

Use of nCounter mRNA Profiling to Identify at-arrival Gene Expression Patterns for Predicting Bovine Respiratory Disease in Beef Cattle

Matthew Scott (✉ matthewscott@tamu.edu)

Texas A&M University and West Texas A&M University

Amelia Woolums

Mississippi State University

Cyprianna Swiderski

Mississippi State University

Alexis Thompson

Mississippi State University

Andy Perkins

Mississippi State University

Bindu Nanduri

Mississippi State University

Brandi Karisch

Mississippi State University

Dan Goehl

Professional Beef Services, LLC

Research Article

Keywords: bovine respiratory disease, NanoString, mRNA, gene expression, biomarkers, disease prediction, infectious disease, host immunology, beef cattle, bioinformatics

Posted Date: September 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-923361/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Veterinary Research on February 23rd, 2022. See the published version at <https://doi.org/10.1186/s12917-022-03178-8>.

Abstract

Background

Transcriptomics has identified at-arrival differentially expressed genes associated with bovine respiratory disease (BRD) development; however, their use as prediction molecules necessitates further evaluation. Therefore, we aimed to selectively analyze and corroborate at-arrival mRNA expression from multiple independent populations of beef cattle. In a nested case-control study, we evaluated the expression of 56 mRNA molecules from at-arrival blood samples of 234 cattle across seven populations via NanoString nCounter gene expression profiling. Analysis of mRNA was performed with nSolver Advanced Analysis software ($p < 0.05$), comparing cattle groups based on the diagnosis of clinical BRD within 28 days of facility arrival ($n=115$ Healthy; $n=119$ BRD); BRD was further stratified for severity based on frequency of treatment and/or mortality (Treated_1, $n=89$; Treated_2+, $n=30$). Gene expression homogeneity of variance, receiver operator characteristic (ROC) curve, and decision tree analyses were performed between severity cohorts.

Results

Increased expression of mRNAs involved in specialized pro-resolving mediator synthesis (*ALOX15*, *HPGD*), leukocyte differentiation (*LOC100297044*, *GCSAML*, *KLF17*), and antimicrobial peptide production (*CATHL3*, *GZMB*, *LTF*) were identified in Healthy cattle. BRD cattle possessed increased expression of *CFB*, and mRNA related to granulocytic processes (*DSG1*, *LRG1*, *MCF2L*) and type-I interferon activity (*HERC6*, *IFI6*, *ISG15*, *MX1*). Healthy and Treated_1 cattle were similar in terms of gene expression, while Treated_2+ cattle were the most distinct. ROC cutoffs were used to generate an at-arrival treatment decision tree, which classified 90% of Treated_2+ individuals.

Conclusions

Increased expression of complement factor B, pro-inflammatory, and type I interferon-associated mRNA hallmark the at-arrival expression patterns of cattle that develop severe clinical BRD. Here, we corroborate at-arrival mRNA markers identified in previous transcriptome studies and generate a prediction model to be evaluated in future studies. Further research is necessary to evaluate these expression patterns in a prospective manner.

Background

Bovine respiratory disease (BRD) is the leading clinical disease in post-weaned beef production systems throughout North America [1, 2, 3]. BRD is a multifactorial disease that culminates from complex interactions between viral and bacterial etiologies, the host, and adverse environmental circumstances, such as novel housing and feeding situations [3, 4, 5, 6, 7]. While novel diagnostic and therapeutic programs have been developed over the past decade, rates of morbidity and mortality related to BRD remains relatively unchanged [2, 8, 9, 10]. Ultimately, antemortem diagnostics of BRD is imprecise, and no

gold standard modality currently exists for determining real-time therapeutic intervention [11, 12, 13, 14]. Subsequently, at-arrival mass medication of cattle with antimicrobials (e.g., metaphylaxis) is commonly employed to improve cattle herd health and long-term economic performance [15, 16, 17]. Metaphylaxis is a proven technique in reducing morbidity and mortality among high-risk beef cattle, but its usage is highly variable when evaluated across populations and may influence the development of multidrug antimicrobial resistance [17, 18, 19, 20]. Therefore, the evaluation of gene expression and associated mechanisms of beef cattle at arrival may lead to the development of rapid diagnostics which accurately predict BRD development and severity. The development of a predictive test could allow for better assessment of BRD risk and reduce overall antimicrobial usage.

RNA sequencing (RNA-Seq) has become an invaluable tool for evaluated genomic and gene-by-environment associations related to clinical disease [21, 22, 23]. Previous studies have detected differentially expressed genes and enriched mechanisms associated with peak clinical disease after experimental challenge with known BRD-associated pathogens and related to naturally occurring disease [24, 25, 26, 27, 28, 29]. Additionally, our previous research has defined genes and pathways related to pro-inflammatory mediation/resolution and antiviral defense to be increased at arrival in high-risk beef cattle that resist or succumb to naturally occurring BRD, respectively [30, 31, 32]. These studies have provided a better understanding of the host-disease complex relationship and created a defined record of target molecules involved in BRD, but their potential use as gene expression markers requires substantiation through extensive testing for classification accuracy and mechanistic delineation.

Candidate biomarker development remains a leading area of medical research, as technologies continue to improve, and the understanding of patient heterogeneity and dynamic changes associated with pathology are primary focuses of therapeutic advancement. Modern panels are developed in a multi-marker approach, as single-molecule diagnostics lack clinical significance and pathological representation [33]. Furthermore, single-molecule diagnostics often fail to define individuals within subgroups or cohorts of disease, as the intricacies of pathological severity and temporal association are frequently missed [33, 34]. Thus, we developed a multi-marker mRNA expression panel to evaluate previously identified differentially expressed genes and enriched mechanisms in high-risk post-weaned beef cattle at facility arrival. We hypothesized and corroborated that specific genes related to host immune and inflammatory responses are differentially expressed in equivocal manner to previous BRD host transcriptome studies. Here, we provide mRNA expressional information that represents BRD acquisition and severity across multiple populations of cattle, serving as a platform for the development of predictive diagnostics at arrival.

Methods

Animal use and enrollment

All animal use and procedures were approved by the Mississippi State University Animal Care and Use Committee (IACUC protocols #15 – 003 and #18–529) and carried out in accordance with relevant IACUC

and agency guidelines and regulations. The information reported here is in accordance with Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines (<https://arriveguidelines.org>). Whole blood samples were acquired from 397 cattle at facility arrival, spanning seven independent populations; five populations were cattle purchased from commercial livestock auctions within the state of Mississippi or neighboring states and housed at the H. H. Leveck Animal Research Center at Mississippi State University (Starkville, MS, USA; VD_15, n = 14; VD_17, n = 71; PS_19, n = 72; MH_19, n = 83; TA_20, n = 84) and two populations were cattle purchased from commercial livestock auctions within the state of Missouri or neighboring states and housed at the Professional Beef Services Research Facility in northeast Missouri (Canton, MO, USA; DG_17, n = 33; DG_18, n = 40). On day 0 for each population, blood samples were collected into RNA preservation tubes (Tempus Blood RNA Tubes, Thermo Fisher Scientific, Waltham, Massachusetts, USA) via jugular venipuncture and stored at -80° C until analysis. All cattle were given identification ear tags and confirmed to be negative for persistent infection with bovine viral diarrhea virus (BVDV) via ear notch antigen capture ELISA. Due to this study being performed in conjunction with seven independent experiments, metadata collection (average daily weight gain, pen assignment, vaccination/anthelmintic administration, etc.) varied by population.

Cattle within each population were assessed daily for clinical signs of BRD by trained staff. Cattle identified with clinical BRD were assigned a severity score of 0–4, adapted from the scoring system described by Holland and colleagues [35]. Necessary antimicrobial therapy for cattle housed in Mississippi was instituted as previously described [20]. Necessary antimicrobial therapy for cattle housed in Missouri followed a regimen of gamithromycin (Zactran; Boehringer-Ingelheim Animal Health USA, Duluth, GA, USA) and vitamin C, enrofloxacin (Baytril, Bayer Animal Health, Shawnee Mission, KS, USA) and vitamin C, or florfenicol with flunixin meglumine (Resflor Gold, Merck Animal Health, Madison, NJ, USA); all drugs were given at labeled dosage and route of administration. After the third antimicrobial treatment cattle with signs of predetermined endpoints indicative of unlikely recovery were evaluated by project veterinarians, and euthanized via intravenous administration of pentobarbital if necessary. All cattle that died or were euthanized were subjected to necropsy to identify grossly visible abnormalities of lungs and other organ systems. Clinical endpoints included severe dyspnea, dull mentation, and/or weakness leading to inability to take in feed and water adequate to sustain life.

***A priori* and *post-hoc* power calculations**

A priori and *post hoc* power analyses were performed with G*Power v3.1.9.7 [36] for sample size estimation to assess Healthy versus BRD gene expression and retrospective power estimation for the constructed decision tree model. Means and standard deviation observed between Δ Ct values generated from RT-qPCR analysis of *ALOX15* [30] was utilized for *a priori* power calculation, accounting for an alpha of 0.05 and power (1- β error probability) of 0.95; power was calculated with a two-tailed t-Test of the mean differences between Healthy and BRD cattle. Two test *post hoc* analysis for the decision tree models utilized the proportions of correctly identified treated (BRD) and non-treated (Healthy) cattle.

Sample selection and RNA processing

A total of 240 samples were selected for RNA isolation and gene expression analysis. 116 samples from cattle never receiving treatment throughout their associated study (Healthy) were randomly selected across all seven populations with the RANDBETWEEN function in Excel 2016 (Microsoft Corp, Redmond, WA, USA). 124 samples from cattle treated at least once within 28 days of arrival (BRD) were selected and further stratified for severity based on treatment frequency. Of the 124 BRD samples, 33 samples from cattle treated twice or more and/or died (Treated_2+) were available and prioritized for analysis; the remaining 91 BRD samples were randomly selected with the RANDBETWEEN function in Excel 2016 from cattle treated once within 28 days of arrival and never treated again (Treated_1). Additional information regarding cattle selected for gene expression analysis is found in Supplementary Table S1.

