Genome-wide association study and functional analysis of infectious and horn type hoof lesions in Canadian Holstein cattle

F. Malchiodi1, L. F. Brito1, F. S. Schenkel1, A. M. Christen2, D. F. Kelton3 & F. Miglior1,4

1CGIL, Department of Animal Biosciences, University of Guelph, Guelph, Ontario, Canada
fmalchio@uoguelph.ca (Corresponding author)
2Valacta, Sainte-Anne-De-Bellevue, Quebec, Canada
3Department of Population Medicine, OVC, University of Guelph, Guelph, Ontario, Canada
4Canadian Dairy Network (CDN), Guelph, Ontario, Canada

Summary

The objectives of this study were to perform a genome-wide association study and a functional analysis to unravel genomic regions associated with hoof lesions and potential key regulator genes affecting the resistance to infectious and horn lesions in Canadian Holstein cattle. A total of 249,709 observations from 105,450 animals were recorded by 51 hoof trimmers during the trimming activity in 1,080 Canadian herds from 2009 to 2016. For horn lesions, a total of 51 20-adjacent-SNP moving windows explained more than 0.30% of the genetic variance and these most important peaks were observed on chromosomes BTA1, BTA5, BTA6, BTA9, BTA15, BTA16, BTA20, BTA21, BTA25, and BTA28. For infectious lesions, the six most important windows (>0.30%) were observed on chromosomes BTA2, BTA3, BTA10, and BTA14. The highest peaks were observed in BTA10 (0.47%) and BTA5 (1.23%) for infectious and horn lesions, respectively.

Keywords: dairy cattle, hoof health, weighted single step GBLUP, ssGWAS

Introduction

Reducing the incidence of hoof lesions is a key goal of dairy farmers and it can be achieved both by improving management practices, and through genetic selection. Hoof lesions can be classified as infectious (IL) or horn type lesions (HL) and weak genetic correlations have been reported between IL and HL (van der Waaij et al., 2005; van der Speck et al., 2013; Malchiodi et al., 2017). Previous studies suggested that lesions should be considered individually according to their etiology and pathogenesis. In order to achieve faster genetic progress for these traits, it is key to understand the genetic architecture of these groups of traits and the genomic relationship among them via genome-wide association studies (GWAS) and functional analyses. Unraveling genomic regions associated with resistance to hoof lesions and defining candidate genes harboring important regions will allow a more comprehensive understanding of the mechanisms involved in these different groups of lesions and potentially contribute to a more accurate genetic selection for these traits. Therefore, the main objectives of this study were: 1) to perform a GWAS aiming to identify genomic regions associated with hoof health traits; and 2) to perform functional analysis to identify potential key regulator genes affecting the resistance to IL and HL in Canadian dairy cattle using a large dataset.

Materials and methods

Hoof lesions were recorded by 51 hoof trimmers during the routinely trimmed
activity in 1,080 herds located in Alberta, British Columbia, New Brunswick, Ontario, and Quebec from 2009 to 2016. The hoof trimmers were trained to use a rugged touch-screen computerized lesion recording system (Hoof Supervisor, KS, Dresser, WI). Lesions were classified according to their etiology and pathogenesis into IL or partly IL (digital and interdigital dermatitis, foot rot, and heel erosion), and HL (sole and toe ulcer, sole haemorrhage, and white line disease), as previously reported by Chapinal et al. (2013). Hoof lesions were considered as binary variables (0 or 1), where 1 was assigned to the presence of a lesion. The final phenotypic dataset contained 249,709 observations from 105,450 animals, with a pedigree file of 331,587 animals.

