Genome-Wide Association and Functional Annotation of Positional Candidate Genes for Immune Response in Canadian Holstein Cattle

B. Mallard[1,2], S.L. Cartwright[1], S. Nayeri[2], M. Emam[1,2], & M. Sargolzaei[2,3]

[1] Ontario Veterinary College, University of Guelph, Department of Pathobiology, 50 Stone Rd E. N1G 2W1, Guelph, Ontario, Canada
bmallard@ovc.uoguelph.ca (Corresponding Author)
[2] Centre of Genetics and Improvement of Livestock, University of Guelph, Department of Animal Biosciences, 50 Stone Rd E N1G 2W1, Guelph Ontario, Canada
[3]Semex Alliance, 5653 Hwy 6, N1H 6J2, Guelph, Ontario, Canada

Summary

Genetic selection for health traits in the Canadian dairy industry began well over a decade ago but there are still improvements to be made to disease resistance. Previous work by our group has shown animals can be classified based on their immunocompetence and that high immune responders (HIR) have reduced disease incidence compared to their herd mates. Therefore, the objectives of this study were to perform a genome-wide association study (GWAS) on 4400 Holsteins, using a high density panel (HD), to identify genomic regions associated with immune response traits and subsequently perform a functional analysis. In this study genomic estimates of heritability were 0.37 and 0.16 and pedigree estimates were 0.45 and 0.18 for antibody-mediated immune response (AMIR) and cell-mediated immune response (CMIR), respectively. Results also revealed significant SNPs associated with AMIR and CMIR. Functional analysis indicated that candidate genes identified within or close (±100kb) to these markers regulate unique functions that are associated with AMIR and CMIR, respectively. Therefore, genomic selection to improve immune response traits is a viable option for reducing disease incidence in dairy cattle.

Keywords: genome wide association study, dairy cattle, immune response

Introduction

Selection in the Canadian Holstein breed has mainly been based on the lifetime profitability index (LPI). For many years the Canadian LPI only included production and confirmation traits, with health traits not being added until 2001. Currently health and fertility accounts for 19% of the LPI, with the main focus of the health traits being on resistance to mastitis. Although health traits have been included in the index for a number of years now, to date we have only seen a reduction in actual somatic cell count with little to no reduction in mastitis incidence over the years (NAHMS, 2007; Koeck et al., 2012). The recent addition of genomics to breeding programs has led to an increase in accuracy of selecting for certain traits in the LPI and therefore has improved the genetic gain in these traits (CDN, 2017). This, however, has not necessarily been the case for all health traits as disease incidence continues to increase or remain at current levels (NAHMS, 2007; Koeck et al., 2012). Therefore there is still room for further improvement to dairy health.

Numerous studies have shown animals classified as HIR using the University of Guelph patented test system have less disease compared to those identified as average and low (Mallard et al., 2015). Specifically, dairy cattle identified as high responders have significantly lower disease incidence for a number of different economically important diseases (Thompson-Crispi *et al*., 2012b). This is also the case for Immunity+^{TM} daughters (Larmer and Mallard, 2016). Additionally, the heritability for immune response is moderately high indicating genetic gains can be made (Thompson-Crispi *et al*., 2012a). Therefore breeding for high immune response is an alternative solution for reducing disease in dairy cattle versus selecting for resistance to a single disease. Since greater genetic gain and increased reliability is achieved with genomic selection compared to traditional selection based solely on estimated breeding values (EBVs) we evaluated if single nucleotide polymorphisms (SNPs) could be found that associated with AMIR and CMIR traits using GWAS on over 4400 IR phenotyped Holsteins. Additionally a functional enrichment analysis was performed on the GWAS result to investigate relevant biological pathways and gene-networks enriched for these immune response traits.