RNA from samples collected into RNA preservation tubes (3 mL blood + 6 mL buffer) was isolated with Tempus Spin RNA Isolation kits (Thermo Fisher Scientific, Waltham, Massachusetts, USA) according to manufacturer's instructions. RNA quantification, quality, and integrity was assessed with a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and Agilent 2100 Bioanalyzer system (Agilent Technologies, Santa Clara, CA, USA) at the Emory Integrated Genomics Core (EIGC; Emory University, Atlanta, GA, USA). RNA isolation integrity and concentration values for all samples are found in Supplementary Table S2. Samples having a concentration below 10 ng/ μ L were adjusted to ≥ 10 ng/ μ L with a SpeedVac vacuum concentrator (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

Gene selection and NanoString nCounter assay

A custom NanoString nCounter (NanoString Technologies, Seattle, WA, USA) mRNA probe set was created to analyze the expression of 56 mRNA molecules of interest, and also four housekeeping genes (Supplementary Table S3). All genes selected were identified as differentially expressed in our previous studies and have known biological relevance related to immune, metabolic, and/or inflammatory systems [30, 31, 32]. Additionally, the panel included six positive and eight negative spike-in controls implemented by EIGC. Samples were prepared onto a 96-well plate and analyzed with the nCounter SPRINT Profiler (NanoString Technologies, Seattle, WA, USA), according to manufacturer's protocol. Briefly, samples were hybridized to target-specific reporter and capture probes overnight at 65°C in a thermocycler. Following hybridization, 30–35 μ L of each hybridized product was pipetted to the nCounter cartridge, sealed, and loaded into the drawer of the SPRINT Profiler. A maximum field of view (FOV) sensitivity of 555 was selected for expressional capture.

nCounter analysis

Reporter code counts (RCC) and reporter library file (RLF) obtained from the nCounter SPRINT Profiler were utilized for gene expression analysis through the nSolver Advanced Analysis Software v4.0 (NanoString Technologies, Seattle, WA, USA; <https://www.nanostring.com/products/analysis-solutions/ncounter-analysis-solutions/>). Background correction, quality assessment, and inter-sample normalization was performed in accordance with the Gene Expression Data Analysis Guidelines (MAN-C0011-04). Imaging quality control measurements were set to flag any lane failing to have a registered

FOV above 75%. Binding density was flagged if any lane recorded an optical feature per square micron outside the range of 0.1 to 1.8. Positive control linearity, used to assess the efficiency of hybridization, was flagged if any resulting count was below a correlation of 0.95. Positive control limit of detection flagged any sample if the 0.5 fM positive control counts were ≤ 2 SD above the mean of the negative controls. Background thresholds were set to the mean of the manufactured negative controls plus two standard deviations. Positive control normalization was performed using the geometric mean to compute normalization factors for each sample; lanes with a normalization factor outside of the 0.3–3.0 range were flagged. Codeset content normalization was performed with the geometric mean of the four housekeeping genes, flagging lanes if the normalization factor was outside of the 0.1–10.0 range. In total, six samples (72_VD_15, 118_VD_17, 147_VD_17, 188_VD_17, 143_PS_19, and 205_TA_20) were removed from further analysis due to low quality. Normalized mRNA counts were analyzed for differential expression in four comparative analyses: Healthy versus BRD, Treated_1 versus Healthy, Treated_2 + versus Healthy, and Treated_2 + versus Treated_1. Differential expression analysis was performed with the normalized count data via the nSolver Advanced Analysis module, accounting for study as a potential confounding variable. Differentially expressed genes (DEGs) were determined to be significant with a p-value of ≤ 0.05 . Gene expression data produced in this study were submitted to the National Center for Biotechnology Information Gene Expression Omnibus (NCBI-GEO) under accession GSE179536.

Statistical analyses and data characterization

Due to discrepancies in metadata recording across all seven studies, average daily gain in pounds at the end of each study, when cattle were sold at study conclusion (ADGend), was assessed for distributive differences associated with frequency of treatment. Residual normality and assumption of equal variance was assessed in R v4.0.4 with the Shapiro-Wilk and Bartlett's test, respectively. When accounting for disease severity, the residuals were found to be non-normally distributed. A one-way Kruskal-Wallis test was used to test the relationship between ADGend and severity of disease, blocking for study as a potential confounding variable. Pairwise comparisons were assessed, accounting for familywise error rate with the Bonferroni correction method; any comparison with an adjusted p-value of ≤ 0.05 was considered significant.

Normalized gene expression values were exported from nSolver as a comma separated values file. Counts were imported into R v4.0.4 and converted to log₂ count-per-million (log₂CPM) values for downstream statistical and classificational analyses. Heatmap visualization and unsupervised hierarchical clustering of z-score-scaled log₂CPM gene expression values were performed with the Bioconductor package pheatmap v1.0.12 [37], utilizing Ward's minimum variance method of unsupervised hierarchical clustering on Minkowski distances and Pearson correlation coefficients for samples and DEGs, respectively. Multi-level modeling (MLwiN v2.25, Centre for Multilevel Modelling, University of Bristol, Bristol, EN, UK) was used to assess the proportion of variance at each level of disease status (BRD, Healthy) within study and severity of disease (Healthy, Treated_1, Treated_2+) within study. Variance partition coefficients were calculated to observe the proportion of the response variation that lay within each level of the hierarchy by dividing the variation within a level by the sum of all the

variation within the model. Equal variance between genes by disease status was assessed using the Levene's test for homogeneity of variance in the general linear model procedure of SAS v9.2 (SAS Institute, Cary, NC, USA), in order to evaluate for outlier-driven gene expression. An *a priori* level of significance was set at an alpha of 0.10. Visual assessment of the conditional residuals confirmed the assumption of normal distribution. Remaining data visualization was conducted with ggplot2 v3.3.3 [38] and UpSetR v1.4.0 [39, 40]. Graphical color scaling was performed with viridis v0.6.1 [41] to allow ease of visual interpretation for individuals with color blindness.

Counts from uniquely identified DEGs were utilized for multiclass receiver operating characteristic (ROC) curve analysis, recording aggregate areas under the curve (AUC) and cutoff values used to distinguish disease severity cohorts. ROC curve analyses were conducted with the Bioconductor package pROC v1.17.0 [42]. Classification between the three groups (Healthy, Treated_1, and Treated_2+) was evaluated as "excellent" (AUC: > 0.900), "good" (AUC: 0.899-0.800), "fair" (AUC: 0.799-0.700), "poor" (AUC: 0.699-0.600), or "failed" (AUC: < 0.600). An empirical decision tree model was constructed to identify the maximum predictive accuracy of disease severity classification from DEGs, utilizing log2CPM value thresholds generated from the ROC analysis of Treated_2 + versus Healthy cattle. Correctly identified and misclassified Healthy and Treated_1 cattle from the resulting decision tree model were assessed for differences in ADGend values via two-tailed Student's t-Test with an *a priori* level of significance set to an alpha of 0.10.

Results

A priori power analysis concluded that 210 samples would be required (Healthy, n = 105; BRD, n = 105) to achieve a power level of 0.95. After removal of samples that failed quality control assessment, 115 Healthy and 119 BRD (Treated_1, n = 89; Treated_2+, n = 30) samples remained for differential expression analysis. 17 genes were identified as differentially expressed between Healthy and BRD cattle (Supplementary Table S4). Healthy cattle, when compared to BRD, possessed increased expression of genes associated with anti-inflammatory processes (*CPB2* and *IL5RA*), specialized pro-resolving mediator (SPM) metabolism (*ALOX15* and *HPGD*), lymphocytic maturation (*LOC100297044*, *GCSAML*, and *KLF17*), and antimicrobial peptide production (*CATHL3*, *GZMB*, and *LTF*); BRD cattle possessed increased expression of genes associated with type I interferon production (*HERC6* and *IFI6*), alternative complement (*CFB*), and granulocyte adhesion/cell-cell interaction (*DSG1*, *LRG1*, and *MCF2L*). Pairwise analysis of each severity cohort revealed 11, 22, and 16 genes to be differentially expressed between Treated_1 versus Healthy, Treated_2 + versus Healthy, and Treated_2 + versus Treated_1 comparisons, respectively (Supplementary Table S5-S7). *CFB* was decreased in expression in Healthy cattle when independently compared to both Treated_1 and Treated_2 + cattle. Type I interferon-associated genes (*HERC6*, *IFI6*, *ISG15*, and *MX1*) were increased in Treated_2 + cattle compared to Healthy, but only *HERC6* was considered differentially expressed between Treated_2 + and Treated_1 cattle. *ALOX15* and *HPGD* were decreased in Treated_1 cattle compared to Healthy, but were not considered differentially expressed between Treated_2 + and Healthy cattle. In total, 30 genes were uniquely identified as differentially expressed in at least one comparative analysis (Fig. 1).

A heatmap was generated for twelve DEGs that best stratified disease severity, based on hierarchical clustering of gene expression patterns across samples (Fig. 2). Clustering of Pearson correlation coefficients was used to stratify the twelve genes into four expressional arrays, which consisted of type I interferon-associated genes (*HERC6*, *IFI6*, *ISG15*, and *MX1*), complement factor B (*CFB*), SPM and leukocyte-associated genes (*ALOX15*, *HPGD*, and *LOC100297044*), and cell adhesion/lymphocyte-associated genes (*ITGA4*, *GZMB*, *GCSAML*, and *LOC100335828*). BRD cattle, specifically within the Treated_2 + classification, appeared to cluster more to the right side of the heatmap associated with higher levels of type I interferon-associated gene expression and lower levels of SPM and cell adhesion/lymphocyte-associated gene expression. However, Healthy and Treated_1 cattle appeared more similar in expressional patterns compared to Treated_2 + cattle.

ADGend values were used to assess growth performance across the three disease severity cohorts. One-way Kruskal-Wallis testing revealed a significant difference in the distribution of average daily weight gain across the three cohorts ($F = 24.699$, $p < 0.001$). Pairwise comparisons of the three groups demonstrated an overall decrease in weight gain with increased frequency of treatment ($p < 0.001$; Fig. 3).

Multi-level modeling identified the proportion of variance was highest when factoring for disease severity, with an average of 92.50% lying between disease severity cohorts within study (Supplementary Table S8). *IFI6* possessed the highest discrepancy in variance, as 16.97% lay between studies and 83.03% lay between disease cohorts within study. *MGC126945* possessed the highest level of variance between disease cohorts within study (100%). Analysis for homogeneity of variance across severity of disease revealed that unequal variance existed for ten DEGs ($p < 0.10$): *CFB*, *DSG1*, *DSP*, *GYPA*, *HPGD*, *KRT10*, *LOC100297044*, *LRG1*, *MCF2L*, and *MGC126945* (Fig. 4). *CFB*, *GYPA*, *LOC100297044*, *LRG1*, and *MCF2L* demonstrated more homogeneous gene expression within the Treated_2 + cohort compared Healthy and Treated_1 cattle. Notably, *DSG1*, *DSP*, *HPGD*, *KRT10*, and *MGC126945* were more variable in terms of gene expression within and between disease cohorts. Complete results for homogeneity of variance analysis across severity cohorts is found in Supplemental Figure S1 and Supplemental Table S8.

All 30 unique DEGs were selected for multiclass ROC curve evaluation to determine log₂CPM cutoffs and discriminative capability for predicting BRD severity outcomes (Table 1). For univariate ROC analysis, all but three DEGs (*DSG1*, *DSP*, and *KRT10*) best discriminated Treated_2 + versus Healthy cattle, as opposed to Treated_1 versus Healthy or Treated_2 + versus Treated_1. Consequently, these three genes were shown to possess high inter-group variation and outlier-driven means associated with Treated_2+ (Fig. 4; mean versus median values). Eleven independently evaluated DEGs failed to discriminate Treated_2 + and Healthy cattle ($AUC < 0.600$). *LOC100297044* demonstrated good discrimination of Treated_2 + and Healthy cattle and was subsequently the top independent classifier ($AUC: 0.868$). To assess pathway-based expressional classification, two multivariable classifier models were constructed: type I interferon-associated (IFN) with *HERC6*, *IFI6*, *ISG15*, and *MX1* and SPM-associated with *ALOX15* and *HPGD*. IFN demonstrated good discrimination of Treated_2 + from Healthy ($AUC: 0.730$). SPM demonstrated poor discrimination of Treated_2 + from Healthy ($AUC: 0.646$). Multiclass ROC curve analyses by population

for DEGs previously evaluated [32] and/or selected for decision tree construction are found in Supplemental Table S9.

