



## OPEN ACCESS

## EDITED BY

Yanjun Zan,  
Swedish University of Agricultural  
Sciences, Sweden

## REVIEWED BY

Tara G. McDanel,  
Agricultural Research Service (USDA),  
United States  
Pablo Fonseca,  
University of Guelph, Canada

## \*CORRESPONDENCE

Graham S. Plastow,  
plastow@ualberta.ca

## SPECIALTY SECTION

This article was submitted to  
Evolutionary and Population Genetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 16 September 2022

ACCEPTED 28 November 2022

PUBLISHED 12 December 2022

## CITATION

Li J, Mukiibi R, Jiminez J, Wang Z,  
Akanno EC, Timsit E and Plastow GS  
(2022), Applying multi-omics data to  
study the genetic background of bovine  
respiratory disease infection in feedlot  
crossbred cattle.  
*Front. Genet.* 13:1046192.  
doi: 10.3389/fgene.2022.1046192

## COPYRIGHT

© 2022 Li, Mukiibi, Jiminez, Wang,  
Akanno, Timsit and Plastow. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which does  
not comply with these terms.

# Applying multi-omics data to study the genetic background of bovine respiratory disease infection in feedlot crossbred cattle

Jiyuan Li<sup>1</sup>, Robert Mukiibi<sup>2</sup>, Janelle Jiminez<sup>1</sup>, Zhiquan Wang<sup>1</sup>,  
Everestus C. Akanno<sup>1</sup>, Edouard Timsit<sup>3</sup> and Graham S. Plastow<sup>1\*</sup>

<sup>1</sup>Livestock Gentec, Department of Agriculture, Food & Nutritional Science, University of Alberta, Edmonton, AB, Canada, <sup>2</sup>The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, Scotland, United Kingdom, <sup>3</sup>Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada

Bovine respiratory disease (BRD) is the most common and costly infectious disease affecting the wellbeing and productivity of beef cattle in North America. BRD is a complex disease whose development is dependent on environmental factors and host genetics. Due to the polymicrobial nature of BRD, our understanding of the genetic and molecular mechanisms underlying the disease is still limited. This knowledge would augment the development of better genetic/genomic selection strategies and more accurate diagnostic tools to reduce BRD prevalence. Therefore, this study aimed to utilize multi-omics data (genomics, transcriptomics, and metabolomics) analyses to study the genetic and molecular mechanisms of BRD infection. Blood samples of 143 cattle (80 BRD; 63 non-BRD animals) were collected for genotyping, RNA sequencing, and metabolite profiling. Firstly, a genome-wide association study (GWAS) was performed for BRD susceptibility using 207,038 SNPs. Two SNPs (Chr5:25858264 and BovineHD1800016801) were identified as associated ( $p$ -value  $< 1 \times 10^{-5}$ ) with BRD susceptibility. Secondly, differential gene expression between BRD and non-BRD animals was studied. At the significance threshold used ( $\log_2FC > 2$ ,  $\log CPM > 2$ , and  $FDR < 0.01$ ), 101 differentially expressed (DE) genes were identified. These DE genes significantly ( $p$ -value  $< 0.05$ ) enriched several immune responses related functions such as inflammatory response. Additionally, we performed expression quantitative trait loci (eQTL) analysis and identified 420 cis-eQTLs and 144 trans-eQTLs significantly ( $FDR < 0.05$ ) associated with the expression of DE genes. Interestingly, eQTL results indicated the most significant SNP (Chr5: 25858264) identified *via* GWAS was a cis-eQTL for DE gene *GPR84*. This analysis also demonstrated that an important SNP (rs209419196) located in the promoter region of the DE gene *BPI* significantly influenced the expression of this gene. Finally, the abundance of 31 metabolites was significantly ( $FDR < 0.05$ ) different between BRD and non-BRD animals, and 17 of them showed correlations with multiple DE genes, which shed light on the interactions between immune response and metabolism. This study identified associations between genome, transcriptome, metabolome, and

BRD phenotype of feedlot crossbred cattle. The findings may be useful for the development of genomic selection strategies for BRD susceptibility, and for the development of new diagnostic and therapeutic tools.

#### KEYWORDS

disease susceptibility, GWAS, differential expressed gene, eQTL, causal SNP, metabolomics

## Introduction

Bovine respiratory disease (BRD) is a worldwide infectious disease that causes a significant economic loss in cattle production due to mortality, increased expenses associated with prevention, treatment and labour, impaired growth of animals, and reduced carcass value (Griffin, 1997; Smith, 2000; Irsik et al., 2006; Schneider et al., 2010). In North America, BRD has been identified as the most expensive infectious disease in the beef cattle industry (Taylor et al., 2010), causing high morbidity rates that can reach up to 80%, and moderate to high mortality in some feedlots (Smith, 1998; Baptista et al., 2017). Newly-received cattle in the feedlot are highly susceptible to BRD due to compromised immunity from a number of stressors including weaning, long-distance transportation, and co-mingling of cattle from different sources especially in auction markets (Taylor et al., 2010). These stressors expose the animals to multiple BRD pathogens and provide a conducive environment for the emergence of opportunistic viral and bacterial infections of the respiratory tract (Griffin et al., 2010; Taylor et al., 2010; Kirchhoff et al., 2014). In feedlots, the most commonly used disease control approach is treating animals with a wide range of antibiotics before or on entry into the feedlots (Ives and Richeson, 2015), however this may lead to the development of antimicrobial resistance which is a major concern for both human and animal health (Klima et al., 2014; Stanford et al., 2020). Additionally, early diagnosis and appropriate treatment of infected animals could increase recovery from infections and potentially reduce negative impacts of the disease on animal performance and productivity. However, most clinical signs of BRD are subjective, difficult to standardize, and non-specific for BRD, which makes the diagnosis of BRD difficult (Buczinski and Pardon, 2020).

Heritability estimates of BRD range from 0.07 to 0.29 (Snowder et al., 2005; Schneider et al., 2010; Neibergs et al., 2014a; 2014b). This suggests there is the potential to breed BRD resistant animals, which would lead to a sustainable reduction in BRD incidence and potentially antimicrobial resistance (Neibergs et al., 2014b; Hoff et al., 2019). Several SNPs and quantitative trait loci (QTLs) have been reported as significantly associated with BRD through genome-wide association studies (GWAS) (Neibergs et al., 2014b; Hoff et al., 2019), however, investigations into the genetic background of resistance or susceptibility to BRD in beef cattle populations is an ongoing

endeavour. In addition, RNA sequencing (RNA-Seq) offers high resolution profiling of transcriptomes of individual animals in a given sample/tissue, hence allowing the discovery of transcriptome-wide expression differences between animals with contrasting phenotypes of interest (e.g., BRD and non-BRD) (Costa-Silva et al., 2017; Hrdlickova et al., 2017). Such differential gene expression between BRD and non-BRD animals could help reveal differences in the host response to BRD infection and help to identify potential biomarkers for BRD diagnosis. Several transcriptomic studies have been conducted to investigate the gene expression differences and host response to BRD infection (Tizioto et al., 2015; Scott et al., 2020; Sun et al., 2020; Jimenez et al., 2021). These studies have shown that host animals have the ability to respond to BRD pathogen infection and the damage caused by these pathogens by altering the expression of certain key immune genes (e.g., *BPI*, *GPR84*, *S100A8*, *S100A9*, and *IL3RA*). These studies further suggest that gene expression alterations may vary with different viral and bacterial pathogen infections as well as the stage of disease development. However, these analyses have mainly focused on the correlation between the transcriptome and BRD infection status, with no consideration of the potential relationship between different omics layers. Changes in gene expression are not only associated with the disease or trait of interest, but are also affected by genomic regulation (i.e., expression QTL, eQTL) (Cookson et al., 2009). Thus, integration of GWAS, differential gene expression analysis and eQTL analysis could aid in interpreting the results of GWAS and identifying functional or causal SNPs. Furthermore, metabolites are small molecules involved in metabolic activities, which could provide additional insights into the host response to disease, and could also be used as biomarkers to indicate the presence or absence of the disease (Xia et al., 2013; Blakebrough-Hall et al., 2020). Identifying metabolites associated with BRD, and studying their correlations with DE genes, may lead to a better understanding of the biological basis for disease response and could help to identify more biomarkers involved in disease pathogenesis. For BRD, however, this knowledge is currently largely unknown.

Therefore, the objectives of this study were to identify DNA markers associated with BRD susceptibility, the DE genes and significant metabolites associated with BRD infection in feedlot cattle. We also aimed to identify cis- and trans-eQTLs associated with DE genes and correlations between DE genes and significant

metabolites in this study. This multi-omics study may shed light on the genetic and molecular architecture of BRD in crossbred feedlot cattle and provide information on biomarker identification in the general beef cattle population in Canada.

## Materials and methods

### Animal population and phenotype collection

A total of 143 crossbred or multi-breed beef cattle were used in this study. Animals were conventionally raised cattle that included heifers ( $n = 87$ ) and steers ( $n = 56$ ). These animals were enrolled into the feedlot during the fall of 2015 at four commercial feedlots in Central/Southern Alberta, Canada. On-arrival processing of the animals was previously described (Jiminez et al., 2021). Briefly, animals were weighed (average body weight:  $571 \pm 106$  Lb) and received a subcutaneous injection of a long-acting macrolide (tulathromycin, Draxxin, 2.5 mg/kg, Zoetis, Kirkland, QC, Canada) and vaccinated against multiple bacterial and viral agents including bovine herpes virus-1 (BoHV-1), bovine viral diarrhea virus (BVDV) (types I and II), bovine parainfluenza-3 (PI3V), bovine respiratory syncytial virus (BRSV), *Mannheimia haemolytica*, *Histophilus somni*, and clostridial pathogens. They were also dewormed with a pour-on ivermectin solution. While in the feedlots, animals were fed twice daily on a concentrate barley-based receiving/growing diet. This diet also contained 25 ppm of monensin (Rumensin 200, Elanco, Guelph, ON, Canada) and 35 ppm of chlortetracycline (Aureomycin 220, Zoetis). Cattle received a growth implant and a second vaccination against infectious viruses at approximately 30 days after arrival to the feedlot. Animals were monitored daily and those that showed signs and symptoms of BRD (depression, nasal or ocular discharge, cough, tachypnea, or dyspnea) within 50 days of arrival at the feedlots were clinically examined by an experienced veterinarian. The details of clinical examinations and case definition (i.e., BRD or non-BRD cattle) were previously described (Jiminez et al., 2021). Briefly, animals were retrospectively identified as BRD positive based on clinical examination and serum haptoglobin concentration. The animal was confirmed as BRD positive if the animal displayed at least one visual BRD symptom, had a rectal temperature  $\geq 40^{\circ}\text{C}$ , abnormal lung sounds detected at auscultation, a serum haptoglobin concentration  $\geq 0.25$  g/L, and had no prior treatment against BRD or other diseases during the feedlot period (i.e., first BRD occurrence). For each animal that was BRD positive blood samples were collected by jugular vein puncture using Tempus (Thermo Fisher Scientific, ON), EDTA and heparin tubes (BD, Mississauga, Canada). In addition, in each

pen that had a BRD positive animal, blood samples were similarly collected from two healthy matched-control or non-BRD cattle (i.e., animals which had no visual signs of BRD or other diseases, a rectal temperature  $< 40^{\circ}\text{C}$ , no abnormal lung sounds detected at auscultation and a serum haptoglobin concentration  $< 0.25$  g/L). A large proportion of apparently healthy pen-mates had rectal temperature  $> 40^{\circ}\text{C}$ , abnormal lung sounds at auscultation and/or serum haptoglobin  $\geq 0.25$  g/L. However, it is expected to find a large proportion of apparently healthy cattle with either fever, abnormal lung sounds at auscultation or high serum haptoglobin as subclinical BRD is very common early in the feeding period (Timsit et al., 2011). Therefore only 63 controls were included in this study. It was also worth noting that these 63 non-BRD cattle did not become BRD positive later during the experiment (i.e., they remained healthy until sent to market). After sample collection, animals identified as BRD positive received an antibiotic treatment subcutaneously in combination with non-steroidal anti-inflammatory drugs, in accordance with feedlot treatment protocols. The blood samples collected from both BRD positive and non-BRD animals were used for DNA extraction for GWAS analysis, transcriptome profiling for gene expression analysis and metabolome abundance profiling for the metabolomics analysis.

