Summary

Increase in milk production as a result of intense genetic selection in dairy cattle has been accompanied by increase in the incidence of several reproductive and metabolic related disorders, including ketosis. Several investigations have studied candidate genes and biological pathways that characterize metabolic status of dairy cattle and ketosis early in lactation, however, little is known about genome-wide regions and key regulatory genes underlying this trait. In this study, we report a genome-wide association analysis on milk BHB concentrations (as an indicator of ketosis) predicted by mid-infrared (MIR) spectroscopy in dairy cattle using a single SNP regression mixed linear model.

Our study detected several significant regions associated with MIR predicted milk BHB concentrations on chromosomes 6 (first lactation) and chromosomes 14 and 20 (later lactations). One highly significant SNP on chromosome 14 was located within *DGAT1* gene, which is known to have significant effects on milk fat/protein yield and other production traits. The significant regions on chromosomes 6 and 20 were not reported to be linked to metabolic associated disorders or ketosis in previous investigations. Enrichment analysis of the list of candidate genes within the identified regions for milk BHB concentrations has yielded molecular functions and biological process that may contribute to inflammatory response and lipid metabolism in dairy cattle. The result of this study can be used for further analysis to identify causal variations and key regulatory genes that affect clinical/ subclinical ketosis.

Key words: milk BHB concentrations, MIR spectroscopy, clinical/subclinical ketosis, genome-wide association, dairy cattle

Genome-wide Association Analysis for β-hydroxybutyrate Concentration in Milk Using Mid-Infrared Spectroscopy in North American Holstein Cattle

S. Nayeri[1], V. Kroezen[1], M. Sargolzaei[12], A. Fleming[1], F. Schenkel[1], C. Baes[1], A. Cánovas[1], J. Squires[1], and F. Miglior[13]
snayeri@uoguelph.ca *(Corresponding Author)*
[1]Centre for Genetic Improvement of Livestock, Department of Animal Bioscience, University of Guelph, ON, Canada
[2]The Semex Alliance, Guelph, ON, Canada
[3]Canadian Dairy Network, Guelph, ON, Canada

Introduction

Increase in milk production as a result of intense genetic selection in dairy cattle has been accompanied by higher incidence of reproductive health issues and production-related diseases including metritis, ketosis and fatty liver (Ha et al., 2015). This can be due to the metabolic changes and challenges in high-producing cows early in lactation and the failure of animals to maintain their internal homeostatic and homeorhetic regulation (Nayeri and Stothard, 2016). Amongst the metabolic diseases, ketosis has been shown to cause a high morbidity and

substantial financial losses to the dairy farmers (Biswal et al., 2016). Symptoms associated with ketosis in dairy cattle are poorly diagnosed in the farm, however, the amount of ketone bodies (acetone, β-hydroxybutric acid (BHB), and acetoacetate) present in milk has been shown to be good indicators of clinical and subclinical ketosis (Tetens et al., 2015; Biswal et al., 2016).

Several studies have investigated candidate genes and biological pathways characterizing the metabolic adaptability of dairy cow pre and post-partum (Loor et al., 2007; Ha et al., 2015). However, significant genome-wide regions harboring candidate genes have not been reported in previous studies. The main goal of this study was to identify genome-wide regions associated with MIR predicted BHB concentrations in milk associated with ketosis in the North American Holstein dairy cattle. Additionally, an *In-Silico* enrichment analysis was performed to identify biologically significant genes and pathways associated with ketosis and subclinical ketosis.

Materials and methods

Animals and data

The Canadian Dairy Network (CDN, Guelph, Ontario, Canada) provided available pedigree, genotype and genetic evaluations of cows and proven bulls for MIR predicted milk BHB records in first lactation (SCK1) and later lactations (SCK2) as separate traits. Milk BHB measurements have been collected as a part of routine phenotyping method via mid-infrared (MIR) spectroscopy by Canadian DHI since December 2016. The accuracy of measurements assessed using receiver operating characteristic curves, as explained by Denis-Robichaud et al. (2014). For the accuracy of milk BHB concentrations the accuracy of optimal threshold values of sensitivity (Se > 80%) and specificity (Sp > 95%) values were considered (Denis-Robichaud et al., 2014). The optimal threshold is the cut-off value providing the maximal sum of sensivity and specificity of the milk BHB measurements (Denis-Robichaud et al., 2014). The genetic evaluations for MIR predicted milk BHB were then de-regressed (Garrick et al., 2009) to be used in the genome-wide association study (GWAS). Animals with a reliability of the de-regressed EBV lower than 10% were removed from the analysis.