An at-arrival treatment decision tree model was constructed from these DEG classifiers, utilizing log2CPM value expression levels for all individuals (Fig. 5A). *Post-hoc* power analysis of the resulting decision tree indicated a power of 0.94 for determining Healthy and BRD cattle. Overall, the treatment model successfully identified 73.9% (88/119) of all BRD cattle and both predictive values, dependent upon the distribution of Healthy and BRD within this study, were above 61.0% (Fig. 5B). Notably, when stratifying for disease severity, the model accurately identified Treated_2 + and Treated 1 individuals with 90.0% and 68.5% accuracy, respectively (Fig. 5C). Statistical assessment of the Healthy cattle misclassified as needing treatment (n = 56) compared to correctly identified Healthy cattle (n = 59) revealed no difference in ADGend ($p = 0.193$). However, Treated_1 cattle misclassified by the decision tree as not needing treatment (i.e., Healthy) possessed a significantly higher ADGend (n = 28; $\bar{x}=2.091$, $\sigma_x = 0.844$) than those correctly identified (n = 61; $\bar{x}=1.728$, $\sigma_x = 0.850$) ($p = 0.084$).

Discussion

Trained individuals within conventional beef production systems handle cattle at a limited number of time points, predominately at arrival, to reduce animal stress and minimize labor demand. Coupled with the complex nature of BRD and lack of a gold standard diagnostic test, these systems have difficulty determining likelihood of disease acquisition and severity on an intra-population scale. Accordingly, our previous RNA-Seq studies were conducted to better characterize the at-arrival host response in cattle related to subsequent development and treatment of naturally-acquired BRD within the first 28 days of facility placement [30, 31, 32]. RNA-Seq studies employ robust statistical and mathematical models in an effort to identify potential molecular targets and mechanisms that may improve disease understanding, detection, and therapy [43, 45]. Such studies reduce the overall dimensionality of host expression and allow for the determination of a finite number of molecules to assess in future studies. However, the development of disease prediction or prognostic models requires rigorous testing and can be resource intensive, as gene-by-environment interactions and disease temporality are difficult to account for in developing predictive models [46, 47]. Therefore, we focused on corroborating differential expression and retrospectively classifying the predictive capability of mRNA molecules previously identified by RNA-Seq, but through a method feasible for use in large numbers of cattle. To this purpose, we utilized NanoString nCounter mRNA profiling as a less variable and higher throughput method of evaluating specific gene expression, compared to the commonly utilized alternative, RT-qPCR [48, 49, 50, 51, 52]. To our knowledge, this study is the first to substantiate selective at-arrival gene expressional patterns previously associated with BRD acquisition and severity, utilizing blood samples from beef cattle across multiple independent populations.

A limitation of this work is that we aimed to identify at-arrival gene expression profiles that predicted future treatment for BRD based on clinical assessment, which can lack diagnostic sensitivity [13, 14]. While we recognize that the visual identification of clinical BRD is relatively insensitive and potentially

confounded by study location, our statistical analyses and modeling were conducted to account for this effect. Furthermore, analysis of weight gain records at time of sale (ADGend) demonstrated an objective loss in production that was associated with increased frequency of treatment, in agreement with previous research (Fig. 3) [53, 54], confirming the significance of the BRD diagnoses in this study.

In total, 30 unique genes were determined to be differentially expressed across all comparisons (Fig. 1; Supplemental Tables S4-S7). Cattle never diagnosed with clinical BRD possessed increased expression of genes broadly associated with leukocyte/granulocyte development and differentiation (*GCSAML*, *IL5RA*, *KLF17*, *LOC100297044* (CCL14), and *LOC100335828* (CD200R1)), SPMs (*ALOX15* and *HPGD*), anti-inflammatory/apoptotic processes (*CPB2* and *ITGA4*), and antimicrobial peptides (*CATHL3*, *GZMB*, and *LTF*). *CD200R1*, *CCL14*, *GCSAML*, and *IL5RA* have previously been identified as chemokines for macrophage activity and enhanced lymphocyte migration and survival [55, 56, 57, 58, 59]. SPMs, while not widely explored in ungulates, are categorized into two primary categories based on the metabolism of arachidonic acid (lipoxins) or essential polyunsaturated fatty acids (resolvins, maresins, and protectins) [60, 61]. In other mammalian species, SPMs are involved in quelling prolonged pro-inflammatory events and promoting cellular/molecular homeostasis in the lower airways; these molecules are currently under investigation as therapeutic modalities in humans with sepsis or viral-induced pneumonia/acute respiratory distress syndrome (ARDS) [62, 63, 64, 65]. *CPB2* and *ITGA4* are key regulators of pro-inflammation, as they are involved in modulating the complement system, degrading induced plasma anaphylatoxins, and reducing nitric oxide accumulation [66, 67, 68, 69, 70, 71]. Cattle are capable of producing a myriad of well conserved innate peptides within the cathelicidin and defensin peptide families, such as *CATHL3*, *GZMB*, and *LTF*, which have oxygen-independent killing capacity against both Gram-positive and Gram-negative pathogens [72, 73]. These amphophilic molecules align specifically to bacterial cell walls via electrostatic interactions and penetrate into the cytoplasmic space, leading to DNA replication disruption and/or bacterial autolysis [74, 75, 76, 77]. Collectively, these findings indicate that cattle that remain clinically healthy possessed enhanced immunological, anti-inflammatory, and microbial killing mechanisms at arrival, compared to cattle that require antimicrobial therapy within the first 28 days of arrival.

Cattle that would go on to develop BRD possessed increase expression of genes at arrival broadly associated with pro-inflammatory, granulocytic processes (*LRG1*, *MCF2L*, *SPP1*), alternative complement (*CFB*), and type I interferon signaling (*HERC6*, *IFI6*, *ISG15*, and *MX1*), when compared to healthy cattle. Notably, Treated_1 cattle only possessed increased expression of *CFB* and *LRG1* when compared to Healthy cattle. Therefore, these aforementioned genes and their associations are predominantly detected within Treated_2 + cattle. *LRG1* encodes for a leucine-rich glycoprotein specifically stored and secreted by neutrophils, which appears to be marker of early differentiation [78, 79]. *MCF2L* explicitly binds with RAC subfamily of Rho GTPases and may be critical in neutrophilic signal transduction [80, 81]. *SPP1* is a major secreted component of pro-inflammatory leukocytes and enhances cellular infiltration and fibrosis of the airways [82, 83, 84]. *CFB*, increased in both Treated_1 and Treated_2 + cattle when independently compared to Healthy cattle, encodes for the proenzyme of alternative complement. In human sera, *CFB* has been indicated in early, severe infectious disease, such as that induced by severe acute respiratory

syndrome coronavirus 2 (SARS-CoV-2) [85, 86]. Additionally, *CFB* production is exaggerated in human airway epithelial cells following experimental asthma and rhinovirus challenge [87]. *HERC6*, *IFI6*, *ISG15*, and *MX1* encode for proteins induced by type I interferons, such as IFN- α , IFN- β , IFN- ω , and are typically produced by host cells in response to viral infection [88, 89, 90]. Cattle experimentally infected with BRD-associated viruses, such as bovine respiratory syncytial virus (BRSV) and bovine herpesvirus 1 (BHV-1), demonstrate increased expression of these genes during peak clinical presentation, when compared to sham controls [24, 25, 26, 27]. While this appears to be a natural host (bovine) response to viral infection, our previous and current findings, coupled with the work of Sun and colleagues, indicate that anti-viral responses are indicative of early stage and, often, severe naturally-acquired BRD [28, 31]. Previous work indicated strong elevation and co-expression of these genes at arrival in cattle that died of BRD, with predicted interactions with pro-inflammatory cytokine production such as interleukin-6 and tumor necrosis factor- α [31, 91]. Taken together, our findings suggest that cattle that required multiple antimicrobial therapies, and/or died of naturally-acquired BRD, entered the production system with increased anti-viral defense mechanisms without the benefit of activation/elevation of mechanisms for pro-inflammatory regulation or resolution.

To further clarify the relationship between DEGs and disease status, we employed hierarchical clustering and visualization with a gene expression heatmap (Fig. 2). Notably, Healthy and Treated_1 cattle were relatively indistinguishable based on the expression patterns of those 12 DEGs, but Treated_2 + cattle tended to cluster with the association of elevated type I interferon-associated gene expression and decreased SPM and lymphocyte proliferation-associated gene expression (Fig. 2, right side). This finding is similar to results found in our previous multipopulational transcriptome analysis, suggesting that Treated_2 + cattle are more distinct in terms of at-arrival gene expression identification compared to Healthy and Treated_1 cattle [32]. This finding was made more apparent when assessing multilevel modeling and inter-cohort variance, as Treated_2 + cattle tended to be more similar in gene expression compared to Healthy and Treated_1 cohorts, regardless of study-level effect (Figs. 4 and S1). *CFB*, a gene identified with unequal variance and representing both BRD acquisition and severity, was relatively homogeneous in terms of gene expression within Treated_2 + cattle, but disproportionate within the Healthy and Treated_1 cohorts. Outliers within the Healthy and Treated_1 cohorts tended to shift mean gene expression higher than the median and drove differences in variance observed between the three groups. The observed results of *CFB* may suggest that some cattle within the Healthy and Treated_1 cohort had subclinical BRD. Further evaluation of variance determined that *DSG1*, *DSP*, *KRT10*, and *MGC126945* appeared to be outlier-driven, implying unproportionate differential gene expression across populations, and ruling them out as candidate predictive classifiers for BRD.

Multiclass ROC curve analysis was utilized to generate cutoffs and assess classificational capability for each unique DEG; however, this demonstrated that a single-molecule classification model was not sufficient to capture the dynamic expressional patterns observed across all seven populations (Table 1). Notably, our findings here contrasts with our previous multiclass ROC evaluation to classify disease based on at-arrival differential expression of *ALOX15*, *CFB*, *MARCO*, *LOC100335828*, and *SLC18A2* [32]. While we previously found good to excellent discrimination of Healthy versus Treated_2 + cattle based on

the expression levels of single genes in cattle from two independent populations [32], the AUCs resulting from ROC curve analysis in this study were considerably lower when results from all seven populations of cattle were assessed together. For example, in our previous assessment of cattle from two populations, the ROC curve based on differential expression of *ALOX15* to distinguish Treated_2 + from Healthy cattle had an AUC of 0.860 [32], while in this study, the AUC was 0.635 (Table 1). The AUCs resulting from ROC curve analysis for each of the seven populations in this study evaluated individually suggests that these molecules may better discriminate BRD in populations with more severe disease, as seen with the DG_17, DG_18, and MH_19 populations compared to VD_17 and PS_19 (Supplemental Tables S1 and S9). For example, in the DG_17, DG_18, and MH_19 populations, the AUCs for the ROC curves differentiating Treated_2 + from Healthy cattle by differential expression of *ALOX15* were 0.806, 0.760, and 0.800, respectively; compared to VD_17 and PS_19, for which the AUCs were 0.632 and 0.529, respectively (Supplemental Tables S1 and S9).