### DNA extraction, genotyping, and genomic breeding composition estimation

DNA was extracted from 50  $\mu\text{L}$  whole blood sample of each animal using the sBeadex Livestock kit (LGC, Berlin, Germany). Four animals were not genotyped because their blood samples were lost. The extracted DNA was used to genotype each animal for 100,000 SNPs using Illumina's GGP Bovine 100 K microarray SNP chip (Illumina, San Diego, CA, United States) by Neogen (Lincoln, NE, United States). The SNPs located on the sex chromosomes were excluded from the analysis. In addition, SNPs that had minor allele frequency  $< 5\%$ , or missing allele rate  $> 10\%$  and those that failed to pass the Hardy-Weinberg equilibrium test ( $p$ -value  $< 0.0001$ ) were also excluded from the analysis. The remaining SNPs (i.e., 85,100 SNPs) were used to predict breed composition of each animal using ADMIXTURE software (v1.3.0) (Alexander and Lange, 2011). The reason for this analysis is due to the crossbred nature of the animals in our population, which should be considered in the statistical analysis. The ancestry value of  $K = 3$  was used to define the source of genetic makeup because it had the smallest cross-validation error and yielded the most accurate breed composition prediction based on prior knowledge of breed composition for a subset of animals. The genomic breed composition of each animal was shown in Supplementary Figure S1.

Additional SNP markers and their genotypes were called from the RNA-Seq data of all animals using the Genome Analysis Toolkit best practices (GATK v3.8) (Van der Auwera and O'Connor, 2020). Prior to the variant calling processes, the mapped reads from two-pass STAR alignment were sorted, had read groups added, and duplicates identified using the Picard tools package (v2.20.6). A series of processing steps including splitting “N” cigar reads (i.e., splice junction reads), reassigning mapping quality score, and base quality score recalibration were performed to improve variant calling accuracy using GATK. After data preprocessing, variants were called using the HaplotypeCaller algorithm in Genomic Variant Call Format (GVCF) mode, which included two steps: (i) variants were called individually on each sample, generating one GVCF file per sample that lists genotype likelihoods and their genome annotations; (ii) variants were called from the GVCF file through a joint genotyping analysis. The joint genotyping method is more flexible and technically easier, and is recommended for variant calling in RNA-Seq experiments (Poplin et al., 2017; Brouard et al., 2019). Stringent filtering procedures were applied to variants using the GATK Variant Filtration tool and VCFtools (v0.1.14) (Danecek et al., 2011). Indels, non-biallelic SNPs and SNPs on sex chromosomes were excluded. Then SNPs with QD < 3.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0, SOR > 3.0, minor allele frequency < 5%, missing allele rate > 10% and severe departure from Hardy-Weinberg equilibrium ( $p$ -value < 0.0001) were removed.

Finally, the two SNP datasets (genotype derived vs. RNAseq derived) were merged based on the position of SNPs on the chromosome using Plink (v1.90b6.7) (Chang et al., 2015). For the overlapping SNPs in the two SNP datasets, the SNPs derived from genotyping were used. A total of 207,038 SNPs were available for 138 animals (79 BRD; 59 non-BRD animals as one sample failed to yield quality genotypes from the RNA-seq data) and used in GWAS and eQTL analysis.

## Genome-wide association analysis for BRD susceptibility

Prior to performing the GWAS for BRD susceptibility, the phenotype of BRD (BRD status of the animal, as a binary trait) was first fitted into a logistic model with a fixed effect of “feedlot,” and two covariates including days on feed, and genomic breed composition from admixture breed composition analysis. This modeling was used to determine which fixed effect and covariates had significant effects on the phenotype. Next, the GWAS analysis between SNP marker genotypes (from SNP chip and RNA-seq data) and adjusted BRD status was performed using the single SNP-based mixed linear model association (mlma), as implemented in the GCTA package (v1.93.2) (Yang et al., 2011). The linear mixed model can be described as follows:

$$y_{ij} = \mu + b_j x_{ij} + a_i + e_{ij}$$

where  $y_{ij}$  is the adjusted phenotypic value of the  $i$  th animal with the  $j$  th SNP (i.e., the  $ij$  th animal),  $b_j$  is the allele substitution effect of the  $j$  th SNP,  $x_{ij}$  is the  $j$  th SNP genotype of animal  $i$  coded as 0, 1, 2 for genotypes  $A_1A_1$ ,  $A_1A_2$ , and  $A_2A_2$ , respectively,  $a_i$  is the additive polygenic effect of the  $i$  th animal  $\sim N(0, G\sigma_a^2)$ , and  $e_{ij}$  is the random residual effect of the  $ij$  th animal  $\sim N(0, I\sigma_e^2)$ . The genomic relationship matrix  $G$  was derived based on total filtered SNP markers (207,038 SNPs) as described by Yang et al. (2014), which is essentially the same as VanRaden’s second formulation (VanRaden, 2008). The SNP allele substitution effect was estimated, and the significance test of the SNP allele substitution effect was conducted *via* a generalized least square F-test as implemented in the GCTA package. The phenotypic variance explained by each significant SNP was calculated by  $\frac{2pq\beta^2}{S^2} \times 100\%$ , where  $p$  and  $q$  denote the SNP allele frequency of  $A_1$  and  $A_2$ , respectively;  $\beta$  is the SNP allele substitution effect;  $2pq\beta^2$  is the additive variance of the SNP, and  $S^2$  is the phenotypic variance.

Those SNP with FDR < 0.05 were identified as significant SNPs. The threshold of  $p$ -value <  $1 \times 10^{-5}$  were considered as the suggestive line. To visually summarize the GWAS results both the quantile-quantile (Q-Q) plot and Manhattan plot were generated using the R package qqman (Turner, 2014).

## RNA isolation, cDNA library preparation and sequencing

Tempus tubes were thoroughly mixed by 20 inversions and stored on ice until transported to the laboratory and stored at  $-20^\circ\text{C}$ . Total RNA was extracted in two batches (Batch 1,  $n = 47$ ; Batch 2,  $n = 96$ ). Similar procedures were performed on all samples in both batches. Initially, total RNA was isolated from blood using a Preserved Blood RNA Purification Kit (Norgen Biotek Corp, Thorold, ON, Canada), and the quality of RNA was measured using the 2200 RNA ScreenTape TapeStation System (Agilent Technologies Inc. Cedar Creek, TX, United States) producing RNA integrity numbers (RIN) ranging from 8.0 to 9.8. Thereafter, cDNA libraries were prepared for sequencing for each individual animal from the high-quality RNA using the TruSeq RNA Library Preparation kit v2 (Illumina, San Diego, CA, United States) and the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (New England Biolabs Ltd. Whitby, ON, Canada). Samples in both batches used the stranded library preparation process. Paired-end sequencing was performed using the HiSeq 4000 platform and Novaseq 6000 for batch 1 and batch 2 samples respectively to generate paired-end sequences of 100 bp read length. Sequencing of samples in the two batches was performed at the McGill University and Genome Quebec Innovation Center (Montreal, QC, Canada). Finally, the raw

reads of 143 samples (80 BRD; 63 non-BRD) were obtained and used for downstream analyses.

## Sequence data processing, alignment and counting

Raw reads for each sample were assessed for quality using FastQC (v0.11.8) (Andrews, 2010). The bases with low quality score (Phred quality score <20) and 3' adapter sequences on raw reads were removed using Trimmomatic (v0.39) (Bolger et al., 2014). These cleaned-up sequences were aligned to the *Bos taurus* reference genome (ARS-UCD1.2.98, downloaded from Ensembl genome browser) using a short read alignment software STAR (v2.7.1a) with paired-end default parameters (Dobin et al., 2013). FeatureCounts (SubRead v1.6.4) was used to count the reads that aligned to a particular annotated gene in the bovine genome (Liao et al., 2014) and these counts were consequently used for differential gene expression analysis between BRD and non-BRD animals.

## Differential gene expression analysis and functional enrichment analysis

Read counts per gene generated by FeatureCounts as described above were utilized for differential gene expression analysis between BRD and non-BRD animals using the R Bioconductor package edgeR (McCarthy et al., 2012). Firstly, lowly expressed (count per million or CPM <0.5 in at least 63 samples) genes were filtered out from the analysis. Counts of the retained genes were then normalized using the trimmed mean M values (TMM) method (Robinson and Oshlack, 2010), to account for the technical variations between samples that may have been caused by the RNA extraction, cDNA library construction, and differences in sequencing depth (Robinson and Oshlack, 2010). The normalized counts were then modeled for differential gene expression between BRD and non-BRD animals using generalized linear models (GLM) under a negative binomial distribution assumption that considered feedlot, genomic breed composition, and sequencing batch as fixed effects. To test for significance of differential expression of a gene between BRD and non-BRD groups, a likelihood ratio test was performed, and those genes with Benjamini-Hochberg false discovery rate (FDR) < 0.01, log fold change ( $\log_2FC$ ) > 2, and log counts per million ( $\log_2CPM$ ) > 2 were identified as significant differentially expressed genes between BRD and non-BRD animals. The non-BRD/healthy control animals were set as the reference group in the contrast analysis, therefore, DE genes that were upregulated or downregulated in BRD animals were relative to the non-BRD animals.

Functional enrichment for the DE genes was performed using the Ingenuity Pathway Analysis software (IPA; [www.ingenuity.com](http://www.ingenuity.com)) using their Ensembl gene ID and their  $\log_2$  fold change as the inputs. In this study, biological functions were considered significantly enriched if the  $p$ -value for the overlap comparison test between the input gene list and the knowledge base of IPA for a given biological function was less than 0.05.

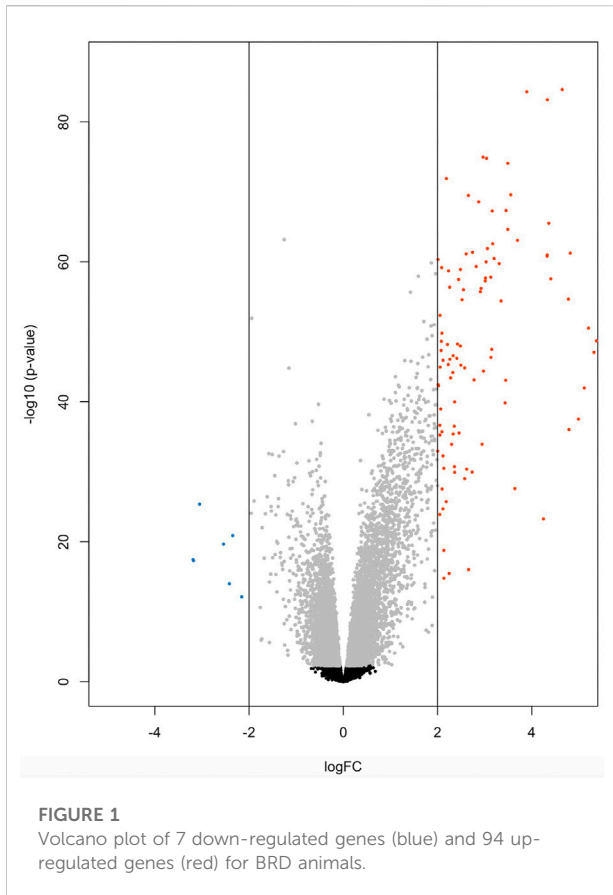
## eQTL analysis and eQTL annotation

We further performed eQTL analysis to identify associations between expression of differentially expressed genes and SNP genotypes. Log transformed normalized counts ( $\log_2CPM$ ) values of 93 protein-coding DE genes on autosomes and 207,038 SNPs were used in the eQTL analysis. The analysis of the linear model was fitted to test the association of each single gene expression and genotype classes of a SNP implemented in the R package MatrixEQTL (Shabalin, 2012). “Feedlot”, “sequencing batch” and “genomic breed composition of animals” were also fitted in the model to correct for any variability in gene expression that could have been due to these factors. SNPs located within 1 Mbp around the gene transcription starting site (TSS) were tested for cis-associations, while SNPs located further than 1 Mbp or on other chromosomes were tested for trans-associations. Only those associations with FDR <0.05 were considered significant cis- or trans-eQTLs. The significant eQTLs were then annotated as located in the TSS-promoter, exonic, intronic, transcription termination site (TTS) or intergenic regions using the annotatePeaks.pl script of HOMER software (<http://homer.ucsd.edu/homer/ngs/annotation.html>).

## Sample preparation and nuclear magnetic resonance spectroscopy

The plasma metabolome was generated at The Metabolomics Innovation Centre (TMIC, Edmonton, AB, Canada). Samples underwent a deproteinization step involving ultra-filtration as previously described by Psychogios et al. (2011) to remove proteins whose presence affects the identification of small molecular weight metabolites by NMR spectroscopy. Prior to filtration, 3 KDa cut-off centrifugal filter units (Amicon Microcon YM-3) were rinsed five times each with 0.5 ml of H<sub>2</sub>O and centrifuged (10,000 rpm for 10 min) to remove residual glycerol bound to the filter membranes. Aliquots of each plasma sample were then transferred into the centrifuge filter devices and spun (10,000 rpm for 20 min) to remove macromolecules (primarily protein and lipoproteins) from the sample. The filtrates were checked visually for any evidence that the membrane was compromised and for such samples the filtration process was repeated with a different filter and the





filtrate inspected again. The subsequent filtrates were collected, and the volumes were recorded. If the total volume of the sample was under 250  $\mu\text{L}$  an appropriate amount of 150 mM  $\text{KH}_2\text{PO}_4$  buffer (pH 7) was added until the total volume of the sample was 173.5  $\mu\text{L}$ . Any sample that had to have buffer added to bring the solution volume to 173.5  $\mu\text{L}$  was annotated with the dilution factor and metabolite concentrations were corrected in the subsequent analysis. Subsequently, 46.5  $\mu\text{L}$  of a standard buffer solution (54%  $\text{D}_2\text{O}$ :46% 1.75 mM  $\text{KH}_2\text{PO}_4$  pH 7.0 v/v containing 5.84 mM DSS (2,2-dimethyl-2-silcepentane-5-sulphonate)) was added to the sample.

