J. Dairy Sci. 106
https://doi.org/10.3168/jds.2022-22573

Associations among the genome, rumen metabolome, ruminal bacteria, and milk production in early-lactation Holsteins

H. M. Golder,[1,2] J. Thomson,[3] J. Rehberger,[4] A. H. Smith,[4] E. Block,[4] and I. J. Lean[1,2]*
[1]Scibus, Camden, NSW, Australia, 2570
[2]Sydney Institute of Agriculture, School of Life and Environmental Sciences, Faculty of Science, The University of Sydney, Camden, NSW, Australia, 2570
[3]Department of Animal and Range Sciences, Montana State University, Bozeman 59717
[4]Arm & Hammer Animal and Food Production, Princeton, NJ 08540

ABSTRACT

A multicenter observational study to evaluate genome-wide association was conducted in early-lactation Holstein cows (n = 293) from 36 herds in Canada, the USA, and Australia. Phenotypic observations included rumen metabolome, acidosis risk, ruminal bacterial taxa, and milk composition and yield measures. Diets ranged from pasture supplemented with concentrates to total mixed rations (nonfiber carbohydrates = 17 to 47, and neutral detergent fiber = 27 to 58% of dry matter). Rumen samples were collected <3 h after feeding and analyzed for pH, ammonia, D- and L-lactate, volatile fatty acid (VFA) concentrations, and abundance of bacterial phyla and families. Eigenvectors were produced using cluster and discriminant analyses from a combination of pH and ammonia, D-lactate, and VFA concentrations, and were used to estimate the probability of the risk of ruminal acidosis based on proximity to the centroid of 3 clusters, termed high (24.0% of cows), medium (24.2%), and low risk (51.8%) for acidosis. DNA of sufficient quality was successfully extracted from whole blood (218 cows) or hair (65 cows) collected simultaneously with the rumen samples and sequenced using the Geneseek Genomic Profiler Bovine 150K Illumina SNPchip. Genome-wide association used an additive model and linear regression with principal component analysis (PCA) population stratification and a Bonferroni correction for multiple comparisons. Population structure was visualized using PCA plots. Single genomic markers were associated with milk protein percent and the center logged ratio abundance of the phyla *Chloroflexi*, *SR1*, and *Spirochaetes*, and tended to be associated with milk fat yield, rumen acetate, butyrate, and isovalerate concentrations and with the probability of being in the low-risk acidosis group. More than one genomic marker was associated or tended to be associated with rumen isobutyrate and caproate concentrations, and the center log ratio of the phyla *Bacteroidetes* and *Firmicutes* and center log ratio of the families *Prevotellaceae*, *BS11*, *S24-7*, *Acidaminococcaceae*, *Carnobacteriaceae*, *Lactobacillaceae*, *Leuconostocaceae*, and *Streptococcaceae*. The provisional *NTN4* gene, involved in several functions, had pleiotropy with 10 bacterial families, the phyla *Bacteroidetes* and *Firmicutes*, and butyrate. The *ATP2CA1* gene, involved in the ATPase secretory pathway for Ca^{2+} transport, overlapped for the families *Prevotellaceae*, *S24-7*, and *Streptococcaceae*, the phylum *Bacteroidetes*, and isobutyrate. No genomic markers were associated with milk yield, fat percentage, protein yield, total solids, energy-corrected milk, somatic cell count, rumen pH, ammonia, propionate, valerate, total VFA, and D-, L-, or total lactate concentrations, or probability of being in the high- or medium-risk acidosis groups. Genome-wide associations with the rumen metabolome, microbial taxa, and milk composition were present across a wide geographical and management range of herds, suggesting the existence of markers for the rumen environment but not for acidosis susceptibility. The variation in pathogenesis of ruminal acidosis in the small population of cattle in the high risk for acidosis group and the dynamic nature of the rumen as cows cycle through a bout of acidosis may have precluded the identification of markers for acidosis susceptibility. Despite a limited sample size, this study provides evidence of interactions between the mammalian genome, the rumen metabolome, ruminal bacteria, and milk protein percentage.

Key words: genome-wide association study, ruminal acidosis, microbiome, quantitative trait loci

INTRODUCTION

Ruminal acidosis is a prominent and complex nutritional disorder of ruminants. The condition is associ-

Received July 25, 2022.
Accepted November 19, 2022.
*Corresponding author: ianl@scibus.com.au

ated with an accumulation of organic acids within the rumen initiated by the combination of consumption of large amounts of readily fermentable carbohydrates and insufficient intake of physically effective fiber (Nagaraja and Titgemeyer, 2007; Bramley et al., 2008). Firkins and Yu (2015) noted that differences in the microbiome structure among cattle fed the same diet often exceed those fed on different diets. Current control strategies are unlikely to manage ruminal acidosis in the entire herd due to the considerable variation in the microbiome (Golder et al., 2014). Genetic selection is commonly used in the cattle industry to improve a range of production, fertility, and longevity traits and minimize health disorders (Cole et al., 2021). The history of the evolution of genetic evaluation in the dairy industry, current traits, and their weighting in selection indices, evaluation methods, and benefits of genetic improvement have been well described in reviews (Wiggans et al., 2017; Cole et al., 2021). No estimated breeding values are currently available for rumen disorders, acidosis susceptibility, or rumen microbiome composition or function.

Host-microbe interactions are increasingly being explored (Weimer, 2015; Myer, 2019). Host-microbe specificity has been demonstrated through the near re-establishment of bacterial communities after near-total exchanges of rumen contents of dairy cattle (Weimer et al., 2010; Weimer et al., 2017) and rumen transplantation in sheep under acidotic conditions (Liu et al., 2019). Breed of cattle has also been found to influence the rumen community (Clemmons et al., 2019; Li et al., 2019a,b), further demonstrating links between the genetics of the host and its microbiota. Bovines appear to have a core microbiota that differs among individuals (Li et al., 2009; Welkie et al., 2010; Jami and Mizrahi, 2012; Henderson et al., 2015). Wallace et al. (2019) found that a heritable subset of the core rumen microbiota influenced the productivity and emissions of cows. Quantifying the components of variation in the core microbiota of cattle that reflect the mammalian genetics, influence of diet, and other environmental influences may provide means to reduce susceptibility to ruminal acidosis and other associated disorders. Genome-wide association studies are a widely used approach to identify QTL associated with phenotypes (Jiang et al., 2019).

A GWAS pilot study found associations between the mammalian genetics, bacterial phyla, and rumen fermentation products in a small population (n = 33) of nulliparous Holstein heifers that were subjected to an acidosis challenge (Golder et al., 2018). This suggested the potential to use genetic markers to select for cattle at lower risk of acidosis. Because this was a pilot study with a very small study population size that used nulliparous heifers from the same herd with closed genetics, a need remained to explore such associations in a larger population size, with a range of genetics and management systems, and to incorporate associations with production measures (Golder et al., 2018). To address this gap in knowledge, we used the model by Bramley et al. (2008) as described in Golder et al. (2023) to classify 293 early-lactation Holstein cattle from 4 geographical regions [Australia (**AU**), Canada (**CAN**), California (**CA**), and Wisconsin (**WI**)] into 3 acidosis risk groups. In the study population of Golder et al. (2023; n = 263) that excluded cows from the WI region, 26.1% of the cows were classified in the high-risk group, 26.8% in the medium-risk group, and 47.1% in the low-risk group. The overall bacterial community composition differed among the high- and low-risk groups and regions, and associations were identified among bacterial community composition and measures of rumen metabolites and milk production (Golder et al., 2023; our unpublished data). Exploration of the associations with the host genome is required and described herein.

This paper is part of a series (Golder et al., 2023; our unpublished data) designed to improve our understandings of the risk of acidosis and ultimately control of ruminal acidosis through evaluation of associations with host genomics, diet, rumen metabolite, ruminal bacterial taxa, and production characteristics in early-lactation Holsteins. The objective of this study was to explore the aforementioned associations by GWAS in early-lactation dairy cattle from a range of management systems and geographical spread, so that genetic markers may be used to select cattle with a reduced genetic predisposition to ruminal acidosis. We hypothesized that associations would be found between the bovine genome and rumen metabolome, acidosis risk, rumen microbiota, and production of lactating dairy cattle.

MATERIALS AND METHODS

This study was carried out in accordance with the recommendations of the Australian Code for Care and Use of Animals for Scientific Purposes, Scibus Animal Ethics Committee (Scibus 0618-1219), the University of California Davis Institutional Animal Care and Use Committee (protocol number 20729), the University of Wisconsin, College of Agriculture and Life Sciences Animal Care and Use Committee (approval A006225), and Animal Care Committee at the University of Guelph (Animal Utilization Protocol 4124).