Corresponding with our differential expression and hierarchical clustering analyses, Treated_2 + cattle were the most dissimilar and routinely possessed the highest AUC values for genes independently evaluated, when compared to Healthy. Given this, we employed a multivariable ROC curve model for the two most explainable mechanisms (IFN and SPM; Table 1) and developed an expressional decision tree model using the ROC-generated cutoffs from Treated_2 + and Healthy comparisons (Fig. 5). While the combination of *ALOX15* and *HPGD* (SPM) made no discernable difference in performance as compared to evaluation of each gene independently, combining the four IFN genes increased the single-test sensitivity and specificity between Treated_2 + and Healthy cattle. The resulting decision tree possessed relatively low accuracy for discerning Healthy from BRD (51.3%), but possessed excellent accuracy in Treated_2 + categorization (90.0%). While Treated_1 individuals are similar to Healthy in terms of at-arrival gene expression, classificational accuracy (68.5%) is moderately higher than the estimated sensitivity of diagnosing clinical BRD via visual assessment alone (22.0–62.0%) [12, 13]. Furthermore, Treated_1 cattle misclassified as not requiring treatment (Healthy) were sold having higher total average daily gains compared to those classified correctly; this may indicate the ability for this decision tree model to better stratify for disease severity compared to visual assessment and treatment frequency associated with BRD.

This report represents our third evaluation of differential gene expression in whole blood at arrival to predict subsequent BRD treatment in cattle, with progressively larger numbers of cattle from more groups evaluated in each study. Genes that were strongly differentially expressed between Healthy and BRD cattle in our earlier work were not always significantly differentially expressed in subsequent studies. Given the heterogeneity of the cattle within and between these populations, and likely variation in the time of sample collection relative to disease onset, perhaps this finding is no surprise. However, some genes, such as *ALOX15* and *CFB*, have been consistently differentially expressed between Healthy and BRD cattle across our studies, suggesting that further evaluation of these molecules and related pathways may improve prediction and understanding of BRD risk. The influence of the imprecise nature of current BRD assessment systems must also be acknowledged, especially when considering individuals without

more definitive clinical endpoints such as mortality. The development of a more accurate method of BRD diagnosis is essential to support efforts to accurately predict the disease.

Conclusions

We sought to evaluate and corroborate the at-arrival expression patterns of mRNA previously identified as differentially expressed between cattle subsequently treated for BRD and cattle not treated. Here, we profiled at-arrival whole blood samples of 234 cattle across seven independent populations via NanoString nCounter and nSolver analysis. At arrival expression levels of mRNA associated with specialized pro-resolving mediator metabolism, anti-inflammatory/apoptotic processes, immune function, and antimicrobial peptide production were increased in cattle never requiring antimicrobial therapy for BRD. Cattle requiring antimicrobial therapy for BRD within 28 days of arrival, especially those needing two or more treatments, possessed increased expression of complement factor B, pro-inflammatory, and type I interferon-associated mRNA. We further evaluated these mRNA expression levels for individual classificational accuracy through receiver operator characteristic curves and the development of a decision tree model. Cattle that would require multiple antimicrobial therapies for BRD were classified with 90.0% accuracy. These findings serve as a foundation for developing an at-arrival mRNA prediction model in high-risk populations of post-weaned beef cattle. As this study assessed at-arrival expression levels, future studies evaluating expression longitudinally are necessary to infer relationships between gene expression and long-term protection against or facilitation of BRD.

Declarations

Ethics approval and consent to participate

The study was conducted according to the Mississippi State University Animal Care and Use Committee (IACUC protocols #15-003, approved January 23, 2015, and #18-529, approved December 4, 2018) and carried out in accordance with relevant IACUC and agency guidelines and regulations.

Consent for publication

Not applicable.

Availability of data and materials

The data presented in this study are openly available in the National Center for Biotechnology Information Gene Expression Omnibus repository (NCBI-GEO; <https://www.ncbi.nlm.nih.gov/geo>) under the accession GSE179536. All remaining data generated or analyzed during this study are included in this published article and its additional information files.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the Animal Health and Disease Program [Grant no. 2020-67016-31469] from the USDA National Institute of Food and Agriculture.

Authors' contributions

Conceptualization: MAS, ARW, CES; Methodology: MAS; Sample and data curation: MAS, ARW, ACT, BBK, DRG; Computational and statistical analysis: MAS, ACT; Animal maintenance and treatment supervision: MAS, ARW, BBK, DRG; Writing – original draft preparation: MAS; Writing – Review & Editing: All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank Elliot Sandhu, Lyra Griffiths, and Emory Integrated Genomic Core (EIGC) staff for their technical assistance and support. The authors additionally thank Marguerite Yelverton, William Crosby, Richard Wagner, Kirsten Midkiff, Merrilee Thoresen, Brianna Furman, Kelsey Shields, and the staff at the Mississippi Agriculture and Forestry Experiment Station (MAFES) and Professional Beef Services, LLC for their assistance in animal care, sample collection, and data preparation.

References

1. USDA. Part IV: Health and Health Management on U.S. Feedlots with a Capacity of 1,000 or More Head. Ft. Collins: USDA-APHIS-VS-CEAH-NAHMS; 2011.
2. Wilson BK, Richards CJ, Step DL, Krehbiel CR. BEEF SPECIES SYMPOSIUM: Best management practices for newly weaned calves for improved health and well-being¹. *Journal of Animal Science*. 2017 May 1;95(5):2170–82.
3. Taylor JD, Fulton RW, Lehenbauer TW, Step DL, Confer AW. The epidemiology of bovine respiratory disease: what is the evidence for preventive measures? *Can Vet J*. 2010 Dec;51(12):1351–9.
4. Grissett GP, White BJ, Larson RL. Structured literature review of responses of cattle to viral and bacterial pathogens causing bovine respiratory disease complex. *J Vet Intern Med*. 2015 May;29(3):770–80.
5. Fulton RW. Viruses in bovine respiratory disease in North America. *Veterinary Clinics of North America: Food Animal Practice*. 2020 Jul;36(2):321–32.
6. Pratelli A, Cirone F, Capozza P, Trotta A, Corrente M, Balestrieri A, et al. Bovine respiratory disease in beef calves supported long transport stress: An epidemiological study and strategies for control and prevention. *Research in Veterinary Science*. 2021 Mar; 135:450–5.
7. Kudirkiene E, Aagaard AK, Schmidt LMB, Pansri P, Krogh KM, Olsen JE. Occurrence of major and minor pathogens in calves diagnosed with bovine respiratory disease. *Veterinary Microbiology*. 2021 Aug;259:109135.

8. Blakebrough-Hall C, Dona A, D'occhio MJ, McMeniman J, González LA. Diagnosis of Bovine Respiratory Disease in feedlot cattle using blood 1H NMR metabolomics. *Sci Rep.* 2020 Dec;10(1):115.
9. Kelly AP, Janzen ED. A review of morbidity and mortality rates and disease occurrence in North American feedlot cattle. *Can Vet J.* 1986 Dec;27(12):496–500.
10. Engler M, Defoor P, King C, Gleghorn J. The impact of bovine respiratory disease: the current feedlot experience. *Anim Health Res Rev.* 2014 Dec;15(2):126–9.
11. Buczinski S, Pardon B. Bovine respiratory disease diagnosis. *Veterinary Clinics of North America: Food Animal Practice.* 2020 Jul;36(2):399–423.
12. White BJ, Renter DG. Bayesian estimation of the performance of using clinical observations and harvest lung lesions for diagnosing bovine respiratory disease in post-weaned beef calves. *J VET Diagn Invest.* 2009 Jul;21(4):446–53.
13. Timsit E, Dendukuri N, Schiller I, Buczinski S. Diagnostic accuracy of clinical illness for bovine respiratory disease (BRD) diagnosis in beef cattle placed in feedlots: A systematic literature review and hierarchical Bayesian latent-class meta-analysis. *Preventive Veterinary Medicine.* 2016 Dec;135:67–73.
14. Caswell JL, Hewson J, Slavić Đ, DeLay J, Bateman K. Laboratory and postmortem diagnosis of bovine respiratory disease. *Veterinary Clinics of North America: Food Animal Practice.* 2012 Nov;28(3):419–41.
15. Dennis EJ, Schroeder TC, Renter DG. Net return distributions when metaphylaxis is used to control bovine respiratory disease in high health-risk cattle. *Translational Animal Science.* 2020 Apr 1;4(2):1091–102.
16. Baptiste KE, Kyvsgaard NC. Do antimicrobial mass medications work? A systematic review and meta-analysis of randomised clinical trials investigating antimicrobial prophylaxis or metaphylaxis against naturally occurring bovine respiratory disease. *Pathogens and Disease.* 2017 July;75(7):ftx083.
17. Ives SE, Richeson JT. Use of antimicrobial metaphylaxis for the control of bovine respiratory disease in high-risk cattle. *Veterinary Clinics of North America: Food Animal Practice.* 2015 Nov;31(3):341–50.
18. Abell KM, Theurer ME, Larson RL, White BJ, Apley M. A mixed treatment comparison meta-analysis of metaphylaxis treatments for bovine respiratory disease in beef cattle. *Journal of Animal Science.* 2017 Feb 1;95(2):626–35.
19. Taylor JD, Fulton RW, Lehenbauer TW, Step DL, Confer AW. The epidemiology of bovine respiratory disease: what is the evidence for preventive measures? *Can Vet J.* 2010 Dec;51(12):1351–9.
20. Woolums AR, Karisch BB, Frye JG, Epperson W, Smith DR, Blanton J, et al. Multidrug resistant *Mannheimia haemolytica* isolated from high-risk beef stocker cattle after antimicrobial metaphylaxis and treatment for bovine respiratory disease. *Veterinary Microbiology.* 2018 Jul;221:143–52.

21. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nat Rev Genet.* 2019 Nov;20(11):631–56.
22. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 2016 Dec;17(1):13.
23. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends in Genetics.* 2014 Sep;30(9):418–26.
24. Tizioto PC, Kim J, Seabury CM, Schnabel RD, Gershwin LJ, Van Eenennaam AL, et al. Immunological response to single pathogen challenge with agents of the bovine respiratory disease complex: an RNA-sequence analysis of the bronchial lymph node transcriptome. Harrod K, editor. *PLoS ONE.* 2015 Jun 29;10(6):e0131459.
25. Behura SK, Tizioto PC, Kim J, Grupioni NV, Seabury CM, Schnabel RD, et al. Tissue tropism in host transcriptional response to members of the bovine respiratory disease complex. *Sci Rep.* 2017 Dec;7(1):17938.
26. Johnston D, Earley B, McCabe MS, Lemon K, Duffy C, McMenamy M, et al. Experimental challenge with bovine respiratory syncytial virus in dairy calves: bronchial lymph node transcriptome response. *Sci Rep.* 2019 Dec;9(1):14736.
27. Johnston D, Earley B, McCabe MS, Kim J, Taylor JF, Lemon K, et al. Messenger RNA biomarkers of Bovine Respiratory Syncytial Virus infection in the whole blood of dairy calves. *Sci Rep.* 2021 Dec;11(1):9392.
28. Sun H-Z, Srithayakumar V, Jiminez J, Jin W, Hosseini A, Raszek M, et al. Longitudinal blood transcriptomic analysis to identify molecular regulatory patterns of bovine respiratory disease in beef cattle. *Genomics.* 2020 Nov;112(6):3968–77.
29. Jiminez J, Timsit E, Orsel K, van der Meer F, Guan LL, Plastow G. Whole-blood transcriptome analysis of feedlot cattle with and without bovine respiratory disease. *Front Genet.* 2021 Mar 8;12:627623.
30. Scott MA, Woolums AR, Swiderski CE, Perkins AD, Nanduri B, Smith DR, et al. Whole blood transcriptomic analysis of beef cattle at arrival identifies potential predictive molecules and mechanisms that indicate animals that naturally resist bovine respiratory disease. Loor JJ, editor. *PLoS ONE.* 2020 Jan 13;15(1):e0227507.
31. Scott MA, Woolums AR, Swiderski CE, Perkins AD, Nanduri B, Smith DR, et al. Comprehensive at-arrival transcriptomic analysis of post-weaned beef cattle uncovers type I interferon and antiviral mechanisms associated with bovine respiratory disease mortality. Ortega-Villaizan M del M, editor. *PLoS ONE.* 2021 Apr 26;16(4):e0250758.
32. Scott M, Woolums A, Swiderski C, Perkins A, Nanduri B, Smith D, et al. Multipopulational transcriptome analysis of post-weaned beef cattle at arrival further validates candidate biomarkers for predicting clinical bovine respiratory disease. In Review; 2021 Jun. Preprint at: <https://www.researchsquare.com/article/rs-600364/v1>.
33. Adamcova M, Šimko F. Multiplex biomarker approach to cardiovascular diseases. *Acta Pharmacol Sin.* 2018 Jul;39(7):1068–72.

34. Mackey HM, Bengtsson T. Sample size and threshold estimation for clinical trials with predictive biomarkers. *Contemporary Clinical Trials*. 2013 Nov;36(2):664–72.
35. Holland BP, Step DL, Burciaga-Robles LO, Fulton RW, Confer AW, Rose TK, et al. Effectiveness of sorting calves with high risk of developing bovine respiratory disease on the basis of serum haptoglobin concentration at the time of arrival at a feedlot. *American Journal of Veterinary Research*. 2011 Oct;72(10):1349–60.
36. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*. 2009 Nov;41(4):1149–60.
37. Kolde R. Pheatmap: pretty heatmaps [Internet]. 2019. Available from: <https://CRAN.R-project.org/package=pheatmap>
38. Wickham H. *Ggplot2: elegant graphics for data analysis*. 2nd ed. 2016. Cham: Springer International Publishing; Imprint: Springer; 2016. 1 p. (Use R!).
39. Lex A, Gehlenborg N, Strobel H, Vuillemot R, Pfister H. Upset: visualization of intersecting sets. *IEEE Trans Visual Comput Graphics*. 2014 Dec 31;20(12):1983–92.
40. Khan A, Mathelier A. Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. *BMC Bioinformatics*. 2017 Dec;18(1):287.
41. Garnier S, Ross N, Rudis B, Sciaini M, Camargo A, Scherer C. Viridis - colorblind-friendly color maps for r [Internet]. 2021. Available from: <https://cran.r-project.org/web/packages/viridis/index.html>
42. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for r and s+ to analyze and compare roc curves. *BMC Bioinformatics*. 2011 Dec;12(1):77.
43. Skala SL, Wang X, Zhang Y, Mannan R, Wang L, Narayanan SP, et al. Next-generation RNA sequencing–based biomarker characterization of chromophobe renal cell carcinoma and related oncocytic neoplasms. *European Urology*. 2020 Jul;78(1):63–74.
44. Kalender Atak Z, Gianfelici V, Hulselmans G, De Keersmaecker K, Devasia AG, Geerdens E, et al. Comprehensive analysis of transcriptome variation uncovers known and novel driver events in T-cell acute lymphoblastic leukemia. Scott HS, editor. *PLoS Genet*. 2013 Dec 19;9(12):e1003997.
45. Kumar-Sinha C, Kalyana-Sundaram S, Chinnaiyan AM. Landscape of gene fusions in epithelial cancers: SEQ and ye shall find. *Genome Med*. 2015 Dec;7(1):129.
46. Mackey HM, Bengtsson T. Sample size and threshold estimation for clinical trials with predictive biomarkers. *Contemporary Clinical Trials*. 2013 Nov;36(2):664–72.
47. Royston P, Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. *Statist Med*. 2004 Aug 30;23(16):2509–25.
48. Goytain A, Ng T. NanoString nCounter technology: high-throughput RNA validation. *Methods Mol Bio*. 2020;2079:125–139.
49. Ebentier DL, Hanley KT, Cao Y, Badgley BD, Boehm AB, Ervin JS, et al. Evaluation of the repeatability and reproducibility of a suite of qPCR-based microbial source tracking methods. *Water Research*.

- 2013 Nov;47(18):6839–48.
50. Dagnall CL, Hicks B, Teshome K, Hutchinson AA, Gadalla SM, Khincha PP, et al. Effect of pre-analytic variables on the reproducibility of qPCR relative telomere length measurement. Criscuolo F, editor. *PLoS ONE*. 2017 Sep 8;12(9):e0184098.
 51. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*. 2008 Mar;26(3):317–25.
 52. Chen Y, Gelfond JA, McManus LM, Shireman PK. Reproducibility of quantitative RT-PCR array in miRNA expression profiling and comparison with microarray analysis. *BMC Genomics*. 2009;10(1):407.
 53. Holland BP, Burciaga-Robles LO, VanOverbeke DL, Shook JN, Step DL, Richards CJ, et al. Effect of bovine respiratory disease during preconditioning on subsequent feedlot performance, carcass characteristics, and beef attributes^{1,2}. *Journal of Animal Science*. 2010 Jul 1;88(7):2486–99.
 54. Wilson BK, Step DL, Maxwell CL, Gifford CA, Richards CJ, Krehbiel CR. Effect of bovine respiratory disease during the receiving period on steer finishing performance, efficiency, carcass characteristics, and lung scores. *The Professional Animal Scientist*. 2017 Feb;33(1):24–36.
 55. Jaguin M, Houlbert N, Fardel O, Lecureur V. Polarization profiles of human M-CSF-generated macrophages and comparison of M1-markers in classically activated macrophages from GM-CSF and M-CSF origin. *Cellular Immunology*. 2013 Jan;281(1):51–61.
 56. Kotarsky K, Sitnik KM, Stenstad H, Kotarsky H, Schmidtchen A, Koslowski M, et al. A novel role for constitutively expressed epithelial-derived chemokines as antibacterial peptides in the intestinal mucosa. *Mucosal Immunol*. 2010 Jan;3(1):40–8.
 57. Shen Y, Iqbal J, Xiao L, Lynch RC, Rosenwald A, Staudt LM, et al. Distinct gene expression profiles in different B-cell compartments in human peripheral lymphoid organs. *BMC Immunol*. 2004;5(1):20.
 58. NCBI Resource Coordinators, Agarwala R, Barrett T, Beck J, Benson DA, Bollin C, et al. Database resources of the national center for biotechnology information. *Nucleic Acids Research*. 2018 Jan 4;46(D1):D8–13.
 59. Buchheit KM, Dwyer DF, Ordovas-Montanes J, Katz HR, Lewis E, Vukovic M, et al. IL-5Ra marks nasal polyp IgG4- and IgE-expressing cells in aspirin-exacerbated respiratory disease. *Journal of Allergy and Clinical Immunology*. 2020 Jun;145(6):1574–84.
 60. Chiang N, Serhan CN. Structural elucidation and physiologic functions of specialized pro-resolving mediators and their receptors. *Molecular Aspects of Medicine*. 2017 Dec;58:114–29.
 61. Lee CH. Role of specialized pro-resolving lipid mediators and their receptors in virus infection: a promising therapeutic strategy for SARS-CoV-2 cytokine storm. *Arch Pharm Res*. 2021 Jan;44(1):84–98.
 62. Duvall MG, Bruggemann TR, Levy BD. Bronchoprotective mechanisms for specialized pro-resolving mediators in the resolution of lung inflammation. *Molecular Aspects of Medicine*. 2017 Dec;58:44–56.

63. Buechler C, Pohl R, Aslanidis C. Pro-resolving molecules—new approaches to treat sepsis? *IJMS*. 2017 Feb 23;18(3):476.
64. Andreakos E, Papadaki M, Serhan CN. Dexamethasone, pro-resolving lipid mediators and resolution of inflammation in COVID-19. *Allergy*. 2021 Mar;76(3):626–8.
65. Panigrahy D, Gilligan MM, Huang S, Gartung A, Cortés-Puch I, Sime PJ, et al. Inflammation resolution: a dual-pronged approach to averting cytokine storms in COVID-19? *Cancer Metastasis Rev*. 2020 Jun;39(2):337–40.
66. Sharif SA, Du X, Myles T, Song JJ, Price E, Lee DM, et al. Thrombin-activatable carboxypeptidase B cleavage of osteopontin regulates neutrophil survival and synoviocyte binding in rheumatoid arthritis. *Arthritis Rheum*. 2009 Oct;60(10):2902–12.
67. Morser J, Shao Z, Nishimura T, Zhou Q, Zhao L, Higgins J, et al. Carboxypeptidase B2 and N play different roles in regulation of activated complements C3a and C5a in mice. *J Thromb Haemost*. 2018 May;16(5):991–1002.
68. Foley JH, Conway EM. basic weapons to degrade C3a and C5a. *J Thromb Haemost*. 2018 May;16(5):987–90.
69. Maugeri N, Powell JE, 't Hoen PAC, de Geus EJC, Willemsen G, Kattenberg M, et al. LPAR1 and ITGA4 regulate peripheral blood monocyte counts. *Hum Mutat*. 2011 Aug;32(8):873–6.
70. Pearl JE, Torrado E, Tighe M, Fountain JJ, Solache A, Strutt T, et al. Nitric oxide inhibits the accumulation of CD4⁺ CD44^{hi} Tbet⁺ CD69^{lo} T cells in mycobacterial infection: Immunity to infection. *Eur J Immunol*. 2012 Dec;42(12):3267–79.
71. Zhang Y, Han K, Du C, Li R, Liu J, Zeng H, et al. Carboxypeptidase B blocks ex vivo activation of the anaphylatoxin-neutrophil extracellular trap axis in neutrophils from COVID-19 patients. *Crit Care*. 2021 Dec;25(1):51.
72. Gennaro R, Skerlavaj B, Romeo D. Purification, composition, and activity of two bactenecins, antibacterial peptides of bovine neutrophils. *Infect Immun*. 1989 Oct;57(10):3142–6.
73. Kościuczuk EM, Lisowski P, Jarczak J, Strzałkowska N, Józwiak A, Horbańczuk J, et al. Cathelicidins: family of antimicrobial peptides. A review. *Mol Biol Rep*. 2012 Dec;39(12):10957–70.
74. White SH, Wimley WC, Selsted ME. Structure, function, and membrane integration of defensins. *Current Opinion in Structural Biology*. 1995 Aug;5(4):521–7.
75. Gazit E, Boman A, Boman HG, Shai Y. Interaction of the mammalian antibacterial peptide cecropin p1 with phospholipid vesicles. *Biochemistry*. 1995 Sep 12;34(36):11479–88.
76. Brogden KA. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol*. 2005 Mar;3(3):238–50.
77. Seil M, Nagant C, Dehaye J-P, Vandenbranden M, Lensink MF. Spotlight on human LL-37, an immunomodulatory peptide with promising cell-penetrating properties. *Pharmaceuticals*. 2010 Nov 1;3(11):3435–60.