The sample (250  $\mu\text{L}$ ) was then transferred to a 3 mm SampleJet NMR tube for subsequent spectral analysis. All  $^1\text{H}$ -NMR spectra were collected on a 700 MHz Avance III (Bruker) spectrometer equipped with a 5 mm HCN Z-gradient pulsed-field gradient (PFG) cryoprobe.  $^1\text{H}$ -NMR spectra were acquired at 25°C using the first transient of the nuclear Overhauser enhancement spectroscopy (NOESY) pre-saturation pulse sequence (noesy1dpr), chosen for its high degree of quantitative accuracy (Saude et al., 2006). All FID's (free induction decays) were zero-filled to 250 K data points. The singlet produced by the DSS methyl groups was used as an internal standard for chemical shift referencing (set to 0 ppm)

and for quantification all  $^1\text{H}$ -NMR spectra were processed and analyzed using the Chenomx NMR Suite Professional Software package version 7.1 (Chenomx Inc., Edmonton, AB, Canada). The Chenomx NMR Suite software allows for qualitative and quantitative analysis of an NMR spectrum by manually fitting spectral signatures from an internal database to the spectrum. Specifically, the spectral fitting for metabolite was done using the standard Chenomx 700 MHz metabolite library. Typically, 90% of visible peaks were assigned to a compound and more than 90% of the spectral area could be routinely fit using the Chenomx spectral analysis software. Most of the visible peaks are annotated with a compound name. Each spectrum was processed and analyzed by at least two NMR spectroscopists to minimize compound misidentification and misquantification.

## Identification of significant metabolites and their correlations with DE genes

Fifty-four metabolites were quantified by NMR analysis followed by a logarithmic transformation to the raw data. The transformed data followed a normal distribution and were further adjusted by two fixed effects (feedlot and sex), and a covariate of genomic breed composition in linear models. The adjusted values were used to identify metabolites that had concentrations significantly (FDR < 0.05) different between BRD and non-BRD animals using two independent sample t-tests. Furthermore, the correlations between 31 significant metabolites and 93 protein-coding DE genes were computed in R.

## Results

### SNPs associated with BRD susceptibility

In this study, no significant SNPs remained after FDR correction. However, two SNPs (Chr5:25858264 on chromosome 5 and BovineHD1800016801 on chromosome 18) were above the suggestive line ( $p$ -value  $< 1 \times 10^{-5}$ ) (Table 1; Supplementary Figure S2). Both SNPs are in intronic regions of genes, i.e., SNP Chr5:25858264 is located in the first intron of *SMUG1* while BovineHD1800016801 is located in the third intron of *IGLON5*.

### Transcriptomic architecture of BRD infection in feedlots

At the significance threshold of  $\log_2\text{FC} > 2$ ,  $\log\text{CPM} > 2$  and FDR < 0.01, 101 genes were identified as differentially expressed between BRD and non-BRD animals (non-BRD as the reference

TABLE 1 SNPs significantly associated with BRD susceptibility.

SNP	Chromosome	Position (bp)	Minor allele frequency	b	se	p-value
Chr5:25858264	5	25858264	0.066	1.181	0.261	$5.85 \times 10^{-6}$
BovineHD1800016801	18	57400705	0.203	0.674	0.151	$7.65 \times 10^{-6}$

group), of which 7 and 94 were downregulated and upregulated respectively in the infected animals (Figure 1). The full list of DE genes with related descriptions and statistics is provided in Table 2. Our result showed that *interleukin 3 receptor subunit alpha* (*IL3RA*) was the most significant ( $FDR = 6.6 \times 10^{-81}$ ) upregulated gene, whereas *hemoglobin subunit beta* (*HBB*) was the most significant ( $FDR = 1.25 \times 10^{-24}$ ) downregulated gene (Table 2). In terms of expression differences, *leucine rich alpha-2-glycoprotein 1* (*LRG1*) showed the highest expression ( $\log_2FC = 7.35$ ) in the BRD animals relative to non-BRD animals, while *hemoglobin subunit alpha 1* (*HBA1*) showed the lowest expression ( $\log_2FC = -3.19$ ) in the same animal contrast (Table 2).

Of the 101 DE genes, 88 were successfully mapped to the IPA database for functional enrichment analysis. The DE genes were significantly ( $p$ -value  $< 0.05$ ) involved in 17 immune response related biological functions of which inflammatory response was the most significant with 60 DE genes. The top 10 most enriched functions are presented in Table 3, while all 17 functions are presented in Supplementary Table S1. Within the inflammatory response function, 3 DE genes (*ARG1*, *ALOX15*, and *ALAS2*) were downregulated, and 57 DE genes (e.g., *IL3RA*, *LRG1*, *BPI*, *CFB*, *GPR84*, *MMP9*, and *CA4*) were upregulated in the BRD animals. Furthermore, within the inflammatory response function, enriched innate immune response related processes such as leukocyte immune response, activation and migration of macrophages and neutrophils, and antimicrobial response were predicted to be activated or upregulated in the BRD animals (Figure 2). Adaptive immune response related processes such as activation of antigen processing cells, and cellular immune response were also identified as enriched, and predicted to be activated in the BRD animals. Some of the key DE genes as demonstrated by their involvement in numerous immune functions included *LCN2*, *S100A8*, *S100A9*, *S100A12*, *LTF*, *IL12B*, *CHI3L1*, and *DEFB4A* (Figure 2).

## Gene expression and genotype associations

At  $FDR < 0.05$ , we identified 420 cis-eQTLs and 144 trans-eQTLs associated with the expression of DE genes (Supplementary Table S2, S3). Some cis-eQTLs and trans-eQTLs were associated with more than one DE gene associated with BRD. For example, the SNP Chr6:

110850346 was cis-eQTL associated with the expression of the DE gene *BST1* and a trans-eQTL associated with the expression of another 6 DE genes (*GPR84*, *NUPR1*, *ART5*, *CFB*, *SLC6A2*, and *ADGRE1*). Similarly, the expression of a DE gene could also be associated with more than one cis- or trans-eQTLs. Of note, the eQTL analysis showed that the SNP (Chr5:25858264) with the smallest  $p$ -value in GWAS (Table 1) was cis-eQTL associated with the expression of the DE gene *GPR84* (Supplementary Table S2, S3). Additionally, 2 potential trans-eQTL hotspots (rs207554348 on chromosome 3 and Chr2:118164919 on chromosome 2) were observed (Supplementary Table S3). Finally, the eQTL annotation showed that the eQTL SNPs identified in this study were mostly located in intronic and exonic regions rather than intergenic regions (Figure 3; Supplementary Table S4). The high proportion of eQTLs observed in intronic and exonic regions may be due to the uneven distribution of SNPs used in this study, with 58.8% of SNPs derived from RNA-seq data.

## Metabolites associated with BRD infection and correlations between metabolites and DE genes

A total of 31 metabolites showed significant abundance difference between BRD and non-BRD animals (Table 4). Twenty metabolites had lower abundance in BRD animals as compared to non-BRD animals, for example, citric acid showed the most significant difference, and was significantly more abundant in non-BRD animals (Table 4). However, we also observed a few metabolites that were more strongly abundant in BRD animals than non-BRD animals, such as D-mannose, L-phenylalanine, and L-carnitine. In addition, 17 significant metabolites also showed significant ( $FDR < 0.05$ ) correlations with the expression of DE genes (Supplementary Table S5). It is worth noting that 15 metabolites had significant correlations with over 74 DE genes, with citric acid, 3-hydroxybutyric acid, acetic acid, L-glutamic acid, D-mannose, L-carnitine, showing correlation with more than 90 DE genes. For those metabolites that were more significantly associated with BRD, they also have a greater number of correlations with DE genes. With respect to DE genes, our results also showed that most (88) of the DE genes had significant correlations

TABLE 2 Differentially expressed genes between BRD and non-BRD animals.

Gene ID	Gene name	Log2 (fold change)	Log counts per million	p-value	FDR
ENSBTAG00000054844	<i>HBA1</i>	-3.188	11.000	$3.80 \times 10^{-18}$	$4.95 \times 10^{-17}$
ENSBTAG00000051412	<i>HBA</i>	-3.178	10.999	$5.02 \times 10^{-18}$	$6.47 \times 10^{-17}$
ENSBTAG00000037644	<i>HBB</i>	-3.050	11.089	$4.35 \times 10^{-26}$	$1.25 \times 10^{-24}$
ENSBTAG00000013178	<i>ALAS2</i>	-2.542	5.314	$2.27 \times 10^{-20}$	$3.70 \times 10^{-19}$
ENSBTAG00000004824	<i>REEP1</i>	-2.419	2.113	$1.02 \times 10^{-14}$	$9.21 \times 10^{-14}$
ENSBTAG00000011990	<i>ALOX15</i>	-2.347	4.915	$1.35 \times 10^{-21}$	$2.46 \times 10^{-20}$
ENSBTAG00000012403	<i>ARG1</i>	-2.157	2.589	$7.74 \times 10^{-13}$	$5.65 \times 10^{-12}$
ENSBTAG00000011465	<i>MYBPH</i>	2.002	6.087	$1.16 \times 10^{-33}$	$6.80 \times 10^{-32}$
ENSBTAG00000010464	<i>MN1</i>	2.009	2.526	$4.76 \times 10^{-61}$	$3.00 \times 10^{-58}$
ENSBTAG00000006156	<i>BST1</i>	2.012	5.490	$3.76 \times 10^{-43}$	$5.60 \times 10^{-41}$
ENSBTAG00000020430	<i>GLT1D1</i>	2.020	3.082	$5.24 \times 10^{-43}$	$7.63 \times 10^{-41}$
ENSBTAG00000050072		2.047	2.427	$1.30 \times 10^{-24}$	$3.23 \times 10^{-23}$
ENSBTAG00000052465		2.048	6.792	$2.32 \times 10^{-37}$	$1.97 \times 10^{-35}$
ENSBTAG00000013368	<i>ANKRD22</i>	2.049	3.861	$5.75 \times 10^{-36}$	$4.16 \times 10^{-34}$
ENSBTAG00000002148	<i>RAB3D</i>	2.053	7.196	$4.63 \times 10^{-53}$	$1.36 \times 10^{-50}$
ENSBTAG00000023648	<i>ART5</i>	2.055	7.114	$1.13 \times 10^{-45}$	$2.00 \times 10^{-43}$
ENSBTAG00000011037	<i>RBPM52</i>	2.067	2.973	$1.10 \times 10^{-39}$	$1.16 \times 10^{-37}$
ENSBTAG00000039556	<i>WPI1</i>	2.076	4.455	$4.77 \times 10^{-48}$	$1.05 \times 10^{-45}$
ENSBTAG00000006921	<i>ABCA6</i>	2.081	6.472	$2.33 \times 10^{-49}$	$5.61 \times 10^{-47}$
ENSBTAG00000013555	<i>ACVR1B</i>	2.088	5.897	$6.85 \times 10^{-60}$	$3.49 \times 10^{-57}$
ENSBTAG00000000783	<i>TGFA</i>	2.093	2.667	$1.94 \times 10^{-36}$	$1.47 \times 10^{-34}$
ENSBTAG00000004150	<i>NRG1</i>	2.093	4.792	$1.59 \times 10^{-50}$	$4.13 \times 10^{-48}$
ENSBTAG00000022779	<i>OLFM4</i>	2.097	3.277	$2.97 \times 10^{-28}$	$1.04 \times 10^{-26}$
ENSBTAG00000006990	<i>MYRF</i>	2.114	2.578	$2.04 \times 10^{-25}$	$5.47 \times 10^{-24}$
ENSBTAG00000037826		2.115	2.446	$5.59 \times 10^{-33}$	$3.01 \times 10^{-31}$
ENSBTAG00000004716	<i>RETN</i>	2.133	5.036	$1.76 \times 10^{-19}$	$2.63 \times 10^{-18}$
ENSBTAG00000014046	<i>BPI</i>	2.135	5.648	$1.67 \times 10^{-15}$	$1.64 \times 10^{-14}$
ENSBTAG00000018016	<i>NUPR1</i>	2.184	5.146	$1.94 \times 10^{-26}$	$5.74 \times 10^{-25}$
ENSBTAG00000014122	<i>FOXRED1</i>	2.188	6.744	$1.28 \times 10^{-72}$	$3.39 \times 10^{-69}$
ENSBTAG00000054765	<i>PGLYRP4</i>	2.210	2.848	$6.61 \times 10^{-49}$	$1.54 \times 10^{-46}$
ENSBTAG00000011677	<i>H1-2</i>	2.226	6.778	$4.82 \times 10^{-46}$	$8.87 \times 10^{-44}$
ENSBTAG00000013290	<i>DYSF</i>	2.233	7.225	$1.98 \times 10^{-59}$	$9.37 \times 10^{-57}$
ENSBTAG00000002635	<i>PGLYRP1</i>	2.247	2.172	$3.56 \times 10^{-16}$	$3.76 \times 10^{-15}$
ENSBTAG00000018223	<i>CHI3L1</i>	2.257	8.313	$4.32 \times 10^{-57}$	$1.59 \times 10^{-54}$
ENSBTAG00000010065	<i>TRPC5</i>	2.278	3.910	$3.87 \times 10^{-44}$	$6.10 \times 10^{-42}$
ENSBTAG00000007169	<i>P2RX1</i>	2.298	4.356	$1.25 \times 10^{-34}$	$7.89 \times 10^{-33}$
ENSBTAG00000001051	<i>OSCAR</i>	2.324	7.220	$6.67 \times 10^{-45}$	$1.12 \times 10^{-42}$
ENSBTAG00000013205	<i>IL1RAP</i>	2.331	5.612	$2.63 \times 10^{-47}$	$5.44 \times 10^{-45}$
ENSBTAG00000006904	<i>TNS2</i>	2.332	2.086	$4.01 \times 10^{-36}$	$2.92 \times 10^{-34}$
ENSBTAG00000004741	<i>IL12B</i>	2.355	2.380	$3.14 \times 10^{-37}$	$2.63 \times 10^{-35}$
ENSBTAG00000008389	<i>HTRA1</i>	2.358	2.118	$1.89 \times 10^{-31}$	$8.88 \times 10^{-30}$
ENSBTAG00000001785	<i>TGM3</i>	2.362	9.942	$1.06 \times 10^{-40}$	$1.23 \times 10^{-38}$
ENSBTAG00000018446	<i>GCA</i>	2.363	3.848	$1.22 \times 10^{-30}$	$5.35 \times 10^{-29}$
ENSBTAG00000013201	<i>ALOX5AP</i>	2.407	6.937	$6.31 \times 10^{-47}$	$1.23 \times 10^{-44}$
ENSBTAG00000020257	<i>PTPN5</i>	2.423	4.813	$5.81 \times 10^{-49}$	$1.37 \times 10^{-46}$
ENSBTAG00000018134	<i>AREG</i>	2.446	2.160	$3.37 \times 10^{-58}$	$1.31 \times 10^{-55}$
ENSBTAG00000003920	<i>TGM1</i>	2.457	2.683	$2.95 \times 10^{-36}$	$2.17 \times 10^{-34}$
ENSBTAG00000010007	<i>MAPK13</i>	2.484	4.014	$1.08 \times 10^{-48}$	$2.47 \times 10^{-46}$