Material and Methods

Animals and Immunization Protocol

Holstein females (n= 2571) from 60 different commercial farms across Canada and Holstein males (n= 1907) from 4 different Semex facilities in Canada and Hungary were immune response phenotyped using a patented test protocol previously described (Heriazon *et al*., 2009). Briefly, on the initial day of testing a blood sample was taken from the tail vein and cattle were immunized intra-muscularly with type 1 and type 2 antigens. Subsequently AMIR and CMIR responses were quantified to these test antigens. As a source of DNA, hair follicle samples were collected and sent to Zoetis (Kirkland, Quebec) for genotyping using the Bovine Illumina 50K chip.

Genome Wide Association Study

After performing genomic quality control, the 50k genotypes on autosomes were imputed to HD (i.e. 777k panel) using another 2,998 reference animals by FImpute (Sargolzaei et al., 2014). GWAS was carried out on 604k SNPs using single SNP mixed linear model implemented in snp1101 software (Sargolzaei, 2014). Statistical models for AMIR and CMIR were as follows:

$AMIR = \mu + \beta_1 \text{ Age} + \beta_2 \text{ LD0} + \text{Gender} + \text{Season} + \text{Facility} + \text{Preg} + a + \beta_3 \text{ SNP} + e$

$CMIR = \mu + \beta_1 \text{ Age} + \beta_2 \text{ Control} + \text{Gender} + \text{Season} + \text{Facility} + \text{Preg} + a + \beta_3 \text{ SNP} + e$

where, μ = the population mean for both traits, Age = age at day 0 phenotyping in months, LD0 = log10 of day 0 values, Control = log10 (control at 24 hours/control at 0 hours), Gender = male or female, Season = season at day 0 phenotyping (spring, summer, winter, fall), Facility = herd or facility animals came from, Preg = pregnancy status of males was assigned a status of 0 (0 non-pregnant, 1 early, 2 middle, 3 late), a = random additive polygenic effect and B3 is allele substitution effect for the SNP.

Functional Analysis

All of the significant SNPs identified for AMIR and CMIR, at the level of 5% genome-wise false discovery rate (FDR), were assigned to the corresponding genes in Ensemble database (Ensembl 90, *Bos taurus* UMD 3.1, http://useast.ensembl.org/Bos_taurus/Info/Index) using

the getBM() function in R-biomaRt package (https://www.bioconductor.org/) (Durinck et al., 2009). Suggestive significant SNPs at the peak regions (10% FDR) were also included for gene annotations. Genes that were located within the ±100kb from the identified SNPs were considered as positional candidate genes and selected for further functional analysis. The lists of candidate genes identified within the significant peak regions associated with AMIR or CMIR traits were uploaded into the *Qiagen's Ingenuity Pathway Analysis (IPA www.ingenuity.com)*. This analysis was performed to identify potential pathways, gene networks and functional annotations of the identified positional candidate genes.

Results and Discussion

Genomic based heritability (h2) estimates were 0.37 (std. error = 0.025) and 0.16 (std. error = 0.021) for AMIR and CMIR, respectively. Pedigree based estimates for h2 were 0.45 (std. error = 0.037) for AMIR and 0.18 (std. error = 0.032) for CMIR. Due to missing heritability it is common for genomic h2 estimates to be lower than pedigree-based estimates. Results for genomic analysis of CMIR identified peaks on chromosomes 5, 16 and 23 using FDR level of 5% (figure 1). Identified associations in this study support the previously reported regions for CMIR on chromosome 5 and 23 at 5% FDR (Thompson-Crispi et al., 2014). However, the region on chromosome 16 was not reported previously and is likely a novel region for this trait. This additional region is likely due to the increased accuracy achieved with the larger sample size of the current study, different statistical methods and different panel densities. The genomic analysis for AMIR detected four significant SNPs on chromosomes 2, 11, and 14 (5% FDR) as well as associations on chromosomes 23 and 21 (1% FDR) (figure 2). These results are consistent with the previous study performed by Thompson-Crispi et al. (2014) which also reported significant regions on chromosomes 23 and 21. A higher number of SNPs were however detected in this study which may reflect the effect of a larger sample size.