Genotypes of 24,657 Canadian Holstein bulls and cows that had been previously genotyped with the BovineSNP50K (50K, 41,097 SNPs) BeadChip (Illumina, San Diego, CA) were provided by CDN. The SNPs in the 50K panel have passed standard quality control measures used by CDN. These genotypes were then imputed to the high density (HD, 311,725 SNPs) genotypes (with a reference population of 2,507 animals) using FImpute V2.2 software (Sargolzaei et al., 2014). Quality control was performed on the imputed HD genotyping data using the snp1101 software (Sargolzaei, 2014) and SNPs with low call rate < 0.9, low minor allele frequency (MAF) <0.01 and excess of heterozygosity > 0.15 were excluded. The remaining 298,210 SNPs after quality control were used for further GWAS.

Genome wide association study (GWAS)

The association analysis was performed using a single SNP regression mixed linear model implemented in the snp1101 software. In order to account for multiple testing, genome-wise false discovery rate (FDR) of 5% and 1% were used to identify significant and very significant

associations, respectively.

In-Silico enrichment analysis

Significant SNPs at the level of 5% FDR were mapped to the corresponding genes in Ensemble database (Ensembl 90, *Bos taurus* UMD3.1, http://useast.ensembl.org/Bos_taurus/Info/Index) using the getBM() function in R-biomaRt package (https://www.bioconductor.org/) (Durinck et al., 2009). Genes within $\pm$100kb from the identified significant SNPs were selected for further functional analysis.

Functional analysis was performed to identify the biological pathways and gene-networks associated with the list of positional candidate genes using the Ingenuity Pathway Analysis software (IPA; Ingenuity System Inc, USA). Additionally, PANTHER classification System (http://pantherdb.org/) was used to test for the GO term enrichment analysis and to gain a better insight into the functional and biological relatedness of the genes for the trait in this study.

Results and discussion

Association analysis identified strong associations for both milk BHB in first and later lactations. In total 71 and 370 SNPs were detected as significant (5% FDR genome-wise) for SCK1 and SCK2, respectively. The strongest association for SCK1 was on chromosome 6. This region was associated with both SCK1 and SCK2. Two highly significant regions for SCK2 on chromosomes 14 and 20 were also identified. The region identified on chromosome 14, contained one highly significant SNP within the gene *DGAT1*, which has been reported previously to have a major effect on several milk production associated traits, including milk fat and protein yield (Grisart et al., 2004). Furthermore, in a whole-genome association study for energy balance and fat/protein ratio in German Holstein cattle, significant SNPs within the *DGAT1* gene were identified to be associated with fat/protein ratio (Tetens et al., 2013). Tetens et al. (2013) described that fat/protein ratio is an alternative measure of the energy balance early in lactation in dairy cattle, suggesting that the *DGAT1* gene might have an influence on the metabolic status of cows. It has been shown that ketosis is mostly observed during the first 65 days in milk, which is the critical stage of the metabolic load in dairy cattle (Dohoo and Martin, 1984). Presence of QTLs on chromosomes 6 (at 87 Mb) near the casein gene cluster affecting milk production traits has been reported in several breeds (Freyer et al., 2003; Nayeri et al., 2016). This is the first time, however, that this region on chromosome 6 (at 88-98 Mb) has been reported to be associated with milk BHB concentrations in dairy cattle. The region on chromosome 20 (55-63 Mb) has been identified as a novel region affecting milk BHB concentrations.

Thirty-six unique genes corresponding to the identified SNPs were identified for SCK1 and SCK2. Biological pathway analysis and GO term enrichment analysis for the list of positional candidate gens resulted in overrepresentation nine biological processes some with relevant functions to the physiology of the trait including metabolic process (GO:0008152) and response to stimulus (GO:0050896). The top two most significant gene-networks in IPA analysis for SCK1

were associated with infectious diseases and cell to cell signaling and interaction, whereas for SCK2, the top two most significant gene-networks were associated with lipid metabolism and developmental disorders. The result of gene-network analysis suggests the potential association of the identified positional candidate genes with immunity and inflammatory response of the cow to metabolic associated disease.