(Continued on following page)



TABLE 2 (Continued) Differentially expressed genes between BRD and non-BRD animals.

Gene ID	Gene name	Log2 (fold change)	Log counts per million	p-value	FDR
ENSBTAG00000003519	<i>NOL3</i>	2.487	2.563	$1.27 \times 10^{-59}$	$6.23 \times 10^{-57}$
ENSBTAG00000005668	<i>SLC39A8</i>	2.494	2.926	$5.96 \times 10^{-46}$	$1.08 \times 10^{-43}$
ENSBTAG00000012638	<i>S100A12</i>	2.520	11.601	$2.66 \times 10^{-55}$	$8.20 \times 10^{-53}$
ENSBTAG00000008428	<i>UPP1</i>	2.549	7.233	$9.94 \times 10^{-57}$	$3.47 \times 10^{-54}$
ENSBTAG00000003353	<i>SLC6A2</i>	2.576	2.389	$1.48 \times 10^{-45}$	$2.58 \times 10^{-43}$
ENSBTAG00000048737	<i>DEFB10</i>	2.576	3.856	$1.00 \times 10^{-29}$	$4.06 \times 10^{-28}$
ENSBTAG00000006523	<i>SOD2</i>	2.607	9.705	$7.47 \times 10^{-62}$	$5.21 \times 10^{-59}$
ENSBTAG00000016566	<i>ITGA9</i>	2.618	3.070	$4.37 \times 10^{-31}$	$1.98 \times 10^{-29}$
ENSBTAG00000049416	<i>RAB20</i>	2.654	5.400	$3.32 \times 10^{-70}$	$7.33 \times 10^{-67}$
ENSBTAG00000001292	<i>LTF</i>	2.658	3.642	$9.82 \times 10^{-17}$	$1.10 \times 10^{-15}$
ENSBTAG00000021887	<i>DPYS</i>	2.734	2.126	$1.13 \times 10^{-30}$	$4.97 \times 10^{-29}$
ENSBTAG00000019669	<i>CD163</i>	2.741	8.221	$4.50 \times 10^{-62}$	$3.51 \times 10^{-59}$
ENSBTAG00000046152	<i>MGAM</i>	2.775	5.328	$7.60 \times 10^{-44}$	$1.16 \times 10^{-41}$
ENSBTAG00000013706	<i>MEGF9</i>	2.820	5.633	$4.75 \times 10^{-60}$	$2.52 \times 10^{-57}$
ENSBTAG00000017969	<i>CA4</i>	2.873	3.695	$2.72 \times 10^{-69}$	$5.15 \times 10^{-66}$
ENSBTAG00000015592	<i>GPR84</i>	2.910	3.862	$1.86 \times 10^{-56}$	$6.32 \times 10^{-54}$
ENSBTAG00000017251	<i>SLC26A8</i>	2.925	2.737	$6.38 \times 10^{-57}$	$2.28 \times 10^{-54}$
ENSBTAG00000020406	<i>GPC3</i>	2.945	3.509	$1.20 \times 10^{-34}$	$7.61 \times 10^{-33}$
ENSBTAG00000018280	<i>SLC28A3</i>	2.975	6.619	$4.28 \times 10^{-45}$	$7.27 \times 10^{-43}$
ENSBTAG00000012640	<i>S100A8</i>	3.012	10.038	$5.43 \times 10^{-58}$	$2.06 \times 10^{-55}$
ENSBTAG00000020580	<i>TCNI</i>	3.017	8.629	$2.11 \times 10^{-58}$	$8.74 \times 10^{-56}$
ENSBTAG00000006505	<i>S100A9</i>	3.030	11.263	$1.04 \times 10^{-60}$	$6.26 \times 10^{-58}$
ENSBTAG00000031950	<i>RAB3IP</i>	3.040	6.690	$1.64 \times 10^{-75}$	$1.09 \times 10^{-71}$
ENSBTAG00000019330	<i>PROK2</i>	3.058	4.710	$1.28 \times 10^{-62}$	$1.06 \times 10^{-59}$
ENSBTAG00000002233	<i>CPNE2</i>	3.130	4.908	$1.61 \times 10^{-58}$	$6.88 \times 10^{-56}$
ENSBTAG00000021240	<i>DCSTAMP</i>	3.132	4.527	$4.67 \times 10^{-47}$	$9.23 \times 10^{-45}$
ENSBTAG00000006354	<i>HP</i>	3.149	9.301	$3.26 \times 10^{-48}$	$7.32 \times 10^{-46}$
ENSBTAG00000006221	<i>ADGRG3</i>	3.160	5.550	$5.38 \times 10^{-68}$	$7.92 \times 10^{-65}$
ENSBTAG00000006999	<i>RYR1</i>	3.309	5.942	$1.85 \times 10^{-60}$	$1.02 \times 10^{-57}$
ENSBTAG00000007239	<i>TNFAIP6</i>	3.348	3.882	$3.97 \times 10^{-55}$	$1.20 \times 10^{-52}$
ENSBTAG00000020676	<i>MMP9</i>	3.434	5.948	$1.48 \times 10^{-40}$	$1.66 \times 10^{-38}$
ENSBTAG00000014149	<i>LCN2</i>	3.446	6.952	$8.55 \times 10^{-44}$	$1.29 \times 10^{-41}$
ENSBTAG00000007901	<i>ADGRE1</i>	3.454	8.940	$4.73 \times 10^{-68}$	$7.83 \times 10^{-65}$
ENSBTAG00000000377	<i>BMX</i>	3.491	5.409	$8.54 \times 10^{-75}$	$3.77 \times 10^{-71}$
ENSBTAG00000002996	<i>SHROOM4</i>	3.491	3.225	$2.38 \times 10^{-65}$	$2.63 \times 10^{-62}$
ENSBTAG00000053557	<i>DEFB4A</i>	3.640	3.012	$2.60 \times 10^{-28}$	$9.14 \times 10^{-27}$
ENSBTAG00000009773	<i>KREMEN1</i>	3.696	6.824	$8.66 \times 10^{-64}$	$8.20 \times 10^{-61}$
ENSBTAG00000049808	<i>IL3RA</i>	3.894	6.830	$4.98 \times 10^{-85}$	$6.60 \times 10^{-81}$
ENSBTAG00000048720		4.248	2.863	$5.64 \times 10^{-24}$	$1.32 \times 10^{-22}$
ENSBTAG00000008951	<i>ALPL</i>	4.327	7.425	$1.10 \times 10^{-61}$	$7.29 \times 10^{-59}$
ENSBTAG00000046158	<i>CFB</i>	4.360	5.400	$3.19 \times 10^{-66}$	$3.84 \times 10^{-63}$
ENSBTAG00000050618		4.403	4.647	$2.75 \times 10^{-58}$	$1.10 \times 10^{-55}$
ENSBTAG00000019627	<i>THY1</i>	4.777	3.655	$2.17 \times 10^{-55}$	$6.85 \times 10^{-53}$
ENSBTAG00000052012		4.788	5.793	$9.54 \times 10^{-37}$	$7.57 \times 10^{-35}$
ENSBTAG00000010273	<i>EREG</i>	4.817	2.213	$5.84 \times 10^{-62}$	$4.30 \times 10^{-59}$
ENSBTAG00000054882		4.990	4.305	$3.10 \times 10^{-38}$	$2.85 \times 10^{-36}$
ENSBTAG00000051132		5.114	5.241	$1.08 \times 10^{-42}$	$1.51 \times 10^{-40}$
ENSBTAG00000039037	<i>SERPIN4</i>	5.203	6.364	$3.02 \times 10^{-51}$	$8.00 \times 10^{-49}$

(Continued on following page)

TABLE 2 (Continued) Differentially expressed genes between BRD and non-BRD animals.

Gene ID	Gene name	Log2 (fold change)	Log counts per million	p-value	FDR
ENSBTAG00000048835		5.317	7.656	$8.80 \times 10^{-48}$	$1.91 \times 10^{-45}$
ENSBTAG00000049569		5.372	6.973	$1.99 \times 10^{-49}$	$4.88 \times 10^{-47}$
ENSBTAG00000006343	<i>IL1R2</i>	5.595	7.513	$4.97 \times 10^{-56}$	$1.61 \times 10^{-53}$
ENSBTAG00000013356	<i>CATHL3</i>	6.087	4.373	$5.49 \times 10^{-33}$	$2.97 \times 10^{-31}$
ENSBTAG00000031647	<i>LRG1</i>	7.348	4.231	$2.37 \times 10^{-74}$	$7.85 \times 10^{-71}$