Experimental Design

Phenotypic and genotypic data were obtained from 293 lactating Holstein-Friesian cows in a multicenter observational study. The study population was from 36 herds (10 in CA, 12 in CAN, 10 in AU, and 4 in WI), with a target of 8 cows/herd (4 primiparous and 4 of parity >1 and ≤4) between 10 and 100 DIM. The AU herds were in 4 geographical regions (Western districts of Victoria, the South Coast of New South Wales, Finley in New South Wales, and the Macarthur region of New South Wales). The herds ranged in size from 57 to 6,294 cows and are described by Golder et al. (2023) along with their selection criteria. The diets of the North American herds were all TMR, whereas those in AU were primarily pasture-based with supplementary concentrates or silage. The chemical composition of the diets was analyzed by DairyOne Cooperative Inc., Forage Testing Laboratory (Ithaca, NY), as described and reported in Golder et al. (2023). The NFC ranged from 17 to 47% of DM, and the NDF ranged between 27 and 58% of DM.

Phenotypic Data

The phenotypic data included milk production and rumen metabolomic and bacterial taxa abundance data. A summary of the associations between these and the acidotic groups and the regions of CA, CAN, and AU is shown in Figure 1.

Milk. The milk production measures included milk volume, fat percentage and yield, protein percentage and yield, natural log SCC, and ECM. Milk volume, fat, protein, and SCC were recorded from the closest herd test to the rumen sampling date, details of the herd testing agencies are provided in Golder et al. (2023). Energy-corrected milk was calculated as ECM = [(0.3246 × milk yield) + (12.86 × fat yield) + (7.04 × protein yield)], as per NRC (2001).

Rumen Metabolomic Data. The metabolomic measures included total VFA, acetate, butyrate, propionate, valerate, ammonia, total lactate, D-lactate and L-lactate concentrations, the ratio of acetate to propionate, pH, and the probability of being classified based on the discriminant and K-means cluster analysis model developed by Bramley et al. (2008) into the following acidosis risk group classifications: high, medium, and low. The predicted means (±SEM) for these measures are provided in Golder et al. (2023).

Rumen fluid collections were planned to coincide with herd testing, but this was not always feasible. Rumen fluid was collected as described by Golder et al. (2023) within 3 h of feeding, using a customized stomach tube approximately 3 m in length with a multi-holed aluminum probe at one end, into a 500-mL container. Rumen fluid was scored for saliva contamination as described by Bramley et al. (2008) using a 3-point scoring system (0 being the lowest and 2 being the highest level of contamination). The rumen fermentation measures were analyzed according to methods described by Golder et al. (2023).

Acidosis Risk Group Classification. To classify the risk status of acidosis for each cow's rumen fluid sample, the measures for individual VFA, ammonia, D-lactate, and pH for each sample were appended to the existing data set of Bramley et al. (2008), which had been used to develop K-means cluster analysis group allocation. All measures were standardized to a z-statistic and a nearest neighbor nonparametric discriminant analysis (*discrim knn*; Stata Version 16.1, StataCorp.) to classify the sample from each cow into 1 of the 3 acidosis categories based on the K-means clustering of Bramley et al. (2008): category 1, acidotic; category 2, suboptimal rumen function; or category 3, normal. We have termed our categories as follows: high, medium, and low risk for acidosis. A total of 24.0% of the cows were classified into the high-risk group, 24.2% into the medium-risk group, and 51.8% into the low-risk group.

The probability for distance to the center of each of the 3 acidosis risk groups (high, medium, and low) was produced where values were between 0 and 1. These describe how closely the rumen metabolomic profile from each cow in the current study matches the center of the high, medium, or low group. Values approaching 1 are in the centroid of the group and represent cows that are most consistent with the characteristics defining that group.

Golder et al. (2023) proposed that the high-risk group had rumen phyla, fermentation, and production characteristics consistent with a model of acidosis that reflected a rapid rate of carbohydrate fermentation. Namely, an acetate to propionate ratio of 1.98 ± 0.11, concentrations of valerate 2.93 ± 0.14 m*M*, milk fat to protein ratio 1.11 ± 0.047, and a positive association with abundance of the phylum *Firmicutes*. The medium-risk group was proposed to contain cows that may be inappetent, had not eaten recently, or were in recovery from acidosis (Golder et al., 2023). The low-risk group may represent cattle that are well fed, with a stable rumen and a slower rumen fermentation of carbohydrates (Golder et al., 2023).

Rumen Bacterial Data. The ruminal bacterial data included the center log ratio of the relative abundance of the following bacterial phyla: *Actinobacteria*, *Armatimonadetes*, *Bacteroidetes*, *Chloroflexi*, *Cyanobacteria*, *Fibrobacteres*, *Firmicutes*, *Lentisphaerae*, *Planctomycetes*, *Proteobacteria*, *Saccharibacteria*, *Spirochaetes*, *SR1*, *Synergistetes*, *Tenericutes*, and *Verrucomicrobia*.

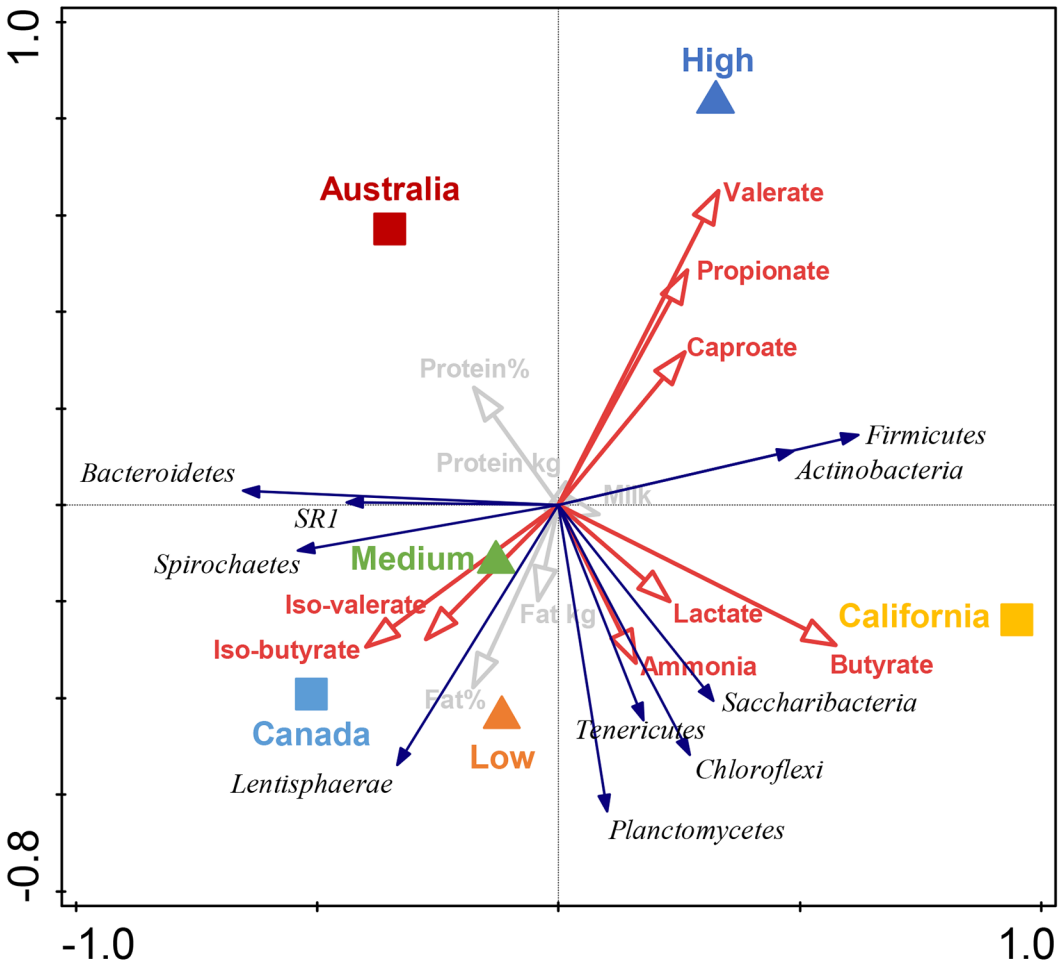

Figure 1. Correlation biplot of the redundancy analysis of bacterial phyla with respect to acidosis risk and region and rumen and milk explanatory variables. The 10 phyla with the best fit to the explanatory variables are displayed, where the length of the arrows are approximate correlation coefficients between the variables and acidosis risk and region. The total variation explained by the explanatory variables is 26.7%, and the eigenvalues for the first 2 axes are 0.12 and 0.09. High-risk and low-risk acidosis group samples were different ($P = 0.038$) and were associated with 5.1% and 3.0% of the variation, respectively. The medium-risk acidosis group was not different and was only associated with 0.4% of the variation ($P = 1.000$). Each region was significantly different ($P = 0.038$), with California associated with 8.8% of the variation, Australia with 4.7%, and Canada with 4.3%. Valerate, butyrate, propionate, isobutyrate, caproate, isovalerate, and ammonia concentrations and milk fat percentage were different ($P = 0.038$) and were associated with 5.3, 4.9, 3.2, 3.2, 2.3, 1.9, 1.9, and 1.8% of the variation, respectively. Total lactate concentration tended to be different ($P = 0.176$), associating with 1.5% of the variation. Milk protein percentage and yields, milk fat yield, and milk volume were not different and were collectively only associated with 3.2% of the variation ($P > 0.100$). Total VFA and acetate concentrations and pH were not significant and are not displayed ($P > 0.100$).