Functional annotations revealed several relevant biological functions, for example "disease" and "bio-functions" were over represented among the positional candidate genes for CMIR and AMIR, respectively (Table 1 and 2). As noted in Table 2, the results for AMIR are similar to those observed for CMIR. Since both AMIR and CMIR typically work together, depending on the nature and stage of infection, to defend against invading pathogens it is not surprising that similar results are seen for both traits.

The investigation of physiological system development and functions of the genes associated with CMIR shows involvement of these genes in trafficking cells of the immune system (figure 3). Many of the functions regulated by these genes are related to T-lymphocytes, which is the major cell associated with CMIR (Chaplin, 2010). Figure 4 shows the significant genes associated with AMIR in relation to immunological disease and the functions regulated by these genes. From this figure it can be seen that one of functions regulated by these genes involves B-cells, the main cell involved in AMIR (Chaplin, 2010).

Results of this study confirm that there are significant markers associated with immune response traits and that these are associated with particular biological processes or functions of immune response relevant to either AMIR or CMIR or both. Taking into account the moderate to high estimates of heritability, the results here demonstrate that a genomics test could be utilised for immune response, leading to more accurate selection of traits that ensure overall resistance to disease in Holstein cattle.

Acknowledgements

This research was supported by grants to B Mallard from NSERC and the Semex Alliance.
This paper is also a contribution to the Food from Thought research program supported by the Canada First Research Excellence Fund and is part of Dairy at Guelph..

.

List of References

Canadian Dairy Network (CDN), 2017. Genetic Gain Before and After Genomics. https://www.cdn.ca/document.php?id=468 posted 27 Apr. 2017.

Chaplin, D.D., 2010. Overview of the immune response. J. Allergy Clin. Immunol. 125(Suppl 2): S3-S23.

Durinck, S., P.T. Spellman, E. Birney & W. Huber, 2009. Mapping identifiers for integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. 4(8):1184-1191.

Heriazon, A., K.A. Thompson, B.N. Wilkie, W. Mathes-Sears, M. Quinton & B.A. Mallard, 2009. Antibody to ovalbumin and delayed-type hypersensitivity to *Candida albicans* a and mycobacteria in lactating Holstein cows using Quil A or Freunds complete adjuvant. Vet. Immunology and Immunopathology. 127:220-227.

Koeck, A., F. Miglior, D.F. Kelton & F.S. Schenkels, 2012. Health recording in Canadian Holsteins: data and genetic parameters. J. Dairy Sci. 95(7):4099-4108.

Larmer, S.G. & B.A. Mallard, 2016. High immune response sires reduce disease incidence in North American large commercial dairy populations. J. Cattle Practice 24:80.

Mallard, B.A., M. Emam, M. Paibomesai, K. Thompson-Crispi & L. Wagter-Lesperance, 2015 Genetic selection of cattle for improved immunity and health. Jpn J. Vet. Res. 63:S37-44.

National Animal Health Monitoring System (NAHMS), 2007. Dairy 2007 Part 2: Changes in the U.S. Dairy Cattle Industry, 1991-2007. https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/monitoring-and-surveillance/nahms/nahms_dairy_studies posted Mar. 2008.

Sargolzaei, M. 2014. SNP1101 User's Guide. Version 1.0. HiggsGene Solutions Inc.

Sargolzaei, M., J. P. Chesnais & F.S. Schenkel, 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15(1):478.

Thompson-Crispi, K.A., A. Sewalem, F. Miglior & B. A. Mallard, 2012a. Genetic parameters of adaptive immune response traits in Canadian Holsteins. J. Dairy Sci. 95:401-409.

Thompson-Crispi, K.A., B. Hine, M. Quinton, F. Miglior & B. A. Mallard, 2012b. *Short Communication*: Association of disease incidence and adaptive immune response in Holstein dairy cows. J. Dairy Sci. 95:3888-3893.

Thompson-Crispi, K.A., M. Sargolzaei, R. Ventura, M. Abo-Ismail, F. Miglior, F. Schenkel & B.A. Mallard, 2014. A genome-wide association study of immune response traits in Canadian Holstein. BMC Genomics 15:559-568.