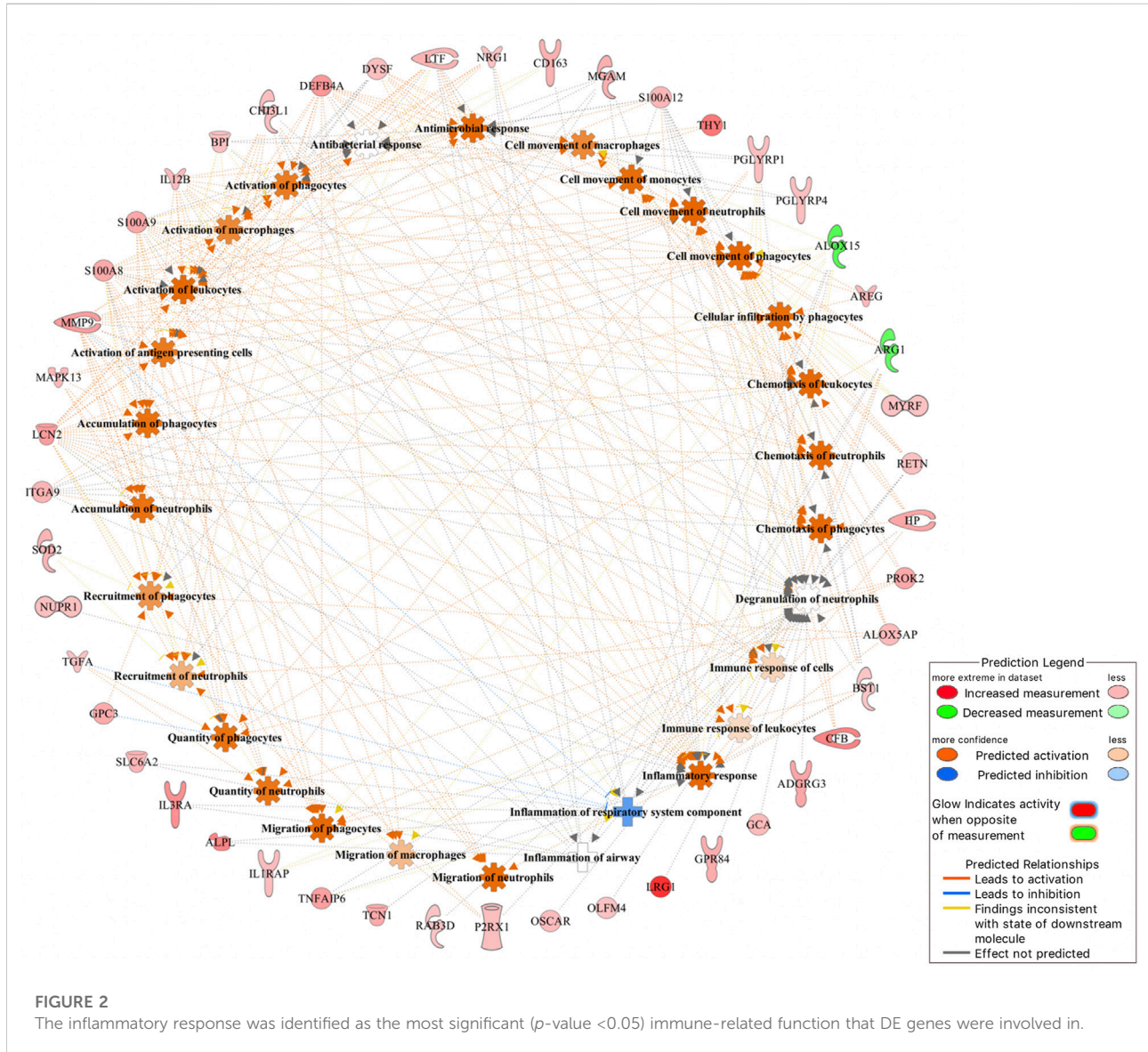
with more than 10 metabolites associated with BRD. Seventy percent of significant correlations between DE genes and metabolites were negative. In other words, in general, BRD animals had higher expression of DE genes and lower concentrations of metabolites in blood tissues. However, there are some interesting exceptions, such as D-mannose, L-phenylalanine, and L-carnitine, which were positively correlated with DE genes and also more abundant in BRD animals.

## Discussion

The BRD and non-BRD animals used in the current study were fed and raised in the same feedlots under similar management and environmental factors. It is therefore expected that all animals in the study were equally exposed to BRD causing pathogens, hence, all BRD animals are assumed to be susceptible while non-BRD animals are resistant. Disease susceptibility and resistance were defined in relation to BRD

TABLE 3 Ten topmost significantly enriched biological functions associated with differentially expressed genes.

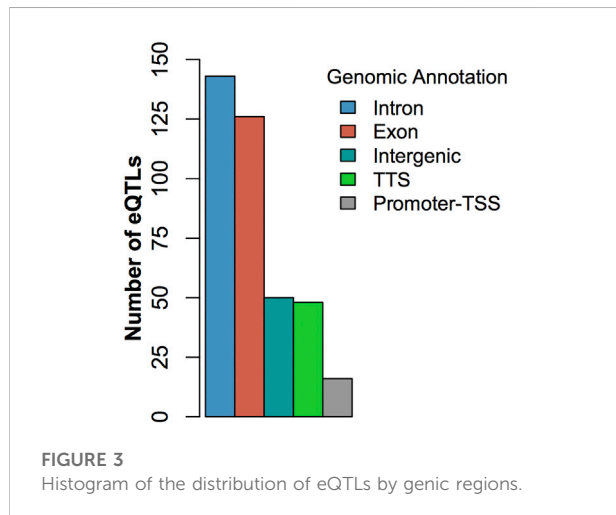
Biological function	p-value range	Genes involved in the biological function
Inflammatory Response	$3.22 \times 10^{-20}$ – $2.06 \times 10^{-3}$	<i>ADGRE1, ADGRG3, ALAS2, ALOX15, ALOX5AP, ALPL, AREG, ARG1, BMX, BPI, BST1, CA4, CD163, CFB, CHI3L1, DEFB4A/DEFB4B, DPYS, DYSF, EREG, GCA, GPC3, GPR84, HP, HTRA1, IL12B, IL1R2, IL1RAP, IL3RA, ITGA9, LCN2, LRG1, LTF, MAPK13, MGAM, MMP9, MYRF, NRG1, NUPR1, OLFM4, OSCAR, P2RX1, PGLYRP1, PGLYRP4, PROK2, RAB3D, RETN, S100A12, S100A8, S100A9, SERPINB4, SLC39A8, SLC6A2, SOD2, TCN1, TGFA, TGM3, THY1, TNFAIP6, TRPC5, UPP1</i>
Connective Tissue Disorders	$1.55 \times 10^{-14}$ – $1.62 \times 10^{-3}$	<i>ALAS2, ALOX15, ALOX5AP, ALPL, AREG, ARG1, BMX, BPI, CA4, CD163, CFB, CHI3L1, DCSTAMP, DPYS, GCA, GPC3, HP, HTRA1, IL12B, IL1R2, IL3RA, ITGA9, KREMEN1, LCN2, LTF, MMP9, PGLYRP1, PROK2, RETN, S100A12, S100A8, S100A9, SLC39A8, SLC6A2, SOD2, TGFA, TNFAIP6</i>
Inflammatory Disease	$1.55 \times 10^{-14}$ – $1.83 \times 10^{-3}$	<i>ADGRE1, ALAS2, ALOX15, ALOX5AP, ALPL, AREG, ARG1, BMX, BPI, CA4, CD163, CFB, CHI3L1, DEFB4A/DEFB4B, DPYS, EREG, GCA, H1-2, HP, HTRA1, IL12B, IL1R2, IL3RA, ITGA9, LCN2, LRG1, LTF, MGAM, MMP9, NRG1, OLFM4, PGLYRP1, PGLYRP4, PROK2, RETN, S100A12, S100A8, S100A9, SERPINB4, SLC39A8, SLC6A2, SOD2, TCN1, TGFA, TGM3, THY1, TNFAIP6</i>
Organismal Injury and Abnormalities	$1.55 \times 10^{-14}$ – $2.07 \times 10^{-3}$	<i>ABCA6, ACVR1B, ADGRE1, ADGRG3, ALAS2, ALOX15, ALOX5AP, ALPL, ANKRD22, AREG, ARG1, ART5, BMX, BPI, BST1, CA4, CD163, CFB, CHI3L1, CPNE2, DCSTAMP, DEFB4A/DEFB4B, DPYS, DYSF, EREG, FOXRED1, GCA, GLT1D1, GPC3, GPR84, H1-2, HBD, HP, HTRA1, IL12B, IL1R2, IL1RAP, IL3RA, ITGA9, KREMEN1, LCN2, LRG1, LTF, MAPK13, MEGF9, MGAM, MMP9, MN1, MYBPH, MYRF, NOL3, NRG1, NUPR1, OLFM4, OSCAR, P2RX1, PGLYRP1, PGLYRP4, PROK2, PTPN5, RAB20, RAB3D, RAB3IP, RBPMS2, REEP1, RETN, RYR1, S100A12, S100A8, S100A9, SERPINB4, SHROOM4, SLC26A8, SLC28A3, SLC39A8, SLC6A2, SOD2, TCN1, TGFA, TGM1, TGM3, THY1, TNFAIP6, TNS2, TRPC5, UPP1, WIPI1</i>
Immunological Disease	$3.58 \times 10^{-11}$ – $2.05 \times 10^{-3}$	<i>ADGRG3, ALAS2, ALOX15, ALOX5AP, ALPL, AREG, ARG1, BMX, BPI, CD163, CFB, CHI3L1, DEFB4A/DEFB4B, GCA, GPC3, GPR84, HP, IL12B, IL1R2, IL3RA, ITGA9, LCN2, LTF, MGAM, MMP9, NRG1, PGLYRP1, PROK2, RETN, S100A12, S100A8, S100A9, SERPINB4, SLC6A2, SOD2, TGFA, TGM3, TNFAIP6</i>
Infectious Diseases	$1.62 \times 10^{-8}$ – $1.42 \times 10^{-3}$	<i>ALOX5AP, ALPL, BPI, CD163, CFB, DEFB4A/DEFB4B, DYSF, GCA, GPC3, H1-2, HP, IL12B, IL1R2, IL3RA, LCN2, LTF, MGAM, MMP9, MYRF, NRG1, OLFM4, P2RX1, PGLYRP1, RAB3D, RETN, S100A12, S100A8, S100A9, SLC6A2, TCN1</i>
Respiratory Disease	$1.62 \times 10^{-8}$ – $1.73 \times 10^{-3}$	<i>ABCA6, ACVR1B, ALAS2, ALOX15, ALPL, ANKRD22, AREG, ARG1, BMX, BPI, BST1, CA4, CD163, CFB, CHI3L1, CPNE2, DCSTAMP, DPYS, DYSF, EREG, FOXRED1, GLT1D1, GPC3, GPR84, H1-2, HP, HTRA1, IL12B, IL1RAP, IL3RA, ITGA9, LCN2, LTF, MAPK13, MEGF9, MGAM, MMP9, MN1, MYBPH, MYRF, NOL3, NRG1, NUPR1, OLFM4, PGLYRP1, PGLYRP4, PTPN5, RETN, RYR1, S100A12, S100A8, S100A9, SERPINB4, SHROOM4, SLC6A2, SOD2, TCN1, TGFA, TGM3, THY1, TNFAIP6, TNS2, TRPC5</i>
Antimicrobial Response	$2.01 \times 10^{-8}$ – $2.57 \times 10^{-4}$	<i>BPI, DEFB4A/DEFB4B, IL12B, LCN2, LTF, PGLYRP1, PGLYRP4, S100A12, S100A8, S100A9</i>
Psychological Disorders	$2.55 \times 10^{-8}$ – $1.83 \times 10^{-3}$	<i>ALOX15, ARG1, CA4, CFB, CHI3L1, DYSF, HP, HTRA1, IL12B, IL1R2, LCN2, LRG1, LTF, MMP9, NRG1, PTPN5, RYR1, S100A9, SLC6A2, SOD2, TGM1, THY1, UPP1</i>
Metabolic Disease	$1.81 \times 10^{-7}$ – $7.5 \times 10^{-4}$	<i>ALOX15, ALOX5AP, ALPL, ARG1, BPI, CA4, CFB, CHI3L1, DYSF, GPC3, HBD, HP, HTRA1, IL12B, IL1R2, IL3RA, LCN2, LTF, MGAM, MMP9, PTPN5, RETN, S100A8, S100A9, SLC6A2, SOD2, TGM1, THY1</i>



in general as a multifactorial and multi-pathogen disease, and not according to specific pathogens. Based on this assumption, two SNPs (Chr5:25858264 and BovineHD1800016801) associated with BRD susceptibility in beef cattle were identified through GWAS (Table 1). The most significant SNP (Chr5:25858264) explained 17% of the phenotypic variance for BRD susceptibility. This implies that this SNP could be a major quantitative trait nucleotide or in linkage disequilibrium with a major QTL for BRD susceptibility in the studied population. However, the proportion of phenotypic variance explained by significant SNPs in the current study might have been overestimated because of the limited number of animals used. In addition, the low coverage depth of SNP calling and low minor allele frequency of Chr5:25858264 may have also led to a false-positive result. Thus, future research utilizing a larger sample size and

higher whole genome sequencing depth are warranted to provide more power for fine mapping this QTL and identifying the causal gene and mutation for BRD resistance.