78. Buchanan SGStC, Gay NJ. Structural and functional diversity in the leucine-rich repeat family of proteins. *Progress in Biophysics and Molecular Biology*. 1996;65(1–2):1–44.
79. Pang KT, Ghim M, Liu C, Tay HM, Fhu CW, Chia RN, et al. Leucine-rich α -2-glycoprotein 1 suppresses endothelial cell activation through ADAM10-mediated shedding of TNF- α receptor. *Front Cell Dev Biol*. 2021 Jul 5;9:706143.
80. Jaiswal M, Dvorsky R, Ahmadian MR. Deciphering the molecular and functional basis of DBL family proteins. *Journal of Biological Chemistry*. 2013 Feb;288(6):4486–500.
81. Dong X, Mo Z, Bokoch G, Guo C, Li Z, Wu D. P-rax1 is a primary RAC2 guanine nucleotide exchange factor in mouse neutrophils. *Current Biology*. 2005 Oct;15(20):1874–9.
82. Bello L, Pegoraro E. The “usual suspects”: genes for inflammation, fibrosis, regeneration, and muscle strength modify duchenne muscular dystrophy. *JCM*. 2019 May 10;8(5):649.
83. Morimoto Y, Hirahara K, Kiuchi M, Wada T, Ichikawa T, Kanno T, et al. Amphiregulin-producing pathogenic memory T helper 2 cells instruct eosinophils to secrete osteopontin and facilitate airway fibrosis. *Immunity*. 2018 Jul;49(1):134-150.e6.
84. Hirahara K, Shinoda K, Morimoto Y, Kiuchi M, Aoki A, Kumagai J, et al. Immune cell-epithelial/mesenchymal interaction contributing to allergic airway inflammation associated pathology. *Front Immunol*. 2019 Mar 26;10:570.
85. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, et al. Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell*. 2020 Jul;182(1):59-72.e15.
86. Yu J, Yuan X, Chen H, Chaturvedi S, Braunstein EM, Brodsky RA. Direct activation of the alternative complement pathway by SARS-CoV-2 spike proteins is blocked by factor D inhibition. *Blood*. 2020 Oct 29;136(18):2080–9.
87. Herbert C, Do K, Chiu V, Garthwaite L, Chen Y, Young PM, et al. Allergic environment enhances airway epithelial pro-inflammatory responses to rhinovirus infection. *Clinical Science*. 2017 Mar 1;131(6):499–509.
88. An D, Guo Y, Bao J, Luo X, Liu Y, Ma B, et al. Molecular characterization and biological activity of bovine interferon-omega3. *Research in Veterinary Science*. 2017 Dec;115:125–31.
89. Pestka S, Krause CD, Walter MR. Interferons, interferon-like cytokines, and their receptors. *Immunol Rev*. 2004 Dec;202(1):8–32.
90. Walker AM, Roberts RM. Characterization of the bovine type I IFN locus: rearrangements, expansions, and novel subfamilies. *BMC Genomics*. 2009;10(1):187.
91. Noppert SJ, Fitzgerald KA, Hertzog PJ. The role of type I interferons in TLR responses. *Immunol Cell Biol*. 2007 Aug;85(6):446–57.

Table

Table 1

Receiver operator characteristic (ROC) curve analysis with area under the curve (AUC), sensitivity, and specificity.

Comparison	Gene	AUC	Cutoff	Sensitivity	Specificity
Treated_1 vs Healthy	ALDH1A1	0.520	8.115	0.360	0.730
Treated_2 + vs Healthy		0.533	8.585	0.667	0.478
Treated_2 + vs Treated_1		0.494	8.143	0.767	0.371
Treated_1 vs Healthy	ALOX15	0.590	12.855	0.966	0.191
Treated_2 + vs Healthy		0.635	8.703	0.400	0.904
Treated_2 + vs Treated_1		0.554	8.415	0.333	0.865
Treated_1 vs Healthy	BGN	0.514	8.115	0.348	0.730
Treated_2 + vs Healthy		0.527	8.585	0.667	0.478
Treated_2 + vs Treated_1		0.508	8.503	0.633	0.472
Treated_1 vs Healthy	CATHL3	0.540	8.388	0.416	0.696
Treated_2 + vs Healthy		0.625	8.585	0.633	0.626
Treated_2 + vs Treated_1		0.577	8.878	0.767	0.436
Treated_1 vs Healthy	CD27	0.524	13.946	0.382	0.730
Treated_2 + vs Healthy		0.659	14.306	0.800	0.557
Treated_2 + vs Treated_1		0.613	14.310	0.800	0.506
Treated_1 vs Healthy	CFB	0.483	9.629	0.483	0.583
Treated_2 + vs Healthy		0.656	10.058	0.667	0.661
Treated_2 + vs Treated_1		0.618	9.673	0.767	0.494
Treated_1 vs Healthy	CPB2	0.588	8.714	0.562	0.652
Treated_2 + vs Healthy		0.679	8.793	0.767	0.635
Treated_2 + vs Treated_1		0.573	9.020	0.833	0.404

ROC evaluation was used to generate log2CPM cutoffs for decision tree construction and assess DEGs as single-molecule predictors of BRD severity.

Comparison	Gene	AUC	Cutoff	Sensitivity	Specificity
Treated_1 vs Healthy	DSG1	0.516	8.115	0.360	0.739
Treated_2 + vs Healthy		0.499	8.792	0.733	0.374
Treated_2 + vs Treated_1		0.475	0.900	0.800	0.281
Treated_1 vs Healthy	DSP	0.499	8.115	0.337	0.739
Treated_2 + vs Healthy		0.494	8.585	0.600	0.487
Treated_2 + vs Treated_1		0.489	8.482	0.567	0.517
Treated_1 vs Healthy	GCSAML	0.543	9.225	0.337	0.774
Treated_2 + vs Healthy		0.750	9.340	0.700	0.739
Treated_2 + vs Treated_1		0.694	9.586	0.800	0.584
Treated_1 vs Healthy	GYPA	0.571	8.800	0.584	0.565
Treated_2 + vs Healthy		0.692	8.801	0.800	0.565
Treated_2 + vs Treated_1		0.612	8.881	0.833	0.404
Treated_1 vs Healthy	GZMB	0.609	10.318	0.618	0.583
Treated_2 + vs Healthy		0.732	9.786	0.700	0.696
Treated_2 + vs Treated_1		0.640	9.488	0.633	0.652
Treated_1 vs Healthy	HERC6	0.575	14.942	0.663	0.504
Treated_2 + vs Healthy		0.741	15.352	0.600	0.835
Treated_2 + vs Treated_1		0.660	15.352	0.600	0.719
Treated_1 vs Healthy	HPDG	0.608	9.582	0.809	0.391
Treated_2 + vs Healthy		0.617	8.964	0.700	0.609
Treated_2 + vs Treated_1		0.522	8.952	0.700	0.494
Treated_1 vs Healthy	IFI6	0.585	15.478	0.494	0.687
Treated_2 + vs Healthy		0.699	15.496	0.733	0.687

ROC evaluation was used to generate log2CPM cutoffs for decision tree construction and assess DEGs as single-molecule predictors of BRD severity.

Comparison	Gene	AUC	Cutoff	Sensitivity	Specificity
Treated_2 + vs Treated_1		0.631	15.630	0.667	0.640
Treated_1 vs Healthy	IL5RA	0.622	10.898	0.775	0.443
Treated_2 + vs Healthy		0.651	8.907	0.433	0.852
Treated_2 + vs Treated_1		0.555	8.921	0.433	0.742
Treated_1 vs Healthy	ISG15	0.545	16.689	0.708	0.417
Treated_2 + vs Healthy		0.699	17.930	0.600	0.800
Treated_2 + vs Treated_1		0.657	17.606	0.733	0.596
Treated_1 vs Healthy	ITGA4	0.540	15.461	0.303	0.817
Treated_2 + vs Healthy		0.676	15.674	0.633	0.713
Treated_2 + vs Treated_1		0.622	15.921	0.733	0.551
Treated_1 vs Healthy	KLF17	0.530	8.115	0.337	0.791
Treated_2 + vs Healthy		0.596	8.589	0.700	0.522
Treated_2 + vs Treated_1		0.552	8.996	0.867	0.303
Treated_1 vs Healthy	KRT10	0.509	8.115	0.337	0.748
Treated_2 + vs Healthy		0.471	8.444	0.500	0.557
Treated_2 + vs Treated_1		0.460	9.283	0.833	0.225
Treated_1 vs Healthy	LOC100297044	0.629	13.045	0.899	0.313
Treated_2 + vs Healthy		0.868	9.952	0.467	0.904
Treated_2 + vs Treated_1		0.577	9.981	0.467	0.764
Treated_1 vs Healthy	LOC100335828	0.529	9.695	0.371	0.730
Treated_2 + vs Healthy		0.731	9.784	0.700	0.704
Treated_2 + vs Treated_1		0.688	10.079	0.800	0.562

ROC evaluation was used to generate log2CPM cutoffs for decision tree construction and assess DEGs as single-molecule predictors of BRD severity.

Comparison	Gene	AUC	Cutoff	Sensitivity	Specificity
Treated_1 vs Healthy	LRG1	0.479	8.116	0.315	0.739
Treated_2 + vs Healthy		0.566	8.585	0.700	0.487
Treated_2 + vs Treated_1		0.579	8.996	0.867	0.337
Treated_1 vs Healthy	LTF	0.613	9.191	0.629	0.635
Treated_2 + vs Healthy		0.727	8.885	0.667	0.551
Treated_2 + vs Treated_1		0.615	9.211	0.800	0.617
Treated_1 vs Healthy	MCF2L	0.510	8.300	0.427	0.635
Treated_2 + vs Healthy		0.532	8.585	0.667	0.487
Treated_2 + vs Treated_1		0.514	8.996	0.867	0.292
Treated_1 vs Healthy	MGC126945	0.532	16.117	0.506	0.643
Treated_2 + vs Healthy		0.638	16.293	0.700	0.591
Treated_2 + vs Treated_1		0.617	15.081	0.267	0.933
Treated_1 vs Healthy	MS4A2	0.506	8.115	0.326	0.783
Treated_2 + vs Healthy		0.608	8.591	0.700	0.548
Treated_2 + vs Treated_1		0.591	8.997	0.867	0.404
Treated_1 vs Healthy	MX1	0.581	17.544	0.427	0.800
Treated_2 + vs Healthy		0.677	17.420	0.667	0.713
Treated_2 + vs Treated_1		0.598	17.643	0.500	0.742
Treated_1 vs Healthy	SLC18A2	0.494	8.998	0.281	0.783
Treated_2 + vs Healthy		0.563	8.585	0.700	0.478
Treated_2 + vs Treated_1		0.552	8.503	0.667	0.494
Treated_1 vs Healthy	SPP1	0.514	8.115	0.348	0.730
Treated_2 + vs Healthy		0.529	8.585	0.667	0.487

ROC evaluation was used to generate log2CPM cutoffs for decision tree construction and assess DEGs as single-molecule predictors of BRD severity.

