genomic positions of SNPs, genes, and QTL were based on the UMD3.1 Bovine Genome Assembly (Zimin et al., 2009). Genomic views (graphical overview of the chromosomal locations) were constructed using the visualization tool Flash GViewer (http://gmod.org/wiki/Flashgviewer/) developed by the GMOD project.

Results and discussion

To identify SNP genetic markers for subclinical mastitis, we carried out an association study on the Slovenian BS cow population with calculated (known) SCCI. For this purpose animals of BS breed from the Slovenian population were genotyped and used in this association study. We performed an association study between SCCI data and 18,136 SNPs. We identified 130 SNPs significantly associated with SCCI (Supplementary Table S1). Linkage disequilibrium (LD) of those 130 SNPs located on the bovine genome is shown in Figure 2. A total of 10 haplotype blocks among 22 SNPs were identified in the SB breed population, with the largest one (block 1), spanning approximately 214,512 kb. Blocks were defined by the four gamete rule with default parameters. The haplotypes are displayed in the left bottom corner of Figure 2, which shows the population frequency of each haplotype in a block, and the connections from one block to the next. In the crossing areas, multiallelic D' values are displayed between blocks (Wang et al., 2002).

The study reveals 51 significant SNPs overlapped with 51 protein-coding genes and five QTL for somatic cells. Most of those SNPs are intergenic, while into 3'UTR mapped three SNPs, three downstream of the gene, 43 into intron, one within coding region and four SNPs mapped upstream of the protein-coding gene (Supplementary Table S2). On chromosome BTA18 three QTL for somatic cell score overlaps with SNP rs41622425 (Figure 2), which is located within intron of a protein-coding gene named "*Pleckstrin Homology Like Domain Family B Member 3"* (*PHLDB3*). Additionally, at the same chromosome rs110437995 overlaps with QTL for somatic cell count. Moreover, three of overlapped genes *GAS6, PKD2,* and *NCAM1* were previously associated with inflammation in human and mouse. Moreover, *Gas6* negatively regulates the *Staphylococcus aureus*- induced inflammatory response in the mouse mammary gland (Zahoor et al., 2020). *Staphylococcus aureus* is the most prevalent contagious mastitis pathogen. Gas6 activates the TAM receptor kinase

pathway, which is related to the inhibition of TLR2- and TLR6- mediated inflammatory pathways. Gas6 absence alone was found to be involved in the downregulation of TAM receptor- mediated anti- inflammatory effects. Gene *NCAM1*, also known as CD56 is a member of the immunoglobulin superfamily and is the phenotypic marker of natural killer cells (Van Acker et al., 2017). It has been reported in patients with various infectious, autoimmune, or malignant diseases. Therefore, it would be reasonable to examine more closely its role in SM as well.

Conclusion

In conclusion, we found significant effects of bovine genomics variations on SCCI in the Slovenian BS population. Some of the polymorphisms included in our study with an effect on SCCI map in protein-coding genes. However, those genes as well as SNPs associated with mastitis related traits from this study can serve as strong candidates for the marker-assisted selection programs in our and possibly other populations of cattle.

Acknowledgements

This work was supported by the Slovenian Research Agency program P4-0133. We thank our co-workers at the Agricultural Institute of Slovenija, MKGP, KGZS, and breeders.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Literature Cited

Abebe, R., H. Hatiya, M. Abera, B. Megersa, and K. Asmare. 2016. Bovine mastitis: prevalence, risk factors and isolation of Staphylococcus aureus in dairy herds at Hawassa milk shed, South Ethiopia. BMC Veterinary Research. 12:270. doi:10.1186/s12917-016-0905-3.

Adkins, P. R. F., and J. R. Middleton. 2018. Methods for Diagnosing Mastitis. Veterinary Clinics of North America: Food Animal Practice. 34:479–491. doi:10.1016/j.cvfa.2018.07.003.