In addition, results from our study revealed substantial expression differences of 101 genes in the blood tissue of BRD and healthy animals. About 93% of these DE genes were upregulated in the BRD animals (Table 2). Among these upregulated genes, *IL3RA* and *LRG1* showed the strongest association with BRD in terms of statistical significance and fold change, respectively. *IL3RA* encodes the protein of interleukin 3 receptor subunit alpha which is a cytokine receptor protein for interleukin 3 (IL3), colony stimulating factor 2 (CSF2/GM-CSF) and interleukin 5 (IL5) (Milatovich et al., 1993). The cytokine IL3 is generated from T-cells and stem cells, and is involved in macrophage activation and regulation of



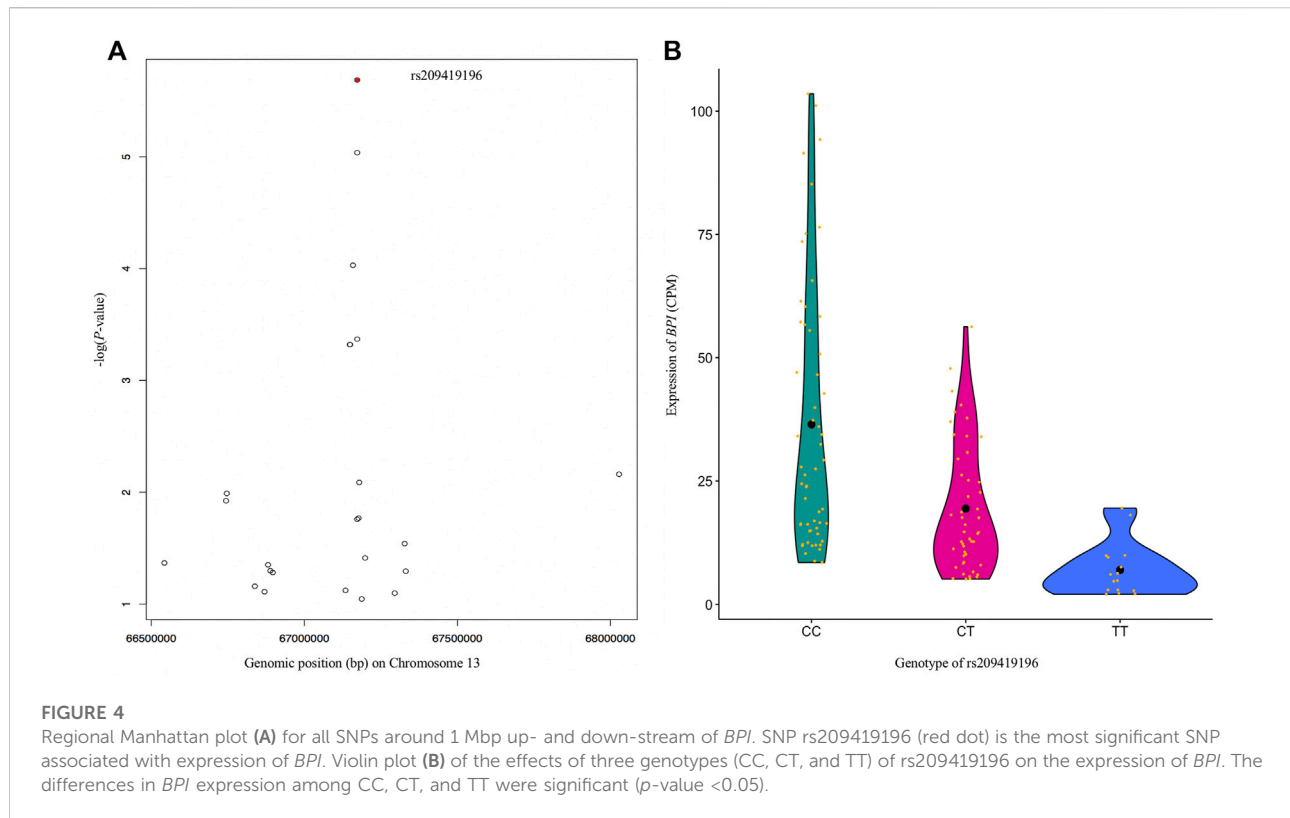
**TABLE 4** Significant metabolites between BRD and non-BRD animals.

Metabolite	t-value	p-value	FDR
Citric acid	-20.226	$5.71 \times 10^{-39}$	$3.08 \times 10^{-37}$
D-Mannose	12.979	$1.28 \times 10^{-21}$	$3.45 \times 10^{-20}$
Acetic acid	-12.277	$7.88 \times 10^{-20}$	$1.42 \times 10^{-18}$
Isobutyric acid	-11.291	$2.62 \times 10^{-17}$	$3.54 \times 10^{-16}$
L-Phenylalanine	97.038	$2.89 \times 10^{-14}$	$3.12 \times 10^{-12}$
3-Hydroxybutyric acid	-94.871	$1.01 \times 10^{-12}$	$9.12 \times 10^{-12}$
L-Glutamine	-89.717	$1.97 \times 10^{-11}$	$1.52 \times 10^{-10}$
L-Glutamic acid	-86.547	$1.20 \times 10^{-10}$	$8.08 \times 10^{-10}$
Tyrosine	-7.910	$7.71 \times 10^{-10}$	$4.63 \times 10^{-8}$
L-Carnitine	78.243	$1.24 \times 10^{-9}$	$6.67 \times 10^{-8}$
Betaine	-71.983	$3.65 \times 10^{-7}$	$1.79 \times 10^{-6}$
L-Threonine	-62.151	$5.78 \times 10^{-5}$	$2.60 \times 10^{-4}$
Isopropanol	38.856	$1.58 \times 10^{-4}$	$4.27 \times 10^{-4}$
Malonate	-36.215	$4.11 \times 10^{-4}$	$1.06 \times 10^{-3}$
Glycine	-35.156	$5.95 \times 10^{-4}$	$1.46 \times 10^{-3}$
L-Acetylcarnitine	57.995	$4.39 \times 10^{-4}$	$1.82 \times 10^{-3}$
L-Serine	-32.186	$1.61 \times 10^{-3}$	$3.78 \times 10^{-3}$
Methionine	-2.899	$4.36 \times 10^{-3}$	$9.81 \times 10^{-3}$
L-Proline	-5.404	$2.81 \times 10^{-3}$	$1.08 \times 10^{-2}$
Dimethylglycine	27.494	$6.78 \times 10^{-3}$	$1.42 \times 10^{-2}$
Choline	-27.472	$6.82 \times 10^{-3}$	$1.42 \times 10^{-2}$
2-Hydroxyisovalerate	27.148	$7.48 \times 10^{-3}$	$1.50 \times 10^{-2}$
L-Aspartate	-26.911	$8.01 \times 10^{-3}$	$1.54 \times 10^{-2}$
L-Lactic acid	25.391	$1.22 \times 10^{-2}$	$2.28 \times 10^{-2}$
Valine	-24.949	$1.38 \times 10^{-2}$	$2.48 \times 10^{-2}$
L-Leucine	24.787	$1.44 \times 10^{-2}$	$2.51 \times 10^{-2}$
Acetone	23.818	$1.86 \times 10^{-2}$	$3.14 \times 10^{-2}$
L-Lysine	22.934	$2.34 \times 10^{-2}$	$3.82 \times 10^{-2}$
Urea	-22.647	$2.51 \times 10^{-2}$	$3.99 \times 10^{-2}$
L-Asparagine	-50.932	$1.14 \times 10^{-2}$	$4.11 \times 10^{-2}$
Sarcosine	-21.973	$2.97 \times 10^{-2}$	$4.58 \times 10^{-2}$

cytokine production (Frendl, 1992). On the other hand, IL-5 is produced by CD4<sup>+</sup> T-cells and causes B-cell growth factor and differentiation, IgA selection, eosinophil activation, and increased production of innate immune cells (Akdis et al., 2011). This study also identified DE genes (*IL1R2*, *IL1RAP*, and *IL12B*) related to interleukin-1 (IL-1) and interleukin-12 (IL-12) which cause lymphocyte activation, macrophage stimulation, increased leukocyte adhesion and release of acute phase proteins by the liver, or induced interferon gamma production by T-cells and natural killer cells (Arena et al., 1998; Akdis et al., 2011; Dinarello, 2018; Jiang et al., 2018). It is worth noting that *IL3RA* and *LRG1* have been reported to be associated with BRD in previous transcriptomic studies (Tizioto et al., 2015; Scott et al., 2020; Jiminez et al., 2021). *LRG1* encodes the protein of leucine rich alpha-2-glycoprotein one that has been reported to be packaged into the granule compartment of human neutrophils and secreted upon neutrophil activation (Druhan et al., 2017). For downregulated genes, the top 3 genes (*HBA1*, *HBA*, and *HBB*) are all related to hemoglobin—the oxygen-carrying protein within red blood cells. Specifically, *HBA1* and *HBA* encode for  $\alpha$ -globin, and *HBB* encodes  $\beta$ -globin, which are the two main globins that compose hemoglobin (Marengo-Rowe, 2006). Thus, the low expressed level of *HBA1*, *HBA*, and *HBB* along with inflammation may indicate anemia of inflammation in infected cattle (BRD susceptible cattle). Additionally, anemia of inflammation could cause normal or sometimes increased amount of iron stored in tissues, but a low level of iron in blood (Nemeth and Ganz, 2014; Fraenkel, 2017). Iron homeostasis is involved in oxygen transport, cellular respiration, and metabolic processes (Ali et al., 2017). The regulation of iron concentration in blood also plays an important role in modulating bacterial infection and contributes to the progression of lung disease (Roehrig et al., 2007; Ali et al., 2017), which could be one of the factors associated with animal susceptibility to BRD. Future studies should determine the relationship between iron levels and susceptibility to BRD in feedlot animals to investigate this hypothesis.

Some of the DE genes identified in the current study have been identified as associated with BRD in beef cattle in other similar studies investigating the lymph node tissue (Tizioto et al., 2015) bronchial epithelial cells (N'jai et al., 2013), and blood (Scott et al., 2020; Jiminez et al., 2021). For example, compared with the results of Tizioto et al. (2015), 26, 35, 29, 39, 20, and 8 of DE genes identified in this study were common with those identified in the lymph node of animals who were challenged by BRSV, IBR, BVDV, *M. haemolytica*, *P. multocida*, and *M. bovis*, respectively (overlapping genes are shown in the Supplementary Table S6). In addition to identifying DE genes specific to individual challenges, Tizioto et al. (2015) found 25 genes expressed differentially in all the infections, of which 5 genes (*S100A8*, *S100A9*, *MMP9*, *TGM3*, and *PGLYRP1*) were also identified as DE genes in the current study. These genes may be differentially expressed in all pathogen challenges because they





are related to innate immune cells. For example, *S100A8* and *S100A9* are expressed in neutrophils and monocytes (Edgeworth et al., 1991) and are known danger-associated molecular patterns that activate innate immune response by binding to pattern recognition receptors of the innate immune cells in response to pathogenic attacks (Schiopu and Cotoi, 2013). Additionally, N'jai et al. (2013) reported the top 70 DE genes identified in bovine bronchial epithelial cells, and three of these genes (*CA4*, *TNFAIP6*, and *HP*) were identified in our current study, as well as previous studies (Tizioto et al., 2015; Jiminez et al., 2021). Comparing our results with DE genes identified in blood samples from other studies (Scott et al., 2020; Jiminez et al., 2021), more common DE genes, such as *LRG1*, *CFB*, and *ALOX15*, were observed. This reveals that the DE genes associated with BRD in different populations are relatively consistent. Furthermore, BRD is a polymicrobial disease that is usually the result of co-infection of several common viral and bacterial pathogens (Dabo et al., 2008; Rice et al., 2008; Griffin et al., 2010; Klima et al., 2014). The infection of different pathogens may cause different immune responses and result in related gene expression in the host (N'jai et al., 2013; Tizioto et al., 2015). When comparing the results of this study to those of Tizioto et al. (2015), the infection process in our population seems to involve multiple pathogens as well. However, the expression of some genes is associated with more than one pathogen and some genes are expressed in response to all

pathogen infections (N'jai et al., 2013; Tizioto et al., 2015) makes it difficult to distinguish specific pathogen infections based on gene expression alone. In this study, the study design and the objectives were to determine common immune responses to BRD infection and to identify DE genes that could be used in different populations and feedlots. Future studies to evaluate the influence of specific pathogens on gene expression in blood are recommended. This may help identify pathogen-specific DE genes for better control and treatment of BRD.

Investigation into the biological involvement of the DE genes revealed inflammatory response as the most significant enriched function. In animals, inflammatory response is a biological response of the immune system to injurious stimuli, such as pathogen presence, damaged cells and toxic compounds (Ferrero-Miliani et al., 2007; Medzhitov, 2010). This response is aimed at clearing the immune insulting agents and initiating healing (Ferrero-Miliani et al., 2007; Medzhitov, 2010). Upon recognition of the pathogenic agents, the immune system responds to such attack by recruiting and activating the phagocytic cells such as macrophages and neutrophils, and those phagocytes that are tasked with the immediate destruction and clearing of the pathogenic agents from the body (Ackermann et al., 2010; Mantovani et al., 2011). Interestingly, activation and recruitment of both neutrophils and macrophages were among the processes identified as enriched within the inflammatory response in the current



study (Figure 2). These processes were predicted to be activated in the BRD animals compared to the non-BRD animals, indicating that the inflammatory response plays a key role in the defense against BRD pathogenic infection. Previous transcriptome studies of blood and other immune organs also demonstrated the significant association of the inflammatory response with BRD status (N'jai et al., 2013; Scott et al., 2020). Some of the interesting inflammatory response genes involved in multiple innate immune responses include *LCN2*, *S100A8*, *S100A9*, *S100A12*, *LTF*, *IL12B*, *CHI3L1*, *DEFB4A*, and *MMP9*. In line with our results and our interpretation of these results, Tizoto et al. (2015) reported DE genes and pathways that were found to be common to all pathogen challenges that were upregulated (e.g., *S100A8*, *S100A9*, and *MMP9*) in the challenged (BRD) animals and appear to primarily be related to the innate immune response. Additionally, we also observed that some genes, such as *LCN2* and *LTF* were predicted to be involved in the cell-mediated immune response, indicating they may be key genes to against viral infections. Therefore, functional enrichment analyses of DE genes provided insights into the biological background of BRD infection and host immune response.