The 29 families included are those that were considered the top 20 most influential families for either acidosis group or region, based on principal component analysis (**PCA**).

DNA was extracted using a DNeasy PowerSoil HTP Kit (Qiagen), sequenced (NovaSeq 6000, Illumina Inc.), and processed (QIIME2 version 2021.2; Bolyen et al., 2019) as described in Golder et al. (2023). Taxonomy was assigned by closed-reference clustering using VSEARCH (Rognes et al., 2016) to the EZBioCloud reference database (https://www.ezbiocloud.net/; Kim et al., 2012). All samples with fewer than 7,000 reads, and all archaea and eukarya were removed from the analysis. The center log ratio of the relative abundance of bacterial phyla and families with an abundance of >0.3% were used.

Beta diversity was examined by redundancy analysis using Canoco5 (Microcomputer Power). The amount of variation for the model was calculated, as well as each axis and each explanatory variable included in the model. The significance of each explanatory variable was determined by Bonferroni-corrected *P*-value.

Genotypic Data

Hair samples were collected from 79 cows in Tulare, CA, only; of these, only 66 had sufficient DNA quality. Blood was collected by tail venipuncture in blood tubes containing no anticoagulant from 214 of the cattle in CAN, AU, and WI, and 11 from CA that had poor-quality DNA from the hair samples. Two cows that had poor-quality DNA from the hair sample did not have blood samples; these cows were excluded from the analysis. Whole blood was frozen at −20°C and shipped on dry ice for DNA extraction.

DNA Extraction. Bovine DNA from hair was extracted by Neogen Geneseek Inc. (Lincoln, NE) using their standard protocols.

Bovine DNA from frozen whole blood samples was extracted using a Maxwell 16 LEV Blood DNA Kit, cat. no. AS1290 (Promega), and Maxwell 16 Semi Automated Nucleic Acid Extractor (Promega) according to manufacturer recommended procedures, as described in detail in Golder et al. (2018), and 30 μL of DNA elutant was frozen at −20°C until further analysis.

Initial DNA quality and quantity was determined by reading the absorbance at 260- and 280-nm frequencies. A desirable ratio for DNA is 1.8. Samples yielding a ratio between 1.65 and 1.95 were considered to have sufficient purity for genotyping, and samples that yielded at least a DNA concentration of 10 ng/μL were genotyped. All samples extracted met these thresholds, averaging 54 ng/μL.

Genotyping. Bovine DNA from hair and blood were shipped to Neogen Geneseek Inc. (Lincoln, NE) for analysis using the Geneseek Genomic Profiler Bovine 150K Illumina SNPchip. Chips were prepared and run according to manufacturer guidelines, and the data were made electronically available to Montana State University (Bozeman, MT).

The SNP genotypes by animal identification, SNP mapping file containing chromosome (**Chr**) and position for each SNP, and phenotype by animal identification files were imported into Golden Helix software.

Data Quality Control. Samples were filtered out if genotypes could be called in <95% of the SNPs on the genotyping array, as this indicates that DNA quality was compromised or contaminated. A total of 6 samples had poor genotype call rates and were excluded from subsequent analysis.

Three criteria were applied for exclusion of SNPs: call rate <0.95, minor allele frequency <0.01, and Hardy-Weinberg equilibrium Fischer's exact test *P*-value <0.0001. After quality control filtering, 119,205 SNPs were included in analysis.

Linkage disequilibrium pruning was performed with a window size of 100 bp and a window increment of 5 bp. The linkage disequilibrium r^2 threshold was 0.5, and the composite haplotype method was used for the linkage disequilibrium computation method (Weir, 1996). This inactivated 8,668 markers and left 110,537 in the analysis.

Genotype Analysis

Population structure was visualized using PCA. Some structure was evident when the first 2 eigenvectors were plotted, and data were coded by herd identification; hence these eigenvectors were included as covariates in the analysis to correct for underlying population structure or any underlying genetic variation such as regional or country variation. This is shown in Figure 2.

An identity-by-descent matrix was constructed by calculating identity-by-descent for each pair of samples. A genome *F*-statistic was calculated for each animal, with overall results of maximum = 0.23, minimum = −0.06, and mean = 0.003 ± 0.037.

Genome-wide association testing was completed on data from individual phenotypes utilizing 2 different models. Genome-wide association testing was completed using an additive model (dd) versus (Dd) versus (DD) and linear regression where the response y was fit to every genetic predictor variable or encoded genotype. The response was represented with the formula $y = b_1x + b_0 + \in$, where the model was represented by the expression $b_1x + b_0$ and the error term $\in$, expressing the difference, or residual, between the model of the response and the response itself (Bůžková, 2013). This model was chosen for its simplicity, to avoid adding bias to the results. Additionally, a correlation/trend test was performed with genotype predictor values classified according to reference or alternate alleles with the same additive model outlined previously, where the count of the alternate allele A, which is 0 within genotype rr, 1 within genotype Ar, and 2 within genotype AA, where r is the reference allele. The correlation/trend test is a package in Golden Helix, which tests the significance of any correlation between 2 numeric variables (or 2 variables that have been encoded as numeric variables). Population stratification was corrected for using PCA by the method of the program EIGENSTRAT, as described by Price et al. (2006). This PCA methodology was applied due to limited data on animal genetic variance related to the phenotypic characteristics studied.

A Bonferroni adjustment was used to account for multiple comparisons. All *P*-values reported are Bonferroni adjusted, with *P*-values of ≤0.05 considered significant and those between 0.05 and 0.15 reported as trends.

Positional candidate genes were identified around markers that were significantly associated with a phe-

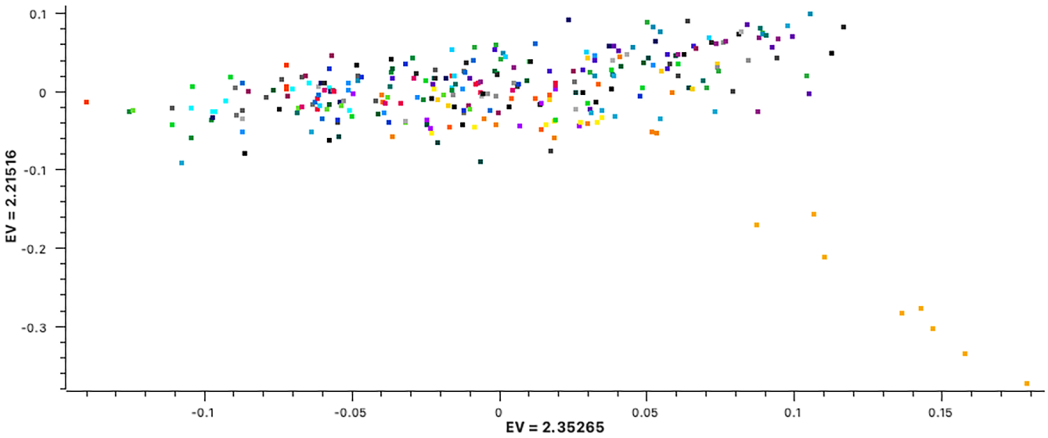

Figure 2. Principal component analysis plot of population structure, with data coded by herd identification. Samples from each of the 36 herds are represented by a unique color, with approximately n = 8 per herd. The herd represented by the orange squares in the lower right-hand corner of the figure that are distinct from the other samples are from a pasture-based herd in the Finley region of New South Wales, Australia, that had long-term use of British Holstein genetics. EV = eigenvector.

notype. A window of 100,000 bp was used in the case of a single associated marker, and where clusters of markers were identified, the positional candidate genes were identified within a window of 3 times the span of the cluster.

RESULTS

A total of 283 samples had both a genotype and phenotype matched. Bovine DNA of sufficient quality was successfully extracted from 217 whole blood samples and 66 hair samples. The herd that differed in population structure (Figure 2) was a pasture-based herd in the Finley region of New South Wales, AU, and had long-term use of British Holstein genetics.

Host-Production Interactions

The association analysis and the correlation/trend test found that a marker on Chr 26 tended to be significant for milk fat yield, with 4 genes identified ($P = 0.064$) and 1 for milk protein percentage on Chr 1 with 2 genes identified ($P = 0.020$; Table 1). No associations were found for milk yield, ECM, milk fat percentage or yield, milk protein yield, total solids, the ratio of milk fat to milk protein, or normal log SCC.