Anderson, C. A., F. H. Pettersson, G. M. Clarke, L. R. Cardon, A. P. Morris, and K. T. Zondervan. 2010. Data quality control in genetic case-control association studies. Nature Protocols. 5:1564–1573. doi:10.1038/nprot.2010.116.

Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 21:263–265. doi:10.1093/bioinformatics/bth457.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 67:1–48. doi:doi:10.18637/jss.v067.i01.

De Vliegher, S., L. K. Fox, S. Piepers, S. McDougall, and H. W. Barkema. 2012. Invited review: Mastitis in dairy heifers: Nature of the disease, potential impact, prevention, and control. Journal of Dairy Science. 95:1025–1040. doi:10.3168/jds.2010-4074.

Hawkins, D. 2019. Use of different somatic cell count cut-points to define intramammary infection at drying off in dairy cows from a herd with a high somatic cell count. N Z Vet J. 67:203–209. doi:10.1080/00480169.2019.1608872.

Jeretina, J., B. Ivanovič, P. Podgoršek, T. Perpar, B. Logar, M. Sadar, J. Jenko, J. Glad, A. Božič, A. Žabjek, D. Babnik, and J. Verbič. 1997. Centralno podatkovna zbirka Govedo. Kmetijski inštitut Slovenije, Ljubljana, Slovenija, http://www.govedo.si. Available from: http://www.govedo.si

Jeretina, J., D. Škorjanc, and D. Babnik. 2017. A new somatic cell count index to more accurately predict milk yield losses. Archives Animal Breeding. 60:373–383. doi:10.5194/aab-60-373-2017.

Martins, S. A. M., V. C. Martins, F. A. Cardoso, J. Germano, M. Rodrigues, C. Duarte, R. Bexiga, S. Cardoso, and P. P. Freitas. 2019. Biosensors for On-Farm Diagnosis of Mastitis. Front. Bioeng. Biotechnol. 7:186. doi:10.3389/fbioe.2019.00186.

Nicolazzi, E. L., A. Caprera, N. Nazzicari, P. Cozzi, F. Strozzi, C. Lawley, A. Pirani, C. Soans, F. Brew, H. Jorjani, G. Evans, B. Simpson, G. Tosser-Klopp, R. Brauning, J. L. Williams, and A. Stella. 2015. SNPchiMp v.3: integrating and standardizing single nucleotide polymorphism data for livestock species. BMC Genomics. 16. doi:10.1186/s12864-015-1497-1. Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-015-1497-1

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet. 81:559–575.

Rahularaj, R., R. M. C. Deshapriya, and R. M. S. B. K. Ranasinghe. 2019. Influence of bovine sub-clinical mastitis and associated risk factors on calving interval in a population of crossbred lactating cows in Sri Lanka. Tropical Animal Health and Production. doi:10.1007/s11250-019-01957-4. Available from: http://link.springer.com/10.1007/s11250-019-01957-4

Rupp, R., and D. Boichard. 2003. Genetics of resistance to mastitis in dairy cattle. Veterinary research. 34:671–88. doi:10.1051/vetres:2003020.

Sender, G., A. Korwin-Kossakowska, A. Pawlik, K. G. A. Hameed, and J. Oprządek. 2013. Genetic Basis of Mastitis Resistance in Dairy Cattle – A Review / Podstawy Genetyczne Odporności Krów Mlecznych Na Zapalenie Wymienia – Artykuł Przeglądowy. Annals of Animal Science. 13:663–673. doi:10.2478/aoas-2013-0043.

Van Acker, H. H., A. Capsomidis, E. L. Smits, and V. F. Van Tendeloo. 2017. CD56 in the Immune System: More Than a Marker for Cytotoxicity? Front. Immunol. 8:892. doi:10.3389/fimmu.2017.00892.

Wang, N., J. M. Akey, K. Zhang, R. Chakraborty, and L. Jin. 2002. Distribution of Recombination Crossovers and the Origin of Haplotype Blocks: The Interplay of Population History, Recombination, and Mutation. The American Journal of Human Genetics. 71:1227–1234. doi:10.1086/344398.