As BRD is caused by multiple viral and bacterial pathogens, the identification of DE genes associated with immune responses to pathogens was in line with our expectations. To study the associations between DNA markers and gene expression, and potential complex regulations of gene expression, we performed eQTL analysis. The information obtained from eQTL analysis could help to understand the GWAS results and illustrate the causality between the significant SNP and BRD susceptibility. For example, the SNP (Chr5:25858264) with the lowest *p*-value in this GWAS showed a cis-effect on the DE gene *GPR84* (Supplementary Table S2). The expression of *GPR84* was mainly observed in bone marrow, lung, and peripheral blood leukocytes (Yousefi et al., 2001), and identified in cells of both the innate and adaptive immune system, which plays a role in pro-inflammatory responses, e.g., cytokine production (Alvarez-Curto and Milligan, 2016; Zhang et al., 2016). The SNP Chr5:25858264 is located downstream of *GPR84*, we speculate the genomic region spanning the SNP might be hosting an enhancer for *GPR84* and hence modulate its expression. However, further experimental evidence is needed to support this speculation. The essential role of eQTL in explaining genetic variance and the shaping of beef cattle phenotypes has been previously reported (Xiang et al., 2022). Additionally, these results obtained from eQTL analysis could help to pinpoint causal SNPs associated with susceptibility to BRD. For example, Neibergs et al. (2014b) reported a genomic region covering *BPI* as associated with BRD susceptibility in Holstein calves, thus indicating that variants within or near this gene have functional relevance in modulating susceptibility to BRD in cattle. *BPI* was also a DE gene associated with BRD in both the current and previous studies (Tizoto et al., 2015; Jiminez et al., 2021). *BPI* encodes the

bactericidal permeability increasing protein, a critical protein involved in neutralizing Gram-negative bacteria lipopolysaccharide antigen while mediating and promoting Gram-negative bacteria recognition by monocytes for phagocytosis (Yu and Song, 2020). Through the eQTL analysis, we further identified another likely causal SNP among all variants within or near the gene *BPI*. The SNP (rs209419196) was the most significant SNP (*p*-value <  $2.1 \times 10^{-6}$ , FDR < 0.006) among six cis-eQTLs associated with the expression of *BPI* (Figure 4A), and the expression of *BPI* was significantly (*p*-value < 0.05) decreased as the number of "T" alleles increased in the genotype (Figure 4B). According to the eQTL annotation analysis, rs209419196 was predicted to be in the promoter region of *BPI*, which is located 92 bp downstream of the 5' end of the transcription start site for the transcript (ENSBTAT00000077785) of *BPI*. Therefore, the results for eQTLs and their annotation not only provide important reference information for GWAS interpretation and causal SNP identification, but also provide additional insights into potential molecular mechanisms of differential gene expression related to disease state. Furthermore, the identification of these SNP markers provides more functional information that can be utilized to enhance genomic selection for BRD resistance in beef cattle.

As BRD is a complex pathogenic interaction with multiple etiologies and risk factors, it is difficult to control and prevent. Conventionally, BRD diagnosis is based on clinical signs, and varies among environment, calf caretakers, producers, and herd veterinarians, often causing a high proportion of false-negative and false-positive diagnoses (Moisá et al., 2019). Such diagnostic inaccuracies lead to the progression of disease, misuse of antimicrobials, production losses, and suboptimal animal welfare outcomes (Moisá et al., 2019). Therefore, accurate diagnostic methods for BRD are still needed. Blood transcriptomic and metabolomic biomarkers have been proposed to be used in the identification of BRD cattle in feedlots (Blakebrough-Hall et al., 2020; Sun et al., 2020). In this study, 101 DE genes were identified and the most informative marker *LRG1* has been previously identified as a potential biomarker for different infections (e.g., active tuberculosis) in humans (Wu et al., 2015; Fujimoto et al., 2020; Yang et al., 2020; Ma et al., 2021). We also found 31 metabolites were significantly associated with BRD infection, 13 of which were consistent with significant metabolites (e.g., citric acid, D-mannose, and acetic acid) identified by Blakebrough-Hall et al. (2020), even though the metabolite profiles used in the two studies were not exactly the same. In addition, 17 significant metabolites were significantly correlated with DE genes and the majority of correlations were negative (Supplementary Table S5), suggesting that generally increased expression of DE genes could result in decreased concentration/abundance of BRD-associated metabolites in blood tissues of BRD animals, thereby negatively impacting

the metabolism of sick animals. Therefore, we suggest that combining such transcriptomic and metabolomic signatures may be useful for BRD identification in feedlots. However, validation in other independent beef cattle populations is required before evaluating their performance and practicality.

## Conclusion

Genomic, transcriptomic and metabolomic data were applied here to elucidate the genetic and molecular background of BRD infection in feedlot beef cattle. Two SNPs associated with BRD susceptibility were identified through GWAS. Transcriptomic and functional analyses revealed 101 DE genes associated with BRD infection. These genes were mainly involved in inflammatory response processes such as recruitment and activation of phagocytes. The most significant SNP (Chr5: 25858264) from the GWAS analysis was also a cis-eQTL associated with a DE gene *GPR84*. This indicates that our integrative analyses could help with the refining of GWAS results and the identification of causal SNPs associated with BRD susceptibility. Additionally, we found 31 metabolites associated with BRD infection and 17 of them were correlated with many DE genes, which indicated the potential biological connections between DE genes and metabolites. Overall, this preliminary multi-omics study illustrates the complex relationships among different omics, which could improve the understanding of genetic and molecular mechanisms underlying the disease.

## Data availability statement

The RNA data presented in the study are deposited in the NCBI Gene Expression Omnibus (GEO) database, accession number GSE217317. The genotype data, phenotype data, and metabolomic data presented in the study are deposited in the Borealis database, doi:10.5683/SP3/ZETWNY.

## Ethics statement

The animal study was reviewed and approved by University of Calgary Veterinary Sciences Animal Care Committee (AC15-0109).

## Author contributions

GP conceived, designed, and oversaw the omics study. ET designed the animal study and was responsible for clinical examination and sample collection. JL performed data

consolidation and data analyses in this study and wrote the draft of the paper. RM performed SNP annotation analyses and helped with the interpretation of IPA results. JJ performed blood sample processing in the laboratory and helped with DE gene analysis. ZW and EA contributed to the statistical analysis. All authors read, edited, and approved the final manuscript.

## Funding

This publication is the result of research supported by the “Genomic Approaches to the Control of Bovine Respiratory Disease Complex (BRDC)” Grant, provided by Genome Alberta, Calgary, AB, Canada, through the Applied Livestock Genomics Program 2014, project code A222. This research is also part of the AMR–One Health Consortium, funded by the Major Innovation Fund Program of the Ministry of Jobs, Economy, and Innovation (JEI), Government of Alberta, Canada, grant number RCP-19-003-MIF.

## Acknowledgments

The authors would like to acknowledge the cooperation of the feedlots and their staff in the performance of the research. Thanks staff at TMIC and Neogen. Special thanks to Cedric Gondro for providing valuable suggestions and comments on this study.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1046192/full#supplementary-material>

## References

- Ackermann, M. R., Derscheid, R., and Roth, J. A. (2010). Innate immunology of bovine respiratory disease. *Vet. Clin. North Am. Food Anim. Pract.* 26, 215–228. doi:10.1016/J.CVFA.2010.03.001
- Akdís, M., Burgler, S., Cramer, R., Eiwegger, T., Fujita, H., Gomez, E., et al. (2011). Interleukins, from 1 to 37, and interferon- $\gamma$ : Receptors, functions, and roles in diseases. *J. Allergy Clin. Immunol.* 127, 701–721. e70. doi:10.1016/j.jaci.2010.11.050
- Alexander, D. H., and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinforma.* 12, 246. doi:10.1186/1471-2105-12-246
- Ali, M. K., Kim, R. Y., Karim, R., Mayall, J. R., Martin, K. L., Shahandeh, A., et al. (2017). Role of iron in the pathogenesis of respiratory disease. *Int. J. Biochem. Cell Biol.* 88, 181–195. doi:10.1016/J.BIOCEL.2017.05.003
- Alvarez-Curto, E., and Milligan, G. (2016). Metabolism meets immunity: The role of free fatty acid receptors in the immune system. *Biochem. Pharmacol.* 114, 3–13. doi:10.1016/J.BCP.2016.03.017
- Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Arena, W. P., Malyak, M., Guthridge, C. J., and Gabay, C. (1998). Interleukin-1 receptor antagonist: Role in biology. *Annu. Rev. Immunol.* 16, 27–55. doi:10.1146/annurev.immunol.16.1.27
- Baptista, A. L., Rezende, A. L., Fonseca, P. de A., Massi, R. P., Nogueira, G. M., Magalhães, L. Q., et al. (2017). Bovine respiratory disease complex associated mortality and morbidity rates in feedlot cattle from southeastern Brazil. *J. Infect. Dev. Ctries.* 11, 791–799. doi:10.3855/jidc.9296
- Blakebrough-Hall, C., Dona, A., D'occhio, M. J., McMeniman, J., and González, L. A. (2020). Diagnosis of Bovine Respiratory Disease in feedlot cattle using blood 1H NMR metabolomics. *Sci. Rep.* 10, 115. doi:10.1038/s41598-019-56809-w
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170
- Brouard, J. S., Schenkel, F., Marete, A., and Bissonnette, N. (2019). The GATK joint genotyping workflow is appropriate for calling variants in RNA-seq experiments. *J. Anim. Sci. Biotechnol.* 10, 44. doi:10.1186/s40104-019-0359-0
- Buczinski, S., and Pardon, B. (2020). Bovine respiratory disease diagnosis: What progress has been made in clinical diagnosis? *Vet. Clin. North Am. Food Anim. Pract.* 36, 399–423. doi:10.1016/J.CVFA.2020.03.004
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* 4, 7. doi:10.1186/S13742-015-0047-8
- Cookson, W., Liang, L., Abecasis, G., Moffatt, M., and Lathrop, M. (2009). Mapping complex disease traits with global gene expression. *Nat. Rev. Genet.* 10, 184–194. doi:10.1038/nrg2537
- Costa-Silva, J., Domingues, D., and Lopes, F. M. (2017). RNA-Seq differential expression analysis: An extended review and a software tool. *PLoS One* 12, e0190152. doi:10.1371/JOURNAL.PONE.0190152
- Dabo, S. M., Taylor, J. D., and Confer, A. W. (2008). *Pasteurella multocida* and bovine respiratory disease. *Anim. Health Res. Rev.* 8, 129–150. doi:10.1017/S1466252307001399
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/BIOINFORMATICS/BTR330
- Dinarello, C. A. (2018). Overview of the IL-1 family in innate inflammation and acquired immunity. *Immunol. Rev.* 281, 8–27. doi:10.1111/IMR.12621
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). Star: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635
- Druhan, L. J., Lance, A., Li, S., Price, A. E., Emerson, J. T., Baxter, S. A., et al. (2017). Leucine rich  $\alpha$ -2 glycoprotein: A novel neutrophil granule protein and modulator of myelopoiesis. *PLoS One* 12, e0170261. doi:10.1371/JOURNAL.PONE.0170261
- Edgeworth, J., Gorman, M., Bennett, R., Freemont, P., and Hogg, N. (1991). Identification of p8, 14 as a highly abundant heterodimeric calcium binding protein complex of myeloid cells. *J. Biol. Chem.* 266, 7706–7713. doi:10.1016/S0021-9258(20)89506-4
- Ferrero-Miliani, L., Nielsen, O. H., Andersen, P. S., and Girardin, S. E. (2007). Chronic inflammation: Importance of NOD2 and NALP3 in interleukin-1 $\beta$  generation. *Clin. Exp. Immunol.* 147, 227–235. doi:10.1111/j.1365-2249.2006.03261.x
- Fraenkel, P. G. (2017). Anemia of inflammation: A review. *Med. Clin. North Am.* 101, 285–296. doi:10.1016/J.MCNA.2016.09.005
- Frendl, G. (1992). Interleukin 3: From colony-stimulating factor to pluripotent immunoregulatory cytokine. *Int. J. Immunopharmacol.* 14, 421–430. doi:10.1016/0192-0561(92)90172-H
- Fujimoto, M., Matsumoto, T., Serada, S., Tsujimura, Y., Hashimoto, S., Yasutomi, Y., et al. (2020). Leucine-rich alpha 2 glycoprotein is a new marker for active disease of tuberculosis. *Sci. Rep.* 10, 3384. doi:10.1038/s41598-020-60450-3
- Griffin, D., Chengappa, M. M., Kuszak, J., and McVey, D. S. (2010). Bacterial pathogens of the bovine respiratory disease complex. *Vet. Clin. North Am. Food Anim. Pract.* 26, 381–394. doi:10.1016/j.cvfa.2010.04.004
- Griffin, D. (1997). Economic impact associated with respiratory disease in beef cattle. *Vet. Clin. North Am. Food Anim. Pract.* 13, 367–377. doi:10.1016/S0749-0720(15)30302-9
- Hoff, J. L., Decker, J. E., Schnabel, R. D., Seabury, C. M., Neibergs, H. L., and Taylor, J. F. (2019). QTL-mapping and genomic prediction for bovine respiratory disease in U.S. Holsteins using sequence imputation and feature selection. *BMC Genomics* 20, 555. doi:10.1186/s12864-019-5941-5
- Hrdlickova, R., Toloue, M., and Tian, B. (2017). RNA-Seq methods for transcriptome analysis. *WIREs RNA* 8, e1364. doi:10.1002/WRNA.1364
- Irsik, M., Langemeier, M., Schroeder, T., Spire, M., and Deen Roder, J. (2006). Estimating the effects of animal health on the performance of feedlot cattle. *Bov. Pract.* 40, 65–74. doi:10.21423/bovine-vol40no2p65-74
- Ives, S. E., and Richeson, J. T. (2015). Use of antimicrobial metaphylaxis for the control of bovine respiratory disease in high-risk cattle. *Vet. Clin. North Am. Food Anim. Pract.* 31, 341–350. doi:10.1016/j.cvfa.2015.05.008
- Jiang, T., Huang, Z., Zhang, S., Zou, W., Xiang, L., Wu, X., et al. (2018). miR-23b inhibits proliferation of SMMC-7721 cells by directly targeting IL-11. *Mol. Med. Rep.* 18, 1591–1599. doi:10.3892/MMR.2018.9151
- Jimenez, J., Timsit, E., Orsel, K., van der Meer, F., Guan, L. L., and Plastow, G. (2021). Whole-blood transcriptome analysis of feedlot cattle with and without bovine respiratory disease. *Front. Genet.* 12, 627623. doi:10.3389/fgene.2021.627623
- Kirchhoff, J., Uhlenbruck, S., Goris, K., Keil, G. M., and Herrler, G. (2014). Three viruses of the bovine respiratory disease complex apply different strategies to initiate infection. *Vet. Res.* 45, 20. doi:10.1186/1297-9716-45-20
- Klima, C. L., Zaheer, R., Cook, S. R., Booker, C. W., Hendrick, S., Alexander, T. W., et al. (2014). Pathogens of bovine respiratory disease in North American feedlots conferring multidrug resistance via integrative conjugative elements. *J. Clin. Microbiol.* 52, 438–448. doi:10.1128/JCM.02485-13
- Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi:10.1093/BIOINFORMATICS/BTT656
- Ma, H., Lu, F., Guo, Y., Shen, Z., and Chen, Y. (2021). The prognostic value of leucine-rich  $\alpha$  2 glycoprotein 1 in pediatric spinal cord injury. *Biomed. Res. Int.* 2021, 7365204. doi:10.1155/2021/7365204
- Mantovani, A., Cassatella, M. A., Costantini, C., and Jaillon, S. (2011). Neutrophils in the activation and regulation of innate and adaptive immunity. *Nat. Rev. Immunol.* 11, 519–531. doi:10.1038/nri3024
- Marengo-Rowe, A. J. (2006). Structure-function relations of human hemoglobins. *Proc. 19, 239–245.* doi:10.1080/08998280.2006.11928171
- McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. doi:10.1093/NAR/GKS042
- Medzhitov, R. (2010). Inflammation 2010: New adventures of an old flame. *Cell.* 140, 771–776. doi:10.1016/j.cell.2010.03.006
- Milatovich, A., Kitamura, T., Miyajima, A., and Francke, U. (1993). Gene for the alpha-subunit of the human interleukin-3 receptor (IL3RA) localized to the X-Y pseudoautosomal region. *Am. J. Hum. Genet.* 53, 1146.
- Moisá, S. J., Aly, S. S., Lehenbauer, T. W., Love, W. J., Rossitto, P. V., Eenennaam, A. L. Van, et al. (2019). Association of plasma haptoglobin concentration and other biomarkers with bovine respiratory disease status in pre-weaned dairy calves. *J. Vet. Diagn. Investig.* 31, 40–46. doi:10.1177/1040638718807242
- Neibergs, H. L., Neibergs, J. S., Wojtowicz, A. J., Taylor, J. F., Seabury, C. M., and Womack, J. E. (2014a). Economic benefits of using genetic selection to reduce the prevalence of bovine respiratory disease complex in beef feedlot cattle. *Proc. Beef Improv. Fed. (BIF) Annu. Meet. Res. Symposium*, 82
- Neibergs, H. L., Seabury, C. M., Wojtowicz, A. J., Wang, Z., Scraggs, E., Kiser, J. N., et al. (2014b). Susceptibility loci revealed for bovine respiratory disease