Host-Rumen Metabolomic Interactions

No associations were found for saliva score, rumen pH, or concentrations of rumen ammonia, propionate, valerate, total VFA, D- or L-lactate, the ratio of acetate to propionate, and the probability of being in the high- or medium-risk acidosis groups. We did find tendencies for a marker for acetate on Chr 6 (not observed in correlation/trend test) and for butyrate and isovalerate on Chr 5 and 26, respectively, each with a single gene identified (Table 2). Caproate had 2 markers (Chr 11 and 2) but only 1 gene. Isobutyrate had an associated marker on Chr 1 with a single gene and tendencies for a further 2 markers from Chr 2 with 3 genes between them, including an overlap for *GALNT13*. These markers were not observed in the correlation/trend test, but the *DHX32* gene had a trend in this analysis ($P = 0.150$; Table 2); this was the same gene that had a tendency for isovalerate. The low-risk acidosis group had a tendency for a QTL on Chr 2 with 6 genes ($P = 0.105$), but this QTL was not identified in the correlation/trend test. All the markers identified from the rumen metabolomics were unique to those identified for the milk parameters.

Host-Rumen Bacterial Interactions

Of the rumen phyla included in the GWAS, only the *Bacteroidetes*, *Chloroflexi*, *Firmicutes*, *SR1*, and *Spirochaetes* had associations, most of which were provisional genes, and all were $P \leq 0.06$ (Table 3). The *Bacteroidetes* had 3 markers from Chr 5, 1, and 4, respectively, with single genes identified for each that all overlapped with other phyla. The marker and gene from Chr 5 demonstrated pleiotropy with overlap for *Firmicutes* and butyrate; the marker and gene from Chr 1 were shared with isobutyrate, and the marker and gene from Chr 4 had overlap with *Firmicutes*. The *Chloroflexi* had a marker on Chr 13 with 1 gene, and *SR1* had a tendency for a single marker association with 2 provisional genes ($P = 0.060$), but neither of these markers were

Table 1. Markers identified in association analysis for milk phenotypes[1]

Phenotype	Marker	Chr	Position	*P*-value		Gene symbol (transcript ID)	Gene name	Position interval (bp)
				LR	Cor			
Fat yield (kg/d)	BovineHD2600012141	26	43,663,927	0.064	0.064	*GPR26*	G protein-coupled receptor 26	43,419,202–43,448,291
						CPXM2	Carboxypeptidase X, M14 family member 2	43,495,815–43,632,782
						CHST15	Carbohydrate sulfotransferase 15	43,712,673–43,793,152
						LOC112444485	Uncharacterized LOC112444485	
Protein (%)	ARS-BFGL-NGS-35604	1	150,084,135	0.020	0.020	*KCNJ6*	Potassium inwardly rectifying channel subfamily J member 6	149,810,382–149,947,602
						LOC112348308	Small nucleolar RNA SNORA72	

[1]The chromosomal position and significance of each marker are provided, along with the genes identified and their position within the quantitative trait loci. *P*-values are presented for both linear regression (LR) and correlation/trend test (Cor) analyses. All genes were protein-coding, except *LOC112444485* and *LOC112348308*, which are noncoding RNA. n = 266. Chr = chromosome.

identified in the correlation/trend test. The *Firmicutes* had 2 associations (Chr 5 and 4) each with single genes, but only provisional gene Netrin 4 (*NTN4*) was identified in the correlation/trend test. The *Spirochaetes* had a single QTL containing 8 primarily provisional genes.

Markers were identified for 14 of the 29 families that underwent association analysis, of which 6 were gram-positive from the *Firmicutes* phylum (Table 4). All except the *Lachnospiraceae*, *Acholeplasmataceae*, and *Coriobacteriaceae* had multiple putative QTL identified per marker or multiple markers. The *Anaerolineaceae*, *Anaerocella*, *BS11* (tendencies), and *Planctomycetaceae* (tendency; $P = 0.060$) were the only families that had QTL that were distinct from all the other phenotypes. The QTL for the *Planctomycetaceae* contained 7 genes.

The most common gene, overlapping in 10 of the families that were from a range of phyla and were not consistent in Gram stain, was *NTN4* from the marker BovineHD0500016819, located on Chr 5. This gene also overlapped for *Firmicutes*, *Bacteroidetes*, and butyrate. The *ATP2CA1* gene, involved in the ATPase secretory pathway for Ca^{2+} transport, overlapped for the *Prevotellaceae* and *S24-7*, both from the *Bacteroidetes* phylum, as well as the *Bacteroidetes* phylum itself, the family *Streptococcaceae* (tendency; $P = 0.055$), and isobutyrate. The *Carnobacteriaceae* had the most QTL identified, with 5 comprising 16 genes. The *GRM8* gene from Chr 4 overlapped between *Prevotellaceae*, *Carnobacteriaceae*, and *Firmicutes*. The gene Neuregulin 3 (*NRG3*) from Chr 28 overlapped for *Prevotellaceae* and *Acidaminococcaceae*. The marker BovineHD0700006320 on Chr 7 had consistent QTL between *Acidaminococcaceae* (tendency; $P = 0.054$), *Carnobacteriaceae*, *Lactobacillaceae*, *Leuconostocaceae*, and *Streptococcaceae*. None of the tendencies that were identified for the bacterial families in the regression analysis were identified in the correlation/trend test. In addition, 2 of the QTL identified, BovineHD0400025098 and BovineHD1100019497, from *Carnobacteriaceae*, were also not identified in the correlation/trend test.

DISCUSSION

As hypothesized, significant QTL were found for rumen metabolites, ruminal microbiota, and milk production. The associations relate to differences in the metabolome, microbiota, and production between the acidosis groups (Golder et al., 2023) and may be attributed to the host genome. This study population, although larger than that of Golder et al. (2018), upon which the hypothesis was based, is still small for a QTL identification study. The possibility of type I error has been reduced by use of a conservative statistical approach, consistent with that used by Golder et al.

Table 2. Markers identified in association analysis for rumen phenotypes[1]

Phenotype	Marker	Chr	Position	*P*-value		Gene symbol (transcript ID)	Gene name	Position interval (bp)
				LR	Cor			
Acetate (m*M*)	Hapmap23971-BTC-035093	6	40,158,436	0.112	—	*SLIT2*	Slit guidance ligand 2	39,779,576–40,184,421
Butyrate (m*M*)	BovineHD0500026819	5	60,120,293	0.067	0.136	*NTN4* (NM_001102203.1)	Provisional. Netrin 4	60,031,675–60,161,682
Isobutyrate (m*M*)	BovineHD0100039652	1	138,862,010	0.006	—	*ATP2C1*	ATPase secretory pathway Ca^{2+} transporting 1	138,871,975–139,025,561
	BovineHD0200012072	2	41,598,134	0.135	—	*KCNJ3*	Potassium voltage-gated channel subfamily J member 3	41,183,416–41,378,419
						GALNT13	Polypeptide *N*-acetylgalactosaminyltransferase 13	41,675,904–423,332,413
	BovineHD0200012088	2	41,673,247	0.151	—	*GALNT13*	Polypeptide *N*-acetylgalactosaminyltransferase 13	41,675,904–423,332,413
	BovineHD2600012739	26	45370168	—	0.150	*DHX32*	DEAH-box helicase32 (putative)	45,341,124–45,394,655
Isovalerate (m*M*)	BovineHD2600012739	26	45,370,168	0.074	0.136	*DHX32*	DEAH-box helicase32 (putative)	45,341,124–45,394,655
Caproate (m*M*)	ARS-BFGL-NGS-32173	11	3,557,549	0.014	0.034	*VWA3B* (NR_144296.2)	*Homo sapiens* von Willebrand factor A domain containing 3B	3,393,946–35,603,113
	BTA-85505-no-rs	2	87,258,372	0.030	0.066	None		
Distance to the center of the low-risk acidosis group[2]	BovineHD0200012905	2	44,482,018	0.105	—	*STAM2* (NM_001076106.1)	Provisional: signal transducing adaptor molecule 2	44,006,701–44,053,728
						CACNB4 (NM_001192104.1)	Provisional: calcium voltage-gated channel auxiliary subunit β 4	44,073,031–44,344,826
						ARL5A (NM_001046245.2)	Provisional: ADP ribosylation factor like GTPase 5A	44,355,896–44,378,846
						TNFAIP6 (NM_001007813.2)	Provisional: TNF α induced protein 6	44,747,346–44,763,738
						NMI (NM_001035098.2)	Provisional: N-myc and STAT interactor	44,826,376–44,846,381
						RBM43 (NM_001099168.1)	Provisional: RNA binding motif protein 43	44,857,602–44,869,114

[1]The chromosomal position and significance of each marker are provided, along with the genes identified and their position within the quantitative trait loci. *P*-values are presented for both linear regression (LR) and correlation/trend test (Cor) analyses. All genes were protein-coding. n = 283. Chr = chromosome.