Genetic parameters and breeding values were estimated with a univariate animal model using the average information-restricted maximum likelihood (AI-REML) procedure, in the derivative-free approach to multivariate analysis (DMU) package (Madsen and Jensen, 2008). The model included the fixed effect of herd-date of hoof trimming, hoof trimmer, parity at trimming, stage of lactation at trimming, and the random additive genetic animal and permanent environmental effects. The estimated breeding values (EBV) for each trait were de-regressed based on the VanRaden et al. (2009) method. Only animals with EBV reliability greater than 0.10 and genotyped (13,657 for HL and 13,834 for IL) were kept for further analyses. Animals were genotyped either with 50K SNP panel or a low-density panel (≥6K SNP) and accurately imputed to 50K. The genotyping quality control was performed using the PREGSF90 software (Aguilar et al., 2014) and filtered out markers located on non-autosomal regions, that mapped at the same position, had $P$-value for Hard-Weinberg equilibrium test less than $10^{-6}$, call rate lower than 0.90 and minor allele frequency lower than 0.01. A total of 44,369 SNPs remained for further analyses.

Direct genomic values (DGV) for each animal and trait were predicted based on the GBLUP method (VanRaden, 2008), using de-regressed EBV as pseudo-phenotypes. The SNP effects were obtained back solving the DGV for each trait using the POSTGFSF90 software (Aguilar et al., 2014), using the formula: $D = Z(Z’Z)^{-1}Z’$, where: $D$ is a diagonal matrix with weights for SNPs, $Z$ is an incidence matrix of genotypes for each locus, and $Z$ is the vector of DGV. The $D$ matrix were iteratively recomputed following an algorithm proposed by Wang et al. (2012) and the algorithm was run over three iterations.

In order to identify important genomic regions related to HL and IL, 20-adjacent-SNP moving windows (SNP-by-SNP) that explained more than 0.30% of genetic variance (calculated by summing the variance of the 20 adjacent SNPs) were used as the criteria to identify genomic regions potentially associated with HL and IL. Genes located within the most significant windows were further investigated to identify candidate genes. Functional analyses were performed using the Ensembl - Variant Effect Predictor tool (www.ensembl.org/Tools/VEP), The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8., BioMart (www.ensembl.org/biomart), and the National Center for Biotechnology Information database (NCBI, www.ncbi.nlm.nih.gov/gene/).

**Results and discussion**

The prevalence of IL and HL lesions was 21.6% and 18.3%, respectively, and the heritability estimates (SE) for IL and HL were 0.087 (0.005) and 0.059 (0.004), respectively. The proportion of genetic variance explained by each genomic region is shown in Figure 1. Van der Speck et al. (2016) performed a GWAS for claw disorders. The authors did not report significant SNP associated with IL, which were defined as a combination of digital and interdigital dermatitis, and heel erosion, but they identified markers mainly located on BTA8 associated with sole ulcer. In the present study, a total of
six 20-adjacent-SNP windows explained more than 0.30% of the total genetic variance for IL and these most important peaks were observed on chromosomes BTA2, BTA3, BTA10 and BTA14. For HL, a total of 51 20-adjacent-SNP windows explained more than 0.30% of the total genetic variance and the most important peaks were observed on chromosomes BTA1, BTA5, BTA6, BTA9, BTA15, BTA16, BTA20, BTA21, BTA25, and BTA28. The highest peaks were observed in BTA10 (0.47%) and BTA5 (1.23%) for IL and HL, respectively.

The most relevant genes located within the most important peak regions (and up to 0.5 Mb from these windows) for HL and IL are reported in Table 1. Based on the functional analysis, some candidate genes with functions related to infectious lesions are PEL12, ZAK, CDCA7, CD1E, and AHSAS1, which are associated with immune system functions.

Table 1. Candidate genes for infectious and horn hoof lesions in Canadian Holstein cattle.