Table 1. Top disease and functions enriched in Ingenuity Pathway Analysis for cell-mediated immune response.

Name	P-value range	# Molecules
Inflammatory Response	3.75E-02 – 1.57E-04	10
Hematological Disease	5.23E-04 – 5.23E-04	3
Immunological Disease	5.23E-04 – 5.23E-04	3
Neurological Disease	4.84E-02 – 1.20E-03	4
Infectious Disease	1.99E-02 – 3.81E-03	3

[1]Qiagen's Ingenuity Pathway Analysis (IPA www.ingenuity.com)

Table 2. Top disease and functions enriched in Ingenuity Pathway Analysis for antibody-mediated immune response.

Name	P-value range	# Molecules
Infectious Disease	4.61E-02 – 2.67E-03	10
Inflammatory Response	4.80-02 – 2.67E-03	11
Immunological Disease	4.61E-02 – 1.1E-02	8
Respiratory Disease	4.39E-02 – 1.14E-02	5
Cancer	1.17E-02 – 1.17E-02	1

[1]Qiagen's Ingenuity Pathway Analysis (IPA www.ingenuity.com)

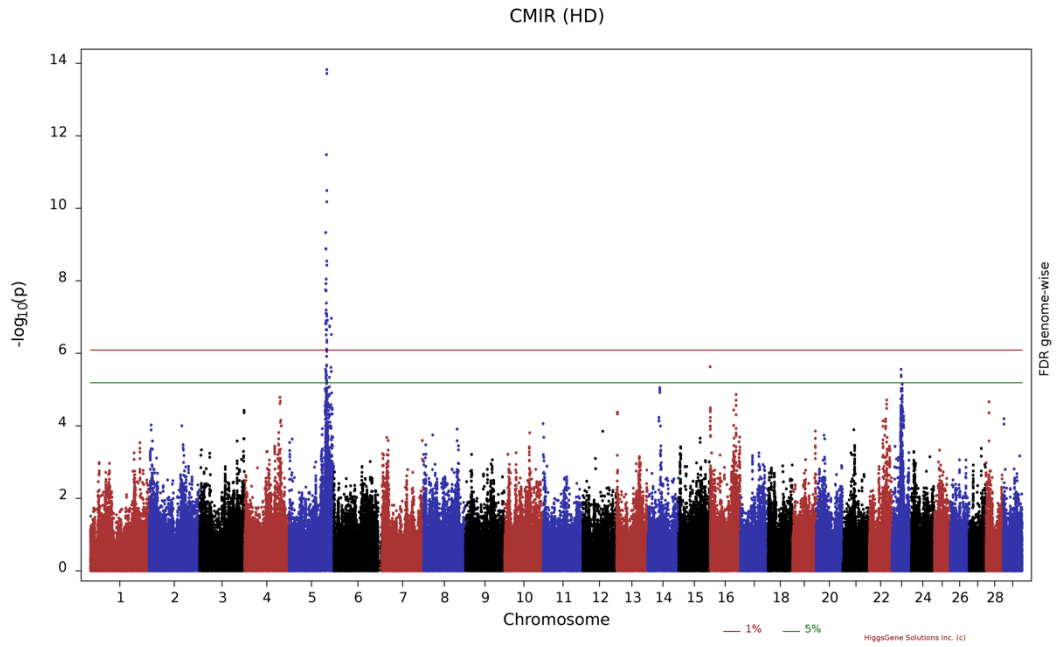

Figure 1. Manhattan plot of –log(p-value) for CMIR. The x-axis is the position of each SNP on the bovine chromosome and the y-axis is the –log$_{10}$P. The red and blue lines indicate 1% and 5% FDR, respectively.