[2]Cattle were classified into 1 of 3 acidosis risk groups (high, medium, or low) using the discriminant analysis and K-means cluster analysis methods described in Bramley et al. (2008) based on standardized individual values of rumen acetate, propionate, butyrate, valerate, isobutyrate, isovalerate, caproate, ammonia, D-lactate, ammonia, and pH. The characteristics of the acidosis groups are described in Golder et al. (2023). The value tested here is the distance from the center of the low-risk acidosis group as a value between 0 and 1, where 1 is the center of the group or cluster; hence, higher values represent those close to the center of the group.

Table 3. Markers identified in association analysis for ruminal bacterial phylum phenotypes[1]

Phylum	Gram	Marker	Chr	Position	*P*-value		Gene symbol (transcript ID)	Gene name	Position interval (bp)
					LR	Corr			
Bacteroidetes	−	BovineHD0500016819	5	60,120,293	<0.001	<0.001	*NTN4* (NM_001102203.1)	Provisional. Netrin 4	60,031,675–60,161,682
		BovineHD0100039652	1	138,862,010	<0.001	<0.001	*ATP2C1*	ATPase secretory pathway Ca^{2+} transporting 1	138,871,975–139,025,561
		BovineHD0400025098	4	90,571,828	0.019	0.046	*GRM8* (NM_002206117.1)	Provisional. Glutamate metabotrophic receptor 8	90,671,110–91,510,132
Chloroflexi	−	BovineHD1300013052	13	44,799,619	0.048	—	*KLF6* (NM_001035271.3)	Provisional. Kruppel like factor 6	44,597,302–44,604,392
Firmicutes	+	BovineHD0500016819	5	60,120,293	0.048	<0.001	*NTN4* (NM_001102203.1)	Provisional. Netrin 4	60,031,675–60,161,682
		BovineHD0400025098	4	490,571,828	0.048	—	*GRM8* (NM_002206117.1)	Provisional. Glutamate metabotrophic receptor 8	90,671,110–91,510,132
SR1	+	BovineHD0200018680	2	64,759,816	0.060	—	*LYPD1* (NM_001046359.1)	Provisional. LY6/PLAUR domain containing 1	64,874,976–64,937,090
							GPR39 (NM_001143744.1)	Provisional. G protein-coupled receptor 39	64,934,063–65,196,636
Spirochaetes	−	BovineHD1200022563	12	79,192,619	0.002	0.007	*FGF14* (NM_001206832.1)	Provisional. Fibroblast growth factor 14	78,130,314–78,245,386 and 78,764,801–78,765,008
							TPP2 (NM_001099034.2)	Provisional. Tripeptidyl peptidase 2	78,857,739–78,915,646
							METTL21C (NM_001100383.1)	Provisional. Methyltransferase like 21C	78,920,742–78,929,824
							TEX30 (NM_001076295.2)	Validated. Testis expressed 30	79,008,350–79,014,993
							POGLUT2 (NM_001075685.1)	Protein O-glucosyltransferase 2 precursor	79,016,631–79,029,810
							BIVM (NM_001206453.2)	Validated. Basic, immunoglobulin-like variable motif containing	79,030,045–79,076,215
							ERCC5 (NM_001035276.2)	Provisional. ERCC excision repair 5, endonuclease	79,079,749–79,107,518
							METTL21E (NM_001014922.1)	Provisional. RIKEN cDNA 4832428D23-like	79,112,916–79,128,997

[1]The chromosomal position and significance of each marker are provided, along with the genes identified and their position within the quantitative trait loci. The Gram stain results of each phylum is noted. All genes were protein-coding. *P*-values are presented for both linear regression (LR) and correlation/trend test (Cor) analyses. n = 277. Chr = chromosome.

Table 4. Markers identified in association analysis for ruminal bacterial family phenotypes[1]

Phylum	Family	Gram	Marker	Chr	Position	*P*-value		Gene symbol (transcript ID)	Gene name	Position interval (bp)
						LR	Corr			
Actinobacteria	*Coriobacteriaceae*	+	BovineHD0500016819	5	60,120,293	0.001	0.003	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
Bacteroidetes	*Anaerocella*	−	BTA-27444-no-rs	5	50,229,041	0.010	0.024	*SRGAP1* (NM_001205668.1)	Provisional: SLIT-ROBO Rho GTPase activating protein 1	49,582,507–49,889,557
								RXYLT1 (NM_001038114.2)	Provisional: Ribitol xylosyltransferase 1	49,919,471–49,945,712
								AVPR1A (NM_001104990.1)	Provisional: Arginine vasopressin receptor 1A	50,363,479–50,366,959
								PPM1H (NM_001193049.1)	Provisional: Protein phosphatase, MG^{2+}/MN^{2+} dependent 1H	50,634,861–50,937,844
Bacteroidetes	*BS11*	−	BovineHD07000016767	7	58,604,611	0.115	—	*STK32A* (NM_001105047.1)	Provisional: serine/threonine kinase 32A	58,457,689–58,586,892
								MIR2284Y-7 (NR_107790.1)	Provisional: microRNA 2284y-7	58,553,819–58,553,895
								DPYSL3 (NM_001101068.1)	Dihydropyrimidinase like 3	58,595,962–58,710,806
			BovineHD0700016607	7	58,081,395	0.137	—	*STK32A* (NM_001105047.1)	Provisional: serine/threonine kinase 32A	58,457,689–58,586,892
								MIR2284Y-7 (NR_107790.1)	Provisional: microRNA 2284y-7	58,553,819–58,553,895
								DPYSL3 (NM_001101068.1)	Dihydropyrimidinase like 3	58,595,962–58,710,806
Bacteroidetes	*Prevotellaceae*	−	BovineHD0500016819	5	60,120,293	<0.001	<0.001	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
			BovineHD0100039652	1	138,862,010	<0.001	0.001	*ATP2C1*	ATPase secretory pathway Ca^{2+} transporting 1	138,871,975–139,025,561
			BovineHD0400025098	4	60,120,293	0.013	0.033	*GRM8* (NM_002206117.1)	Provisional: Glutamate metabotrophic receptor 8	90,671,110–91,510,132
			BovineHD2800010542	28	60,120,293	0.019	0.045	*NRG3* (NM_001205478.1)	Provisional: Neuregulin 3	36,896,430–38,132,526
Bacteroidetes	*S24-7*	−	BovineHD0500016819	5	60,120,293	<0.001	<0.001	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
			BovineHD0100039652	1	138,862,010	0.051	—	*ATP2C1*	ATPase secretory pathway Ca^{2+} transporting 1	138,871,975–139,025,561
Chloroflexi	*Anaerolineaceae*	−	BovineHD1300013052	13	44,799,619	0.051		*KLF6* (NM_001035271.3)	Provisional: Kruppel like factor 6	44,597,302–44,604,392
								PFKP (NM_001193220.2)	Provisional: phosphofructokinase, platelet	45,107,875–45,160,801
Firmicutes	*Acidaminococcaceae*	+	BovineHD0500016819	5	60,120,293	<0.001	<0.001	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
			BovineHD2800010542	28	38,122,718	0.025	0.059	*NRG3* (NM_001205478.1)	Provisional: Neuregulin 3	36,896,430–38,132,526
			BovineHD0700006320	7	23,045,544	0.054	—	*CDC42SE2* (NM_001102537.1)	Provisional: CDC42 small effector 2	23,073,267–23,192,277
								LYRM7 (NM_001076505.2)	Provisional: LYR motif containing 7	23,244,077–23,265,677
								HINT1 (NM_175812.2)	Provisional: Histidine triad nucleotide binding protein 1	23,271,941–23,278,540
Firmicutes	*Carnobacteriaceae*	+	Hapmap43107-BTA-95423	3	87,190,871	<0.001	0.001	*JUN* (NM_001077827.1)	Provisional: Jun proto-oncogene, AP-1 transcription factor	87,265,962–87,268,009
								MYSM1 (NM_001192408.1)	Inferred: Myb like, SWIRM and MPN domains 1	87,349,611–87,391,325

Continued

Table 4 (Continued). Markers identified in association analysis for ruminal bacterial family phenotypes[1]