Zahoor, A., Y. Yang, C. Yang, M. Akhtar, Y. Guo, A. Shaukat, M. Guo, and G. Deng. 2020. Gas6 negatively regulates the *Staphylococcus aureus* -induced inflammatory response via TLR signaling in the mouse mammary gland. J Cell Physiol. jcp.29604. doi:10.1002/jcp.29604.

Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. P. Van Tassell, T. S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian, J. A. Yorke, and S. L. Salzberg. 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol. 10:R42. doi:10.1186/gb-2009-10-4-r42.

FIGURE LEGENDS

Figure 1: Genomic organization of the bovine genome, linkage disequilibrium (r^2) plot for the 130 polymorphisms. The rectangle at the bottom left displays the haplotype blocks with frequencies shown next to each haplotype.

Figure 2: Bovine chromosomes with 130 associated SNPs, overlapping protein-coding genes and QTL for somatic cell. A: Enlargement of the chromosome 18 showing an overlap between four QTL for somatic cells, gene *PHLDB3*, and two SNPs associated with somatic cell count index (SCCI)

Figure 1

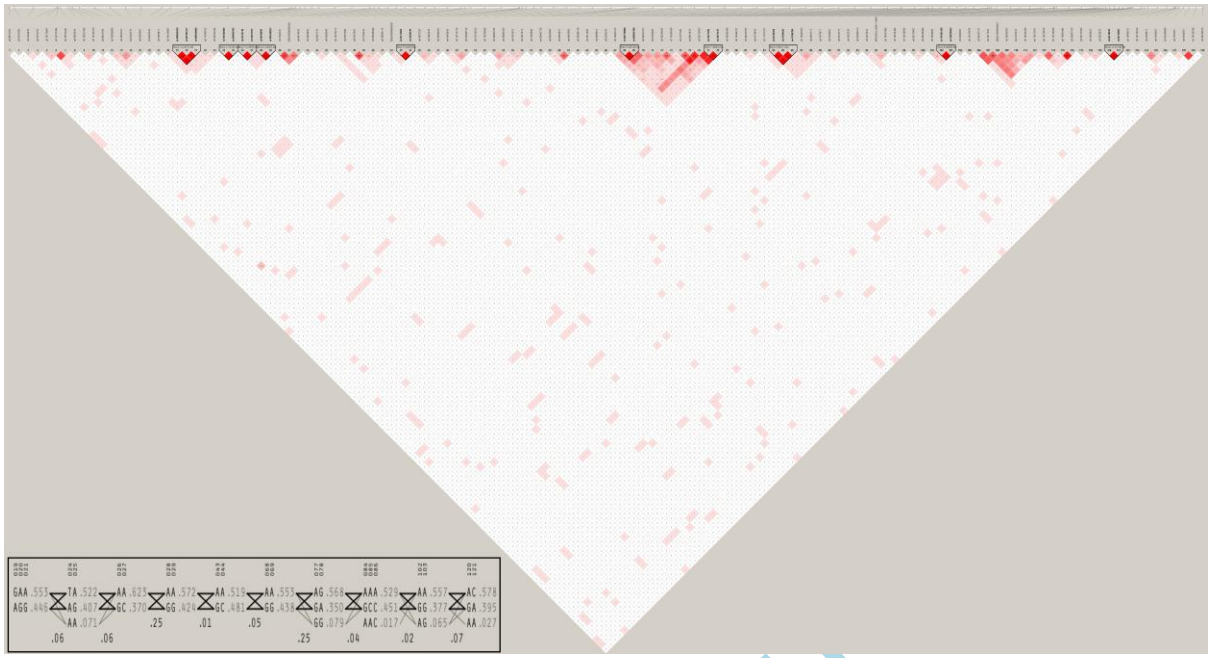

Figure 2