<table>
<thead>
<tr>
<th>Trait</th>
<th>BTA</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infectious lesions</td>
<td>2</td>
<td>CDCA7, EPC2, KIF5C, LYPD6B, RAPGEF4, UBL5, ZAK</td>
</tr>
<tr>
<td>Infectious lesions</td>
<td>3</td>
<td>CD1E, CD5L, FCRL1, FCRL2, FCRL3, KIRREL1, OR10K2, OR10R2, OR10T2</td>
</tr>
<tr>
<td>Infectious lesions</td>
<td>10</td>
<td>ADCK1, AHSAS1, AP5M1, EXOC5, GSTZ1, IRF2BPL, KTN1, NGB, NOXRED1, OTX2, PEL12, POMT2, SAMD15, SNORA72, SNW1, SPTLC2, TMED8, TMEM260, TMEM63C, VIPAS39, ZDHHC22</td>
</tr>
<tr>
<td>Infectious lesions</td>
<td>14</td>
<td>COL22A1</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>1</td>
<td>PCCB, PPP2R3A, PRMT2</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>2</td>
<td>COP57B, PDE6D</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>5</td>
<td>ANKRD54, ANO2, APAF1, ARNTL2, BAIAP2L2, C1QL4, C12orf60, CACNB3, CBY1, CDPF1, CERK, CHST11, DDX23, DMC1, DNAJC22, FMNL3, GALR3, GCAT, GNS, GRAMD4, GRIN2B, GTPBP1, GTSE1, GUCY2C, INTS13, IPO8, JOSD1, KDELR3, KLRK1, KMT2D, LEMD3, LRP6, MAFF, MAGOHB, MANSCI, MCRS1, MSRB3, NCKAP5L, NT5DC3, PIK3C2G, PLBD1, PLCZ1, POLR2F, PPARA, PRKAG1, PTPRO, RERG, RHNO1, RND1, SMC03, SOX10, STAB2, STK38L, STYK1, TBC1D30, TEAD4, TIMP3, TMIBM6, TSPAN9, TUBA8, TULP3, TXNRD1, VWF, WNT7B</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>6</td>
<td>BDH2, CENPE, GSTCD, KCNIP4, MANBA, NFkB1, NPNT, PaCRGL, SCARNA17, SCARNA18, SLC9B1, SLC9B2, SLIT2, TBCK, TKTL2</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>9</td>
<td>ADGRG6, AHI1, MAP7, NMBR, PDE7B, UTRN, VTA1</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>16</td>
<td>HMGB1, KCTD3, USH2A</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>20</td>
<td>CDH18</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>21</td>
<td>BDKRB1, BDKRB2, SERPINA1, SERPINA4, SERPINA5, TTC6</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>25</td>
<td>AUTS2, CUX1, ORAI1, PLD3, PRKRP1, SH2B2, ZNHTI</td>
</tr>
<tr>
<td>Horn lesions</td>
<td>28</td>
<td>ARID5B, C10orf105, CDH23, GALNT2, KCN1, PGBD5, PSA1, RTKN2, UNCS3B, ZNF365</td>
</tr>
</tbody>
</table>

Figure 1. Manhattan plot showing markers and chromosomal regions associated with infectious (a) and horn (b) lesions.
processes, embryonic digit morphogenesis, regulation of cell proliferation and apoptotic processes, antigen processing and presentation, and response to stress, respectively. For horn type lesions, some candidate genes identified are PLOD3, RERG, BDH2, GALNT2, HMGB1, BDKRB1, PPARA, SH2B2, CHST11, TULP3, and STYK1, which are associated with epidermis morphogenesis, collagen fibril organization, collagen metabolic process, regulation of cell growth, epithelial cell differentiation, immunoglobulin biosynthetic process, inflammatory response, activation of innate immune response, epidermis development, regulation of immune response, cartilage development and embryonic digit morphogenesis, and innate immune responses. DAVID functional clusters identified important terms such as calcium and epidermal growth factor-like domain for HL, and immunoglobulin, integral component of membrane and cell membrane terms for IL. The large number of candidate genes identified in this study is probably due to the grouping of various traits in two main categories and large relevant genomic regions due to the use of a medium-density SNP panel.

Conclusions

Important genomic regions associated with infectious and horn lesions were identified and a list of functional candidate genes within or next to these regions was created. The next step is to perform analyses using whole-genome sequence data, alternative statistical methodologies, and individual hoof health traits to ascertain shorter relevant genomic regions and most important candidate genes.

Acknowledgments

The authors are grateful to the hoof trimmers who participated in this study. This research was supported in main part by Agriculture and Agri-Food Canada and by additional contributions from Dairy Farmers of Canada, the Canadian Dairy Network and the Canadian Dairy Commission under the Agri-Science Clusters Initiative and Alberta Milk (Edmonton, AB, Canada). Partial funding was provided by Ontario Genomics.

List of References