Phylum	Family	Gram	Marker	Chr	Position	*P*-value		Gene symbol (transcript ID)	Gene name	Position interval (bp)
						LR	Corr			
								OMA1 (NM_001035033.2)	Provisional: OMA1 zinc metallopeptidase	87,479,311–87,544,530
			BovineHD0500016819	5	60,120,293	<0.001	0.001	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
			BovineHD0700006320	7	23,045,544	0.007	0.017	*CDC42SE2* (NM_001102537.1)	Provisional: CDC42 small effector 2	23,073,267–23,192,277
								LYRM7 (NM_001076505.2)	Provisional: LYR motif containing 7	23,244,077–23,265,677
								HINT1 (NM_175812.2)	Provisional: Histidine triad nucleotide binding protein 1	23,271,941–23,278,540
			BovineHD0400025098	4	90,571,828	0.028	—	*GRM8* (NM_002206117.1)	Provisional: Glutamate metabotrophic receptor 8	90,671,110–91,510,132
			BovineHD1100019497	11	68,913,397	0.032	—	*LYRM7* (NM_001034714.2)	Predicted: Chromosome 11 C2orf42 homolog	68,474,083–68,513,151
								TIA1 (NM_001076109.1)	Provisional: TIA1 cytotoxic granule associated RNA binding protein	68,523,382–68,554,903
								PCYOX1 (NM_001105474.2)	Provisional: Penylcysteine oxidase 1	68,588,248–68,599,766
								SNRPG (NM_001040543.3)	Validated: Small nuclear ribonucleoprotein polypeptide G	68,602,486–68,611,488
								EHD3 (NM_001098004.1)	Provisional: EH domain containing 3	68,659,832–68,688,786
								CAPN14 (NM_001192425.1)	Provisional: Calpain 14	68,706,426–68,742,726
								GALNT14 (NM_001192732.1)	Provisional: Polypeptide N-acetylgalactosaminyltransferase 14	68,772,689–68,987,976
								CAPN13 (NM_001035360.1)	Provisional: Calpain 13	69,068,510–69,144,511
Firmicutes	*Lachnospiraceae*	+	BovineHD0500016819	5	60,120,293	0.099	—	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
Firmicutes	*Lactobacillaceae*	+	BovineHD0700006320	7	23,045,544	0.001	0.002	*CDC42SE2* (NM_001102537.1)	Provisional: CDC42 small effector 2	23,073,267–23,192,277
								LYRM7 (NM_001076505.2)	Provisional: LYR motif containing 7	23,244,077–23,265,677
								HINT1 (NM_175812.2)	Provisional: Histidine triad nucleotide binding protein 1	2,327,194–123,278,540
			BovineHD0500016819	5	760,120,293	0.011	0.028	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
Firmicutes	*Leuconostocaceae*	+	BovineHD0700006320	7	23,045,544	<0.001	0.001	*CDC42SE2* (NM_001102537.1)	Provisional: CDC42 small effector 2	23,073,267–23,192,277
								LYRM7 (NM_001076505.2)	Provisional: LYR motif containing 7	23,244,077–23,265,677
								HINT1 (NM_175812.2)	Provisional: Histidine triad nucleotide binding protein 1	23,271,941–23,278,540
			BovineHD0500016819	5	60,120,293	0.003	0.010	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
Firmicutes	*Streptococcaceae*	+	BovineHD0500016819	5	60,120,293	<0.001	<0.001	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682
			BovineHD0700006320	7	23,045,544	0.003	0.010	*CDC42SE2* (NM_001102537.1)	Provisional: CDC42 small effector 2	23,073,267–23,192,277
								LYRM7 (NM_001076505.2)	Provisional: LYR motif containing 7	23,244,077–23,265,677

Continued

Table 4 (Continued). Markers identified in association analysis for ruminal bacterial family phenotypes[1]

Phylum	Family	Gram	Marker	Chr	Position	*P*-value		Gene symbol (transcript ID)	Gene name	Position interval (bp)
						LR	Corr			
								HINT1 (NM_175812.2)	Provisional: Histidine triad nucleotide binding protein 1	23,271,941–23,278,540
			BovineHD0100013828	1	49,043,446	0.008	0.002	None		
			BovineHD0100039652	1	138,862,010	0.055	—	*ATP2C1*	ATPase secretory pathway Ca^{2+} transporting 1	138,871,975–139,025,561
Planctomycetes	*Planctomycetaceae*	–	ARS-BFGL-NGS-16590	11	44,478,837	0.060	—	*MRPL19* (NM_001046068.2)	Provisional: Mitochondrial ribosomal protein L19	44,019,412–44,030,088
								SEPTIN10 (NM_001046176.1)	Provisional: Septin 10	44,267,491–44,323,708
								RANBP2 (NM_174592.1)	Validated: RAN binding protein 2	44,646,694–44,708,488
								LIMS1 (NM_001083495.1)	Provisional: LIM zinc finger domain containing 1	44,715,273–44,796,595
								GCC2 (NM_001192259.1)	Provisional: Domain containing 2	44,806,992–44,840,325
								SULT1C4 (NM_001080919.1)	Provisional: Sulfotransferase family, cytosolic, 1C, member 4	44,862,157–44,869,843
								CHRNA5.L (NM_001093080.1)	Cholinergic receptor nicotinic α 5 subunit L homeolog	44,906,839–44,921,512
Tenericutes	*Acholeplasmataceae*	–	BovineHD0500016819	5	60,120,293	0.029	0.066	*NTN4* (NM_001102203.1)	Provisional: Netrin 4	60,031,675–60,161,682

[1]The chromosomal position and significance of each marker are provided, along with the genes identified and their position within the quantitative trait loci. The phylum that each family belongs to and its Gram stain status are provided. All genes were protein-coding. *P*-values are presented for both linear regression (LR) and correlation/trend test (Cor) analyses. n = 277. Chr = chromosome.

(2018). It is challenging and costly to obtain data of the quality required to evaluate large populations; smaller studies such as these are important to initiate proof of concept and give direction for larger-scale studies. A similar approach was taken by Benson et al. (2010), who examined factors that influence microbial composition in an intercross line of mice (n = 645). Both Benson et al. (2010) and Golder et al. (2018) reported multiple QTL for each phylum and pleiotrophy with several QTL among microbial taxa. Golder et al. (2018) also reported genetic associations with the metabolome, specifically with rumen function.

The intent of our study was to evaluate variation in rumen conditions that resulted from routine commercial feeding management, not to impose a rumen disturbance or intervention. We anticipated, based on the incidence of acidosis in early-lactation cattle reported by Bramley et al. (2008), that 10% of our sampling population would be acidotic. Classification of cattle using the discriminant analysis and the K-means cluster analysis methodology of Bramley et al. (2008) resulted in 24.0% of the cows classified in the high, 24.2% in the medium, and 51.8% in the low-risk acidosis groups for this study. Differences in rumen metabolomics, bacterial taxa, and milk composition were evident between the acidosis groups (Golder et al., 2023). Thus, our incidence of acidosis and separation of characteristics of cattle between the 3 acidosis risk groups is sufficient to evaluate genomic associations.

Our findings suggest that responses to changes in the rumen could be host-specific. This is supported by associations between the host genome, metabolome, and microbiota described under rumen disturbance conditions in cattle on the same diets (Weimer et al., 2010; Weimer et al., 2017). Those authors demonstrated the capability of the rumen to rapidly revert to pre-exchange VFA concentrations within 24 h, and for microbial composition to revert more variably over a greater period after near-total exchange of rumen fluid. Ruminal microbial homeostasis was restored within 14 d in sheep with acidosis induced by an oligofructose challenge, regardless of rumen fluid transplantation from clinically healthy sheep; however, the rebound from the dysbiotic state was faster in the transplanted sheep (Liu et al., 2019). Breed of cattle also influences the rumen community (Clemmons et al., 2019; Li et al., 2019a,b).

The number of associated markers for the rumen metabolome and microbiota measures were considerably lower than those reported in the pilot study by Golder et al. (2018). This is expected, because the population from the current study was larger and more diverse in genetics and herd-level environmental factors including diet and management, thereby reducing the risk of type I error. Golder et al. (2018) had large numbers of significant markers for both lactate isomers in a pre-acidosis challenge period and for the ratio of acetate to propionate within 4 h of an acidosis challenge. In contrast, no markers were identified for these parameters in the current study, suggesting that these variables have a relatively smaller association with host genomics in wider populations.

The only marker associated with the acidosis risk groups was for the low-risk group, the group that represented most of the cows and which could indicate cows with suboptimal rumen function (Bramley et al., 2008; Golder et al., 2023). It is possible that a genetic construct associated with less optimal rumen function may reflect maintenance of a more diverse rumen population (Golder et al., 2023) that favors herd and individual survival under differing dietary and management systems. The positional candidate genes identified for the low-risk group are primarily involved in protein pathways and signal transduction. The N-myc and STAT interactor, often referred to as the *NMI* gene, is involved in transcription events in tumor growth (Pruitt et al., 2016). The gene TNF α induced protein 6 is an immunomodulatory protein produced in inflammatory diseases and is involved in mediating macrophage plasticity and inhibits neutrophil migration and production of proinflammatory cytokines (Dyer et al., 2016; Mittal et al., 2016). The lack of identified markers for the high- and medium-risk acidosis groups may reflect the size and characteristics of these populations and requires further investigation. Despite the breadth of research in the field of ruminal acidosis several questions remain surrounding its pathogenesis. Several markers and a QTL were associated with the acidosis eigenvalue in Golder et al. (2018), but none of the genes in the QTL overlap with those in the QTL from the current study. The acidosis eigenvalue is a comparable measure to what we termed the high-risk group in the current study. The *Planctomycetaceae* from the *Planctomycetes* phylum are gram-negative and associated with the global carbon and nitrogen cycles (Wiegand et al., 2018), with a QTL containing genes involved in sulfotransferase activity and protein-related functions.