Conclusions

Genome-wide association analysis carried out in this study detected, for the first time, several regions associated with MIR predicted milk BHB in first and later lactations in Canadian Holstein cattle. Functional analysis of the GWAS results could also highlight GO terms and biological pathways associated with metabolic process and immune physiology in dairy cattle. The novel regions identified in this study can be used for further functional analysis to identify key genes and pathways that explain variation for clinical/subclinical ketosis and other associated metabolic diseases.

Acknowledgements

References

Biswal, S., D.C. Nayak, and K.K. Sardar. 2016. Prevalence of ketosis in dairy cows in milk shed areas of Odisha state , 9:1242–1247. doi:10.14202/vetworld.2016.1242-1247.

Denis-Robichaud, J., J. Dubuc, D. Lefebvre, and L. DesCôteaux. 2014. Accuracy of milk ketone bodies from flow-injection analysis for the diagnosis of hyperketonemia in dairy cows. J. Dairy Sci. 97:3364–3370. doi:10.3168/jds.2013-6744.

Dohoo, I.A.N.R., and S.W. Martin. 1984. Disease, production and culling in Holstein-Friesian cows. IV. Effects of disease on production 2:755–770.

Durinck, S., P.T. Spellman, E. Birney, and W. Huber. 2009. Mapping identfiers for integration of genomic datasets with the R/Bioconductor package biomaRt 4:1184–1191. doi:10.1038/nprot.2009.97.Mapping.

Freyer, G., P. Sørensen, C. Kühn, R. Weikard, and I. Hoeschele. 2003. Search for pleiotropic QTL on chromosome BTA6 affecting yield traits of milk production.. J. Dairy Sci. 86:999–1008. doi:10.3168/jds.S0022-0302(03)73683-2.

Garrick, D.J., J.F. Taylor, and R.L. Fernando. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genet. Sel. Evol. 1:41:55. doi:10.1186/1297-9686-41-55.

Grisart, B., F. Farnir, L. Karim, N. Cambisano, J. Kim, A. Kvasz, M. Mni, P. Simon, J.-M. Frere, W. Coppieters, and M. Georges. 2004. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proc. Natl. Acad. Sci. U. S. A. 101:2398–403.

Ha, N.-T., J.J. Gross, A. van Dorland, J. Tetens, G. Thaller, M. Schlather, R. Bruckmaier, and H. Simianer. 2015. Gene-Based Mapping and Pathway Analysis of Metabolic Traits in Dairy Cows. PLoS One 10:e0122325. doi:10.1371/journal.pone.0122325.

Loor, J.J., R.E. Everts, M. Bionaz, H.M. Dann, D.E. Morin, R. Oliveira, S.L. Rodriguez-Zas, J.K. Drackley, and H. a Lewin. 2007. Nutrition-induced ketosis alters metabolic and signaling gene networks in liver of periparturient dairy cows.. Physiol. Genomics 32:105–16. doi:10.1152/physiolgenomics.00188.2007.

Nayeri, S., M. Sargolzaei, M.K. Abo-Ismail, N. May, S.P. Miller, F. Schenkel, S.S. Moore, and P. Stothard. 2016. Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. BMC Genet. 17:75. doi:10.1186/s12863-016-0386-1.

Nayeri, S., and P. Stothard. 2016. Tissues , Metabolic Pathways and Genes of Key Importance in Lactating Dairy Cattle. Springer Sci. Rev. 4:49–77. doi:10.1007/s40362-016-0040-3.

Sargolzaei, M. 2014. SNP1101 User ' s Guide. Version 1.0 0.

Sargolzaei, M., J.P. Chesnais, and F.S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives.. BMC Genomics 15:478. doi:10.1186/1471-2164-15-478.

Tetens, J., C. Heuer, I. Heyer, M.S. Klein, W. Gronwald, W. Junge, P.J. Oefner, G. Thaller, N. Krattenmacher, O. Pj, G. Thaller, and K.N. Polymorphisms. 2015. Polymorphisms within the APOBR gene are highly associated with milk levels of prognostic ketosis biomarkers in dairy cows 129–137. doi:10.1152/physiolgenomics.00126.2014.

Tetens, J., T. Seidenspinner, N. Buttchereit, and G. Thaller. 2013. Whole-genome association study for energy balance and fat/protein ratio in German Holstein bull dams. Anim. Genet. 44:1–8. doi:10.1111/j.1365-2052.2012.02357.x.