Comparison	Gene	AUC	Cutoff	Sensitivity	Specificity
Treated_2 + vs Treated_1		0.508	8.996	0.833	0.281
Treated_1 vs Healthy	IFN (HERC6, IFI6, ISG15, MX1)	0.578	64.739	0.652	0.539
Treated_2 + vs Healthy		0.730	66.093	0.733	0.748
Treated_2 + vs Treated_1		0.661	65.926	0.767	0.618
Treated_1 vs Healthy	SPM (ALOX15, HPGD)	0.598	17.288	0.292	0.887
Treated_2 + vs Healthy		0.646	17.407	0.433	0.878
Treated_2 + vs Treated_1		0.548	17.787	0.467	0.685

ROC evaluation was used to generate log2CPM cutoffs for decision tree construction and assess DEGs as single-molecule predictors of BRD severity.

Figures

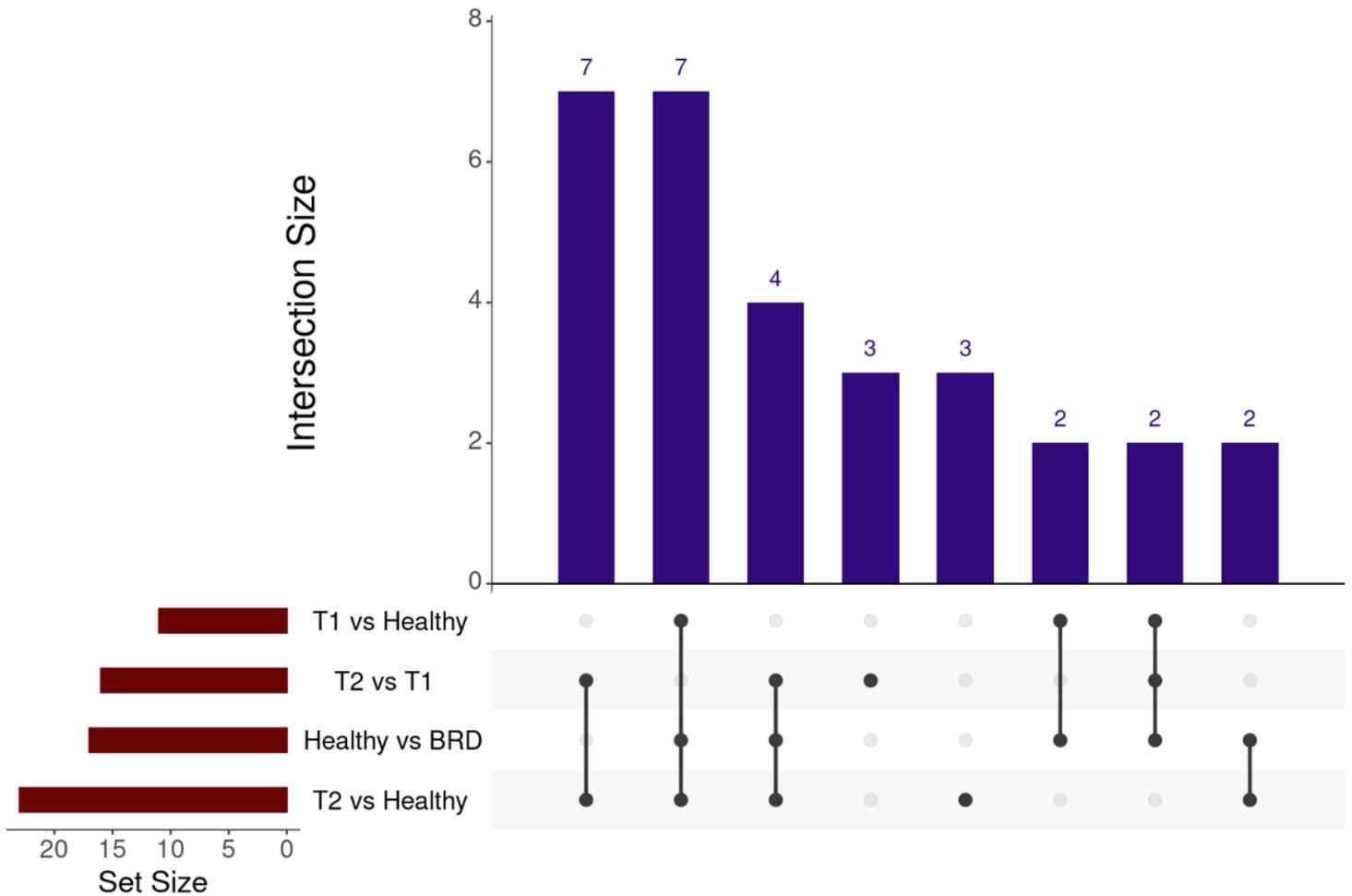


Figure 1

Overlap of the 30 unique DEGs identified from nSolver analysis between the four comparative analyses. Interaction size represents the quantity of overlapping genes within and between each analysis. Set size represents the total number of DEGs identified in each analysis. Healthy are cattle never diagnosed nor received antimicrobial therapy for clinical BRD. T1 are cattle treated for clinical BRD once within 28 days of arrival. T2 are cattle having been treated at least two times within 28 days of arrival for and/or died following a diagnosis of BRD. BRD are T1 and T2 cattle combined into one group.

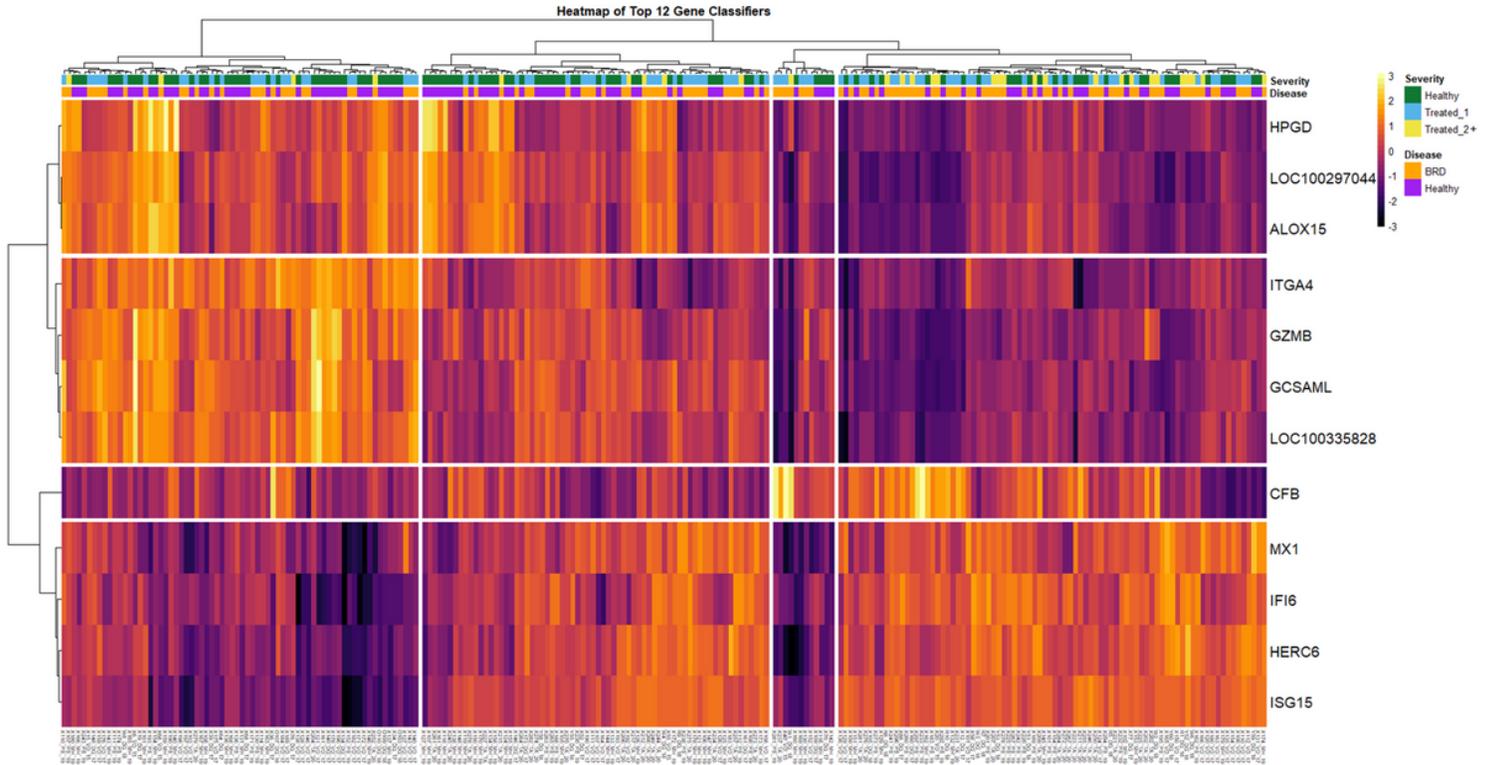


Figure 2

Heatmap and unsupervised hierarchical clustering of the gene expression from twelve select DEGs. Expression values (-3 to +3) were calculated from z-score calculation of nCounter normalized expression values for each gene. Samples were further labeled for disease status (BRD, Healthy) and severity (Healthy, Treated_1, Treated_2+). Yellow/white: relative high expression; purple/black: relative low expression.