Milk production traits are among the 36 traits currently included in the various breeding evaluation indexes for individual dairy breeds across the world (Cole et al., 2021). We demonstrated links with only fat yield and protein percentage, perhaps reflecting the small population size and range of herds sampled. Environmental and management effects create herd-level factors that influence genetic evaluations (Dunne et al., 2018). Wallace et al. (2019) found that a heritable subset of the core rumen microbiota influenced cow productivity and emissions. Associations between milk production

and milk components with bacterial taxa have also been made (Lima et al., 2015), as well as correlations between the relative abundance of taxa and production outcomes (Jami et al., 2014). In contrast, we did not observe any pleiotropy between milk production traits and any other measure.

The identification of an overlapping genetic region of *NTN4* for several bacterial families, from a range of phyla and inconsistent in Gram stain, may be of interest and supports the notion of redundancy of function among bacterial taxa. This region was also associated with the phyla *Firmicutes* and *Bacteroidetes* and the short-chain fatty acid butyrate. This gene may also be prevalent because it has several functions such as promoting neurite extension, regulating pulmonary airway branching, vasculogenesis patterning, endothelial proliferation in pathological angiogenesis, and negative regulation of vascular branching in the retina; it also inhibits osteoclast differentiation and is involved in tumor metastasis (Enoki et al., 2017; Xu et al., 2017). The current known functions do not suggest a direct link to butyrate production. Butyrate is usually the VFA with the third-highest concentration, after acetate and propionate, and is not highly diagnostic for acidosis (Golder et al., 2012). The indicators that were the most diagnostic, valerate and propionate (Bramley et al., 2008; Golder et al., 2012), were not associated with any markers.

The pleiotropy between *the Prevotellaceae*, *S24-7*, and *Streptococcaceae* families and isobutyrate for the *ATP2CA1* gene, which is involved in the ATPase secretory pathway for Ca^{2+} transport, is also interesting, as there is no immediate link to isobutyrate.

Golder et al. (2018) reported very few significant markers for the 2 most prevalent bacterial phyla, *Firmicutes* and *Bacteroidetes*, but both had more than one marker in the current study. The overlap in putative regions between the families suggests redundancy of function, as previously mentioned. The *Streptococcaceae*, *Leuconostocaceae*, *Lactobacillaceae*, *Carnobacteriaceae*, and *Acidaminococcaceae* families that had an overlapping QTL are lactic acid-producing bacteria (Al Jassim et al., 2005; Bergey et al., 2011; Kauter et al., 2019); thus it is logical that they have an association with rumen disturbance and have genomic regions in common. One of the genes in this QTL has already been discussed, *NTN4*, the *LYRM7* gene in human cells; it is involved in mitochondrial complex III assembly (Sánchez et al., 2013). Although *HINT1* is a protein-coding gene for histidine triad nucleotide binding protein 1 (*HINT1*), which is involved in tumor suppression and may have a role in the treatment of neuropsychiatric diseases (Jung et al., 2020).

CONCLUSIONS

Genome-wide associations with the rumen metabolome, microbial taxa, and milk protein percentage were present across a wide geographical and management range of herds, as hypothesized, but not with an acidotic risk status. This suggests that markers for the rumen environment exist, but not for acidosis susceptibility. The variation in pathogenesis of ruminal acidosis in the small population of acidotic cattle, and the dynamic nature of the rumen as cows cycle through a bout of acidosis, may have precluded the identification of markers for acidosis susceptibility. Despite a limited sample size, this study provides evidence of interactions between the mammalian genome, the rumen metabolome, ruminal bacterial, and milk protein percentage.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of Arm & Hammer (Princeton, NJ) and Scibus (Camden, NSW, Australia) and thank all the herds that contributed data to this study. The authors also acknowledge the assistance of S. J. LeBlanc (University of Guelph, Guelph, ON, Canada), T. Duffield (University of Guelph), H. A. Rossow (University of California Davis, Tulare), R. Bogdanich (Cross Street Veterinary Clinic, Tulare, CA), and L. Hernandez (University of Wisconsin-Madison, WI), and their respective staff, for assistance in collecting the samples and accessing the herd recording agency data. Funding was contributed by Arm & Hammer Animal and Food Production (Princeton, NJ). The authors have not stated any other conflicts of interest.

REFERENCES

Al Jassim, R. A., P. T. Scott, A. L. Trebbin, D. Trott, and C. C. Pollitt. 2005. The genetic diversity of lactic acid producing bacteria in the equine gastrointestinal tract. FEMS Microbiol. Lett. 248:75–81. https://doi.org/10.1016/j.femsle.2005.05.023.

Benson, A. K., S. A. Kelly, R. Legge, F. Ma, S. J. Low, J. Kim, M. Zhang, P. L. Oh, D. Nehrenberg, K. Hua, S. D. Kachman, E. N. Moriyama, J. Walter, D. A. Peterson, and D. Pomp. 2010. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc. Natl. Acad. Sci. USA 107:18933–18938. https://doi.org/10.1073/pnas.1007028107.

Bergey, D. H., P. De Vos, D. R. Boone, G. M. Garrity, R. W. Castenholz, D. J. Brenner, N. R. Krieg, and J. T. Staley. 2011. Bergey's Manual of Systematic Bacteriology: Volume 3: The Firmicutes. 2nd ed. Springer.

Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C. T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. K. Cope, R. Da Silva, C. Diener, P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki, J. Fouquier, J. M. Gauglitz, S. M. Gibbons,

D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B. Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. A. Huttley, S. Janssen, A. K. Jarmusch, L. Jiang, B. D. Kaehler, K. B. Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T. Kosciolek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y.-X. Liu, E. Loftfield, C. Lozupone, M. Maher, C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. T. Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L. Peoples, D. Petras, M. L. Preuss, E. Pruesse, L. B. Rasmussen, A. Rivers, M. S. Robeson II, P. Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R. Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft, F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J. Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu, R. Knight, and J. G. Caporaso. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37:852–857. https://doi.org/10.1038/s41587-019-0209-9.

Bramley, E., I. J. Lean, W. J. Fulkerson, M. A. Stevenson, A. R. Rabiee, and N. D. Costa. 2008. The definition of acidosis in dairy herds predominantly fed on pasture and concentrates. J. Dairy Sci. 91:308–321. https://doi.org/10.3168/jds.2006-601.

Bůžková, P. 2013. Linear regression in genetic association studies. PLoS One 8:e56976. https://doi.org/10.1371/journal.pone.0056976.

Clemmons, B. A., B. H. Voy, and P. R. Myer. 2019. Altering the gut microbiome of cattle: Considerations of host-microbiome interactions for persistent microbiome manipulation. Microb. Ecol. 77:523–536. https://doi.org/10.1007/s00248-018-1234-9.

Cole, J. B., J. W. Dürr, and E. L. Nicolazzi. 2021. Invited review: The future of selection decisions and breeding programs: What are we breeding for, and who decides? J. Dairy Sci. 104:5111–5124. https://doi.org/10.3168/jds.2020-19777.

Dunne, F. L., M. M. Kelleher, S. Walsh, and D. Berry. 2018. Characterization of best linear unbiased estimates generated from national genetic evaluations of reproductive performance, survival, and milk yield in dairy cows. J. Dairy Sci. 101:7625–7637. https://doi.org/10.3168/jds.2018-14529.

Dyer, D. P., C. L. Salanga, S. C. Johns, E. Valdambrini, M. M. Fuster, C. M. Milner, A. J. Day, and T. M. Handel. 2016. The anti-inflammatory protein TSG-6 regulates chemokine function by inhibiting chemokine/glycosaminoglycan interactions. J. Biol. Chem. 291:12627–12640. https://doi.org/10.1074/jbc.M116.720953.

Enoki, Y., T. Sato, S. Kokabu, N. Hayashi, T. Iwata, M. Yamato, M. Usui, M. Matsumoto, T. Tomoda, W. Ariyoshi, T. Nishihara, and T. Yoda. 2017. Netrin-4 promotes differentiation and migration of osteoblasts. In Vivo 31:793–799. https://doi.org/10.21873/invivo.11132.

Firkins, J. L., and Z. Yu. 2015. Ruminant nutrition symposium: How to use data on the rumen microbiome to improve our understanding of ruminant nutrition. J. Anim. Sci. 93:1450–1470. https://doi.org/10.2527/jas.2014-8754.