- complex in pre-weaned holstein calves. *BMC Genomics* 15, 1164. doi:10.1186/1471-2164-15-1164
- Nemeth, E., and Ganz, T. (2014). Anemia of inflammation. *Hematol. Oncol. Clin. North Am.* 28, 671–681. doi:10.1016/j.HOC.2014.04.005
- N'jai, A. U., Rivera, J., Atapattu, D. N., Owusu-Ofori, K., and Czuprynski, C. J. (2013). Gene expression profiling of bovine bronchial epithelial cells exposed *in vitro* to bovine herpesvirus 1 and Mannheimia haemolytica. *Vet. Immunol. Immunopathol.* 155, 182–189. doi:10.1016/j.vetimm.2013.06.012
- Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Auwera, G. A. Van der, et al. (2017). *Scaling accurate genetic variant discovery to tens of thousands of samples*. bioRxiv. doi:10.1101/201178
- Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., et al. (2011). The human serum metabolome. *PLoS One* 6, e16957. doi:10.1371/JOURNAL.PONE.0016957
- Rice, J. A., Carrasco-Medina, L., Hodgins, D. C., and Shewen, P. E. (2008). Mannheimia haemolytica and bovine respiratory disease. *Anim. Health Res. Rev.* 8, 117–128. doi:10.1017/S1466252307001375
- Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25. doi:10.1186/GB-2010-11-3-R25
- Roehrig, S. C., Tran, H. Q., Spehr, V., Gunkel, N., Selzer, P. M., and Ullrich, H. J. (2007). The response of Mannheimia haemolytica to iron limitation: Implications for the acquisition of iron in the bovine lung. *Vet. Microbiol.* 121, 316–329. doi:10.1016/J.VETMIC.2006.12.013
- Saude, E. J., Slupsky, C. M., and Sykes, B. D. (2006). Optimization of NMR analysis of biological fluids for quantitative accuracy. *Metabolomics* 2, 113–123. doi:10.1007/s11306-006-0023-5
- Schiopu, A., and Cotoi, O. S. (2013). S100A8 and S100A9: DAMPs at the crossroads between innate immunity, traditional risk factors, and cardiovascular disease. *Mediat. Inflamm.* 2013, 828354. doi:10.1155/2013/828354
- Schneider, M. J., Tait, R. G., Ruble, M. V., Busby, W. D., and Reecy, J. M. (2010). Evaluation of fixed sources of variation and estimation of genetic parameters for incidence of bovine respiratory disease in preweaned calves and feedlot cattle. *J. Anim. Sci.* 88, 1220–1228. doi:10.2527/jas.2008-1755
- Scott, M. A., Woolums, A. R., Swiderski, C. E., Perkins, A. D., Nanduri, B., Smith, D. R., et al. (2020). Whole blood transcriptomic analysis of beef cattle at arrival identifies potential predictive molecules and mechanisms that indicate animals that naturally resist bovine respiratory disease. *PLoS One* 15, e0227507. doi:10.1371/JOURNAL.PONE.0227507
- Shabalina, A. A. (2012). Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 28, 1353–1358. doi:10.1093/BIOINFORMATICS/BTS163
- Smith, R. A. (2000). Effects of feedlot disease on economics, production and carcass value. *Am. Assoc. Bovine Pract. Proc.*, 125–128. doi:10.21423/aabppro20005374
- Smith, R. A. (1998). Impact of disease on feedlot performance: A review. *J. Anim. Sci.* 76, 272–274. doi:10.2527/1998.761272X
- Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., and Bennett, G. L. (2005). Influence of breed, heterozygosity, and disease incidence on estimates of variance components of respiratory disease in preweaned beef calves. *J. Anim. Sci.* 83, 1247–1261. doi:10.2527/2005.8361247x
- Stanford, K., Zaheer, R., Klima, C., McAllister, T., Peters, D., Niu, Y. D., et al. (2020). Antimicrobial resistance in members of the bacterial bovine respiratory disease complex isolated from lung tissue of cattle mortalities managed with or without the use of antimicrobials. *Microorganisms* 8, 288. doi:10.3390/microorganisms8020288
- Sun, H. Z., Srithayakumar, V., Jimenez, J., Jin, W., Hosseini, A., Raszek, M., et al. (2020). Longitudinal blood transcriptomic analysis to identify molecular regulatory patterns of bovine respiratory disease in beef cattle. *Genomics* 112, 3968–3977. doi:10.1016/j.YGENO.2020.07.014
- Taylor, J. D., Fulton, R. W., Lehenbauer, T. W., Step, D. L., and Confer, A. W. (2010). The epidemiology of bovine respiratory disease: What is the evidence for predisposing factors? *Can. Vet. J.* 51, 1095.
- Timsit, E., Bareille, N., Seegers, H., Lehebel, A., and Assié, S. (2011). Visually undetected fever episodes in newly received beef bulls at a fattening operation: Occurrence, duration, and impact on performance. *J. Anim. Sci.* 89, 4272–4280. doi:10.2527/JAS.2011-3892
- Tizioto, P. C., Kim, J. W., Seabury, C. M., Schnabel, R. D., Gershwin, L. J., Van Eenennaam, A. L., et al. (2015). Immunological response to single pathogen challenge with agents of the bovine respiratory disease complex: An RNA-sequence analysis of the bronchial lymph node transcriptome. *PLoS One* 10, e0131459. doi:10.1371/journal.pone.0131459
- Turner, S. D. (2014). *qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots*. bioRxiv, 005165. doi:10.1101/005165
- Van der Auwera, G. A., and O'Connor, B. D. (2020). *Genomics in the cloud: Using docker, GATK, and WDL in terra*. 1st Edition. Sebastapol, California: O'Reilly Media, Inc.
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi:10.3168/jds.2007-0980
- Wu, J., Yin, H., Zhu, J., Buckanovich, R. J., Thorpe, J. D., Dai, J., et al. (2015). Validation of LRG1 as a potential biomarker for detection of epithelial ovarian cancer by a blinded study. *PLoS One* 10, e0121112. doi:10.1371/JOURNAL.PONE.0121112
- Xia, J., Broadhurst, D. I., Wilson, M., and Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics* 9, 280–299. doi:10.1007/s11306-012-0482-9
- Xiang, R., Fang, L., Liu, S., Macleod, I. M., Liu, Z., Breen, E. J., et al. (2022). Gene expression and RNA splicing explain large proportions of the heritability for complex traits in cattle. *bioRxiv*. doi:10.1101/2022.05.30.494093
- Yang, F. J., Hsieh, C. Y., Shu, K. H., Chen, I. Y., Pan, S. Y., Chuang, Y. F., et al. (2020). Plasma leucine-rich  $\alpha$ -2-glycoprotein 1 predicts cardiovascular disease risk in end-stage renal disease. *Sci. Rep.* 10, 5988. doi:10.1038/s41598-020-62989-7
- Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). Gcta: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi:10.1016/j.ajhg.2010.11.011
- Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., and Price, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* 46, 100–106. doi:10.1038/ng.2876
- Yousefi, S., Cooper, P. R., Potter, S. L., Mueck, B., and Jarai, G. (2001). Cloning and expression analysis of a novel G-protein-coupled receptor selectively expressed on granulocytes. *J. Leukoc. Biol.* 69, 1045–1052. doi:10.1189/JLB.69.6.1045
- Yu, Y., and Song, G. (2020). "Lipopolysaccharide-binding protein and bactericidal/permeability-increasing protein in lipid metabolism and cardiovascular diseases," in *Lipid transfer in lipoprotein metabolism and cardiovascular disease*. Editor X. Jiang (Singapore: Springer), 27–35. doi:10.1007/978-981-15-6082-8\_3
- Zhang, Q., Yang, H., Li, J., and Xie, X. (2016). Discovery and characterization of a novel small-molecule agonist for medium-chain free fatty acid receptor G protein-coupled receptor 84. *J. Pharmacol. Exp. Ther.* 357, 337–344. doi:10.1124/JPET.116.232033