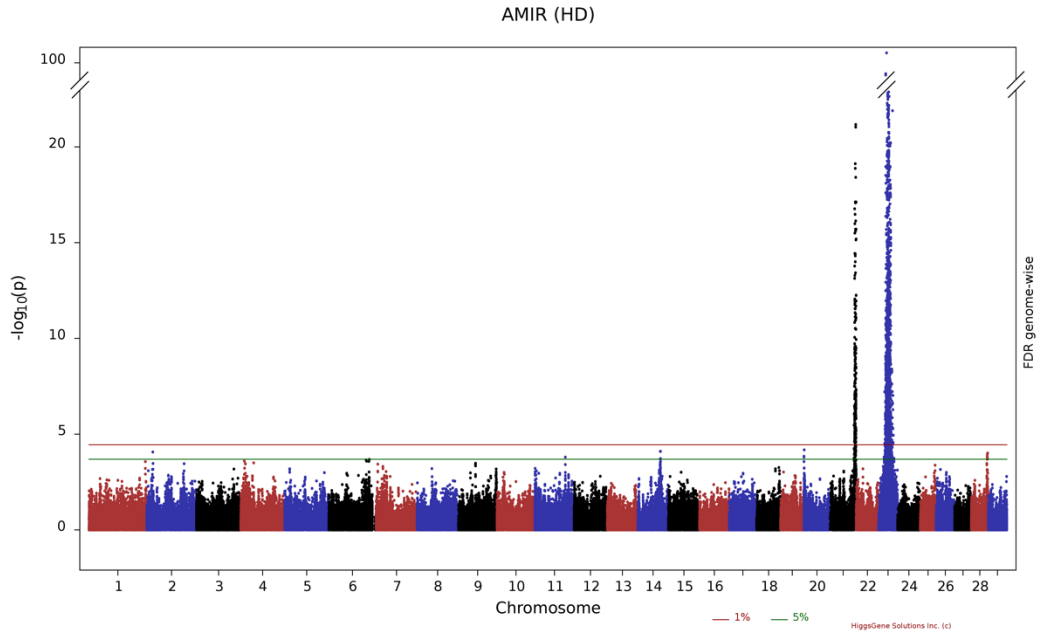

Figure 2. Manhattan plot of –log(p-value) for AMIR. The x-axis is the position of each SNP on the bovine chromosome and the y-axis is the –log$_{10}$P. The red and blue lines indicate 1% and 5% FDR, respectively.