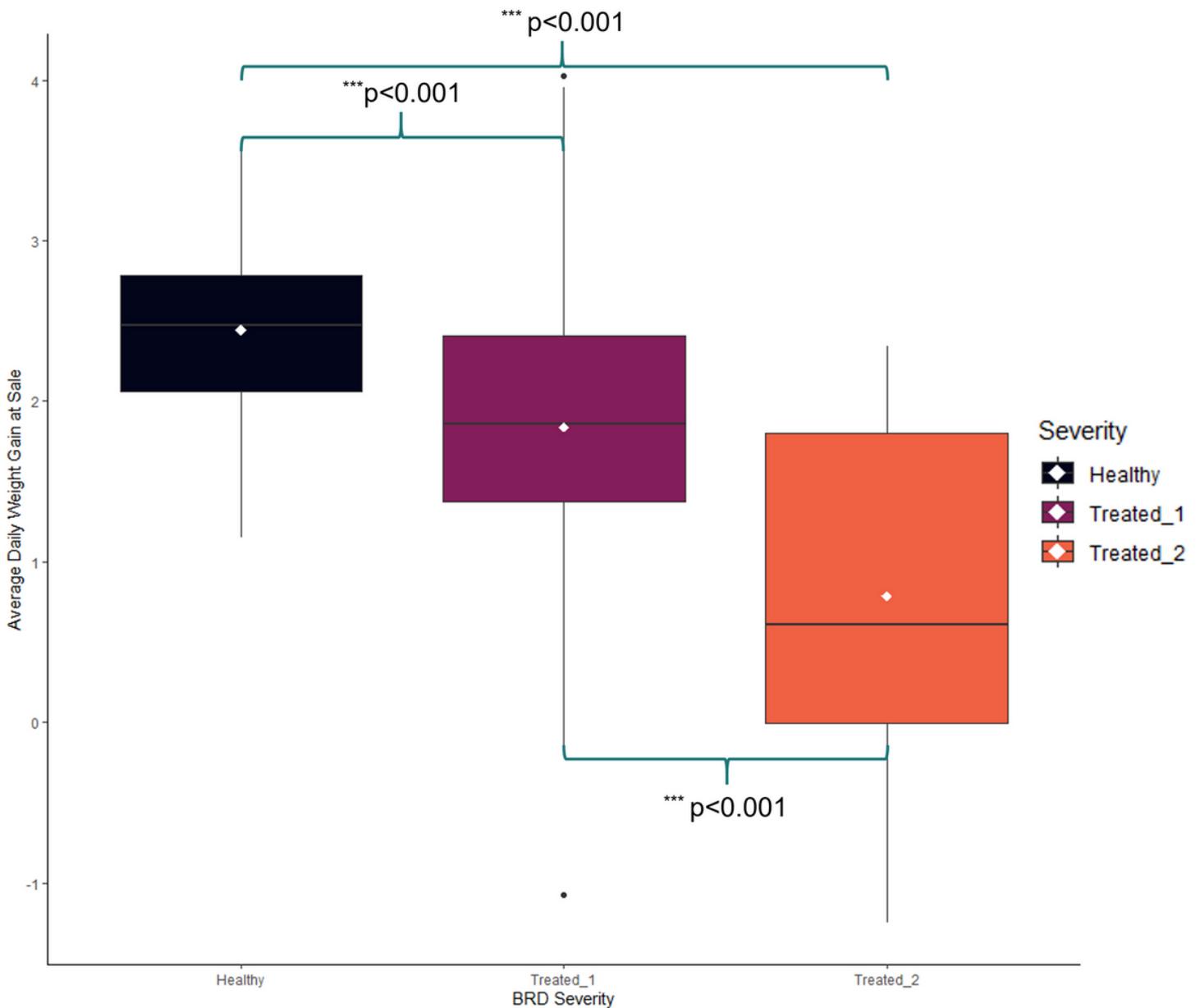


Figure 3

Distributive differences in average daily weight gain in pounds at time of sale (ADGend). Boxplots limits are associated with the first (lower) and third (upper) quartiles. Horizontal lines within the boxplots represent the ADGend median for each cohort. White points within the boxplots represent ADGend mean for each cohort. Whiskers for each boxplot extend to 1.5 times the interquartile ranges. Any black point outside the vertical range of whiskers represents outlier individuals within the associated cohort.

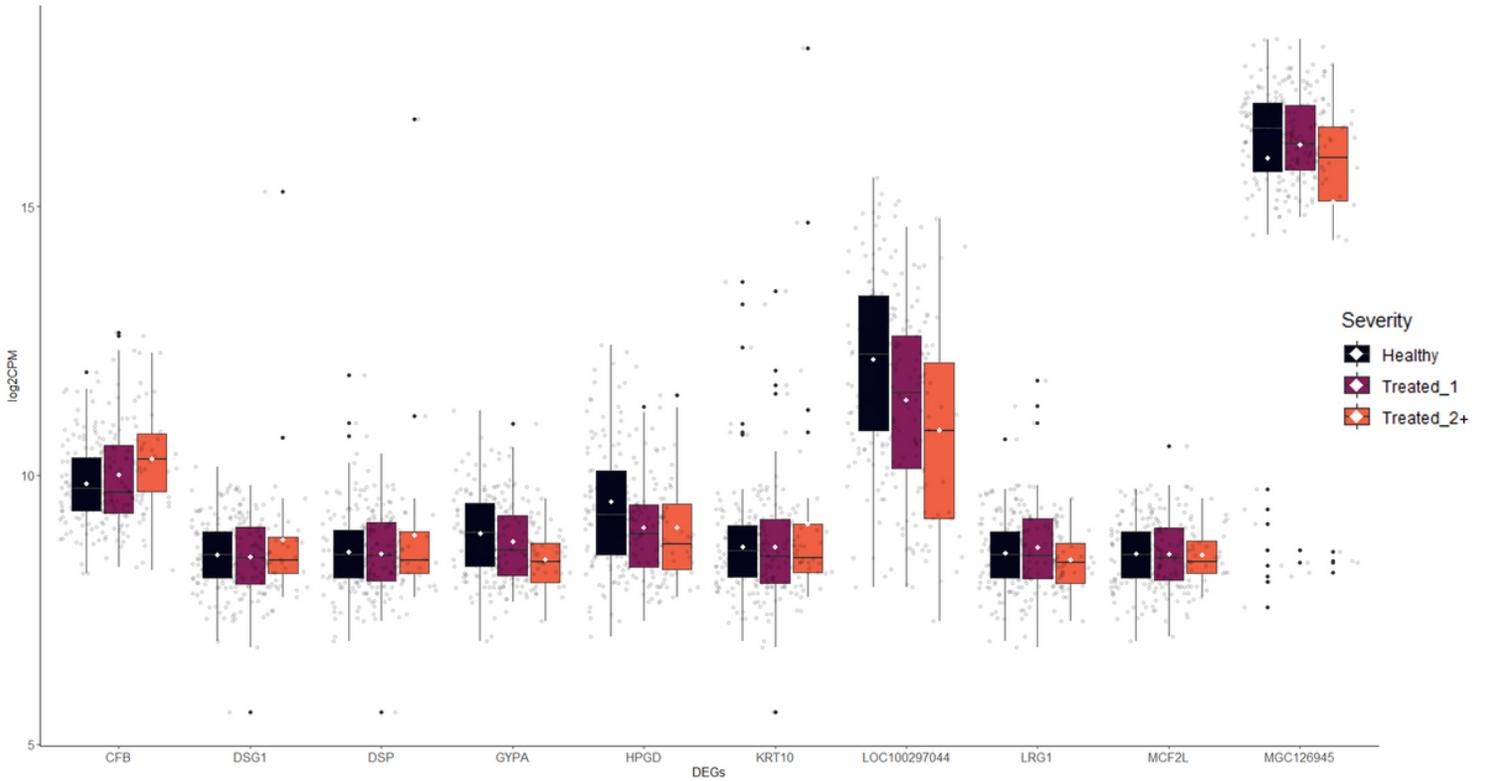


Figure 4

Boxplots of the log₂CPM gene expression for the ten genes with unequal variance across severity. Genes were identified through the Levene's test for homogeneity of variance ($p < 0.10$). Boxplots limits are associated with the first (lower) and third (upper) quartiles. Horizontal lines within the boxplots represent the median log₂CPM for each cohort. White points within the boxplots represent log₂CPM mean for each cohort. Whiskers for each boxplot extend to 1.5 times the interquartile ranges. Any black point outside the vertical range of whiskers represents outlier individuals within the associated cohort.

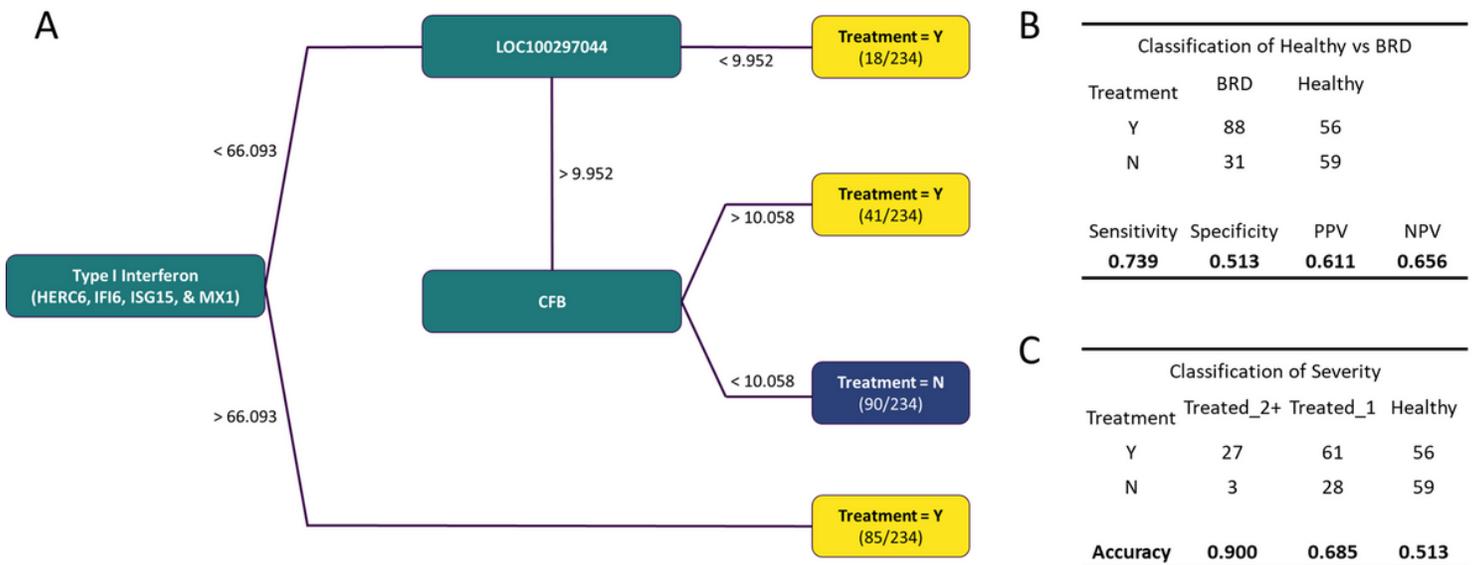


Figure 5

Gene expression decision tree model constructed from log2CPM cutoffs derived from ROC curve evaluation. A) Combined type I interferon genes (HERC6, IFI6, ISG15, and MX1) serve as the root of the decision tree, with an additive log2CPM cutoff of 66.093. Individuals with a value above the threshold are marked as needing BRD treatment, and below the threshold move to the LOC100297044 leaf. Those within the LOC100297044 leaf below a threshold of 9.952 are marked as needing BRD treatment, and above the threshold move to the CFB leaf. Those within the CFB leaf above a threshold of 10.058 are marked as needing BRD treatment, and those below the threshold are considered Healthy. B) Diagnostic accuracy table for the assessment of identified Healthy and BRD individuals. C) Diagnostic accuracy table for the assessment of identified Treated_2+, Treated_1, and Healthy individuals.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.png](#)
- [SupplementalTables.xlsx](#)