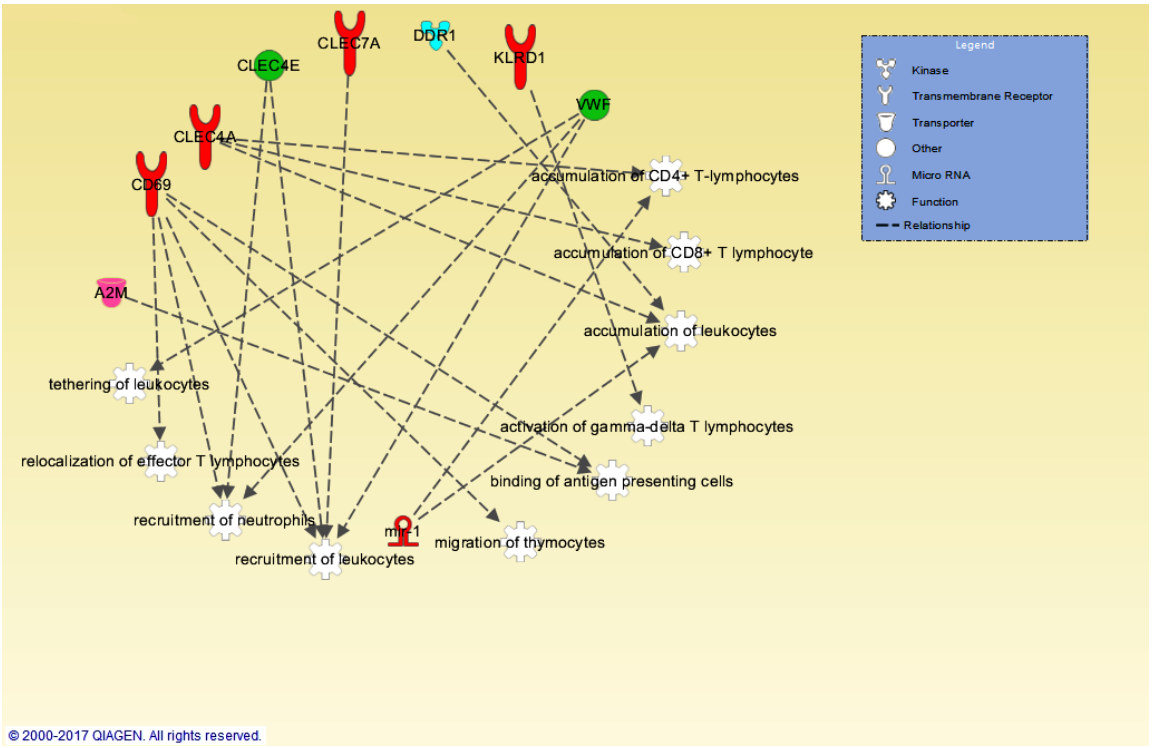

Figure 3. Immune cell trafficking pathways regulated by genes associated with CMIR emphasises the relevance of T-lymphocytes in this response

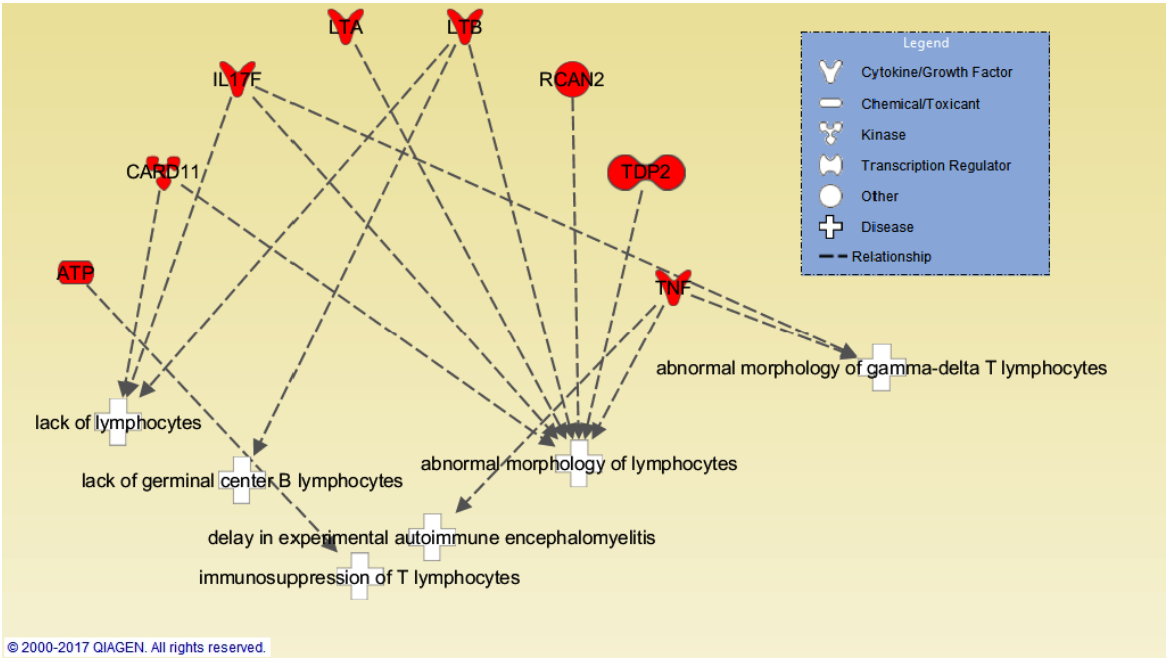

Figure 4. Immunological disease regulated by genes associated with AMIR emphasizes the relevance of cytokines and B-cell germinal centre formation in this response