Golder, H. M., P. Celi, A. R. Rabiee, E. Bramley, and I. J. Lean. 2012. Validation of an acidosis model. Pages 118–122 in Proc. Dairy Research Foundation, Camden, NSW, Australia. University Printing Service Sydney.

Golder, H. M., P. Celi, A. R. Rabiee, and I. J. Lean. 2014. Effects of feed additives on rumen and blood profiles during a starch and fructose challenge. J. Dairy Sci. 97:985–1004. https://doi.org/10.3168/jds.2013-7166.

Golder, H. M., S. J. LeBlanc, T. Duffield, H. A. Rossow, R. Bogdanich, L. Hernandez, E. Block, J. Rehberger, A. H. Smith, J. Thomson, and I. J. Lean. 2023. Characterizing ruminal acidosis risk: A multiherd, multicountry study. J. Dairy Sci. 106:XXX–XXX. https://doi.org/10.3168/jds.2022-22571.

Golder, H. M., J. M. Thomson, S. E. Denman, C. S. McSweeney, and I. J. Lean. 2018. Genetic markers are associated with the ruminal microbiome and metabolome in grain and sugar challenged dairy heifers. Front. Genet. 9:62. https://doi.org/10.3389/fgene.2018.00062.

Henderson, G., F. Cox, S. Ganesh, A. Jonker, W. Young, Global Rumen Census Collaborators, and P. H. Janssen. 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci. Rep. 5:14567. https://doi.org/10.1038/srep14567.

Jami, E., and I. Mizrahi. 2012. Similarity of the ruminal bacteria across individual lactating cows. Anaerobe 18:338–343. https://doi.org/10.1016/j.anaerobe.2012.04.003.

Jami, E., B. A. White, and I. Mizrahi. 2014. Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. PLoS One 9:e85423. https://doi.org/10.1371/journal.pone.0085423.

Jiang, J., L. Ma, D. Prakapenka, P. M. VanRaden, J. B. Cole, and Y. Da. 2019. A large-scale genome-wide association study in U.S. Holstein cattle. Front. Genet. 10:412. https://doi.org/10.3389/fgene.2019.00412.

Jung, T.-Y., G.-R. Jin, Y.-B. Koo, M.-M. Jang, C.-W. Kim, S.-Y. Lee, H. Kim, C.-Y. Lee, S.-Y. Lee, B.-G. Ju, and H.-S. Kim. 2020. Deacetylation by SIRT1 promotes the tumor-suppressive activity of HINT1 by enhancing its binding capacity for β-catenin or MITF in colon cancer and melanoma cells. Exp. Mol. Med. 52:1075–1089. https://doi.org/10.1038/s12276-020-0465-2.

Kauter, A., L. Epping, T. Semmler, E.-M. Antao, D. Kannapin, S. D. Stoeckle, H. Gehlen, A. Lübke-Becker, S. Günther, L. H. Wieler, and B. Walther. 2019. The gut microbiome of horses: Current research on equine enteral microbiota and future perspectives. Anim. Microbiome 1:14. https://doi.org/10.1186/s42523-019-0013-3.

Kim, O.-S., Y.-J. Cho, K. Lee, S.-H. Yoon, M. Kim, H. Na, S.-C. Park, Y. S. Jeon, J.-H. Lee, H. Yi, S. Won, and J. Chun. 2012. Introducing EzTaxon-e: A prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. Int. J. Syst. Evol. Microbiol. 62:716–721. https://doi.org/10.1099/ijs.0.038075-0.

Li, F., T. C. Hitch, Y. Chen, C. J. Creevey, and L. L. Guan. 2019a. Comparative metagenomic and metatranscriptomic analyses reveal the breed effect on the rumen microbiome and its associations with feed efficiency in beef cattle. Microbiome 7:6. https://doi.org/10.1186/s40168-019-0618-5.

Li, F., C. Li, Y. Chen, J. Liu, C. Zhang, B. Irving, C. Fitzsimmons, G. Plastow, and L. L. Guan. 2019b. Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. Microbiome 7:92. https://doi.org/10.1186/s40168-019-0699-1.

Li, M., G. B. Penner, E. Hernandez-Sanabria, M. Oba, and L. L. Guan. 2009. Effects of sampling location and time, and host animal on assessment of bacterial diversity and fermentation parameters in the bovine rumen. J. Appl. Microbiol. 107:1924–1934. https://doi.org/10.1111/j.1365-2672.2009.04376.x.

Lima, F. S., G. Oikonomou, S. F. Lima, M. L. Bicalho, E. K. Ganda, J. C. de Oliveira Filho, G. Lorenzo, P. Trojacanec, and R. C. Bicalho. 2015. Prepartum and postpartum rumen fluid microbiomes: Characterization and correlation with production traits in dairy cows. Appl. Environ. Microbiol. 81:1327–1337. https://doi.org/10.1128/AEM.03138-14.

Liu, J., H. Li, W. Zhu, and S. Mao. 2019. Dynamic changes in rumen fermentation and bacterial community following rumen fluid transplantation in a sheep model of rumen acidosis: Implications for rumen health in ruminants. FASEB J. 33:8453–8467. https://doi.org/10.1096/fj.201802456R.

Mittal, M., C. Tiruppathi, S. Nepal, Y.-Y. Zhao, D. Grzych, D. Soni, D. J. Prockop, and A. B. Malik. 2016. TNFα-stimulated gene-6 (TSG6) activates macrophage phenotype transition to prevent inflammatory lung injury. Proc. Natl. Acad. Sci. USA 113:E8151–E8158. https://doi.org/10.1073/pnas.1614935113.

Myer, P. R. 2019. Bovine genome-microbiome interactions: Metagenomic frontier for the selection of efficient productivity in cattle systems. mSystems 4:e00103-19. https://doi.org/10.1128/mSystems.00103-19.

Nagaraja, T. G., and E. C. Titgemeyer. 2007. Ruminal acidosis in beef cattle: The current microbiological and nutritional outlook. J. Dairy Sci. 90:E17–E38. https://doi.org/10.3168/jds.2006-478.

National Research Council (NRC). 2001. Nutrient Requirements of Dairy Cattle. 7th ed. National Academies Press.

Price, A. L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38:904–909. https://doi.org/10.1038/ng1847.

Pruitt, H. C., D. J. Devine, and R. S. Samant. 2016. Roles of N-Myc and STAT interactor in cancer: From initiation to dissemination. Int. J. Cancer 139:491–500. https://doi.org/10.1002/ijc.30043.

Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahé. 2016. VSEARCH: A versatile open source tool for metagenomics. PeerJ 4:e2584. https://doi.org/10.7717/peerj.2584.

Sánchez, E., T. Lobo, J. L. Fox, M. Zeviani, D. R. Winge, and E. Fernández-Vizarra. 2013. LYRM7/MZM1L is a UQCRFS1 chaperone involved in the last steps of mitochondrial Complex III assembly in human cells. Biochim. Biophys. Acta Bioenerg. 1827:285–293. https://doi.org/10.1016/j.bbabio.2012.11.003.

Wallace, R. J., G. Sasson, P. C. Garnsworthy, I. Tapio, E. Gregson, P. Bani, P. Huhtanen, A. R. Bayat, F. Strozzi, F. Biscarini, T. J. Snelling, N. Saunders, S. L. Potterton, J. Craigon, A. Minuti, E. Trevisi, M. L. Callegari, F. P. Cappelli, E. H. Cabezas-Garcia, J. Vilkki, C. Pinares-Patino, K. O. Fliegerová, J. Mrázek, H. Sechovcová, J. Kopečný, A. Bonin, F. Boyer, P. Taberlet, F. Kokou, E. Halperin, J. L. Williams, K. J. Shingfield, and I. Mizrahi. 2019. A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. Sci. Adv. 5:eaav8391. https://doi.org/10.1126/sciadv.aav8391.

Weimer, P. J. 2015. Redundancy, resilience, and host specificity of the ruminal microbiota: Implications for engineering improved ruminal fermentations. Front. Microbiol. 6:296. https://doi.org/10.3389/fmicb.2015.00296.

Weimer, P. J., M. S. Cox, T. Vieira de Paula, M. Lin, M. B. Hall, and G. Suen. 2017. Transient changes in milk production efficiency and bacterial community composition resulting from near-total exchange of ruminal contents between high- and low-efficiency Holstein cows. J. Dairy Sci. 100:7165–7182. https://doi.org/10.3168/jds.2017-12746.

Weimer, P. J., D. M. Stevenson, H. C. Mantovani, and S. L. C. Man. 2010. Host specificity of the ruminal bacterial community in the dairy cow following near-total exchange of ruminal contents. J. Dairy Sci. 93:5902–5912. https://doi.org/10.3168/jds.2010-3500.

Weir, B. S. 1996. Genetic Data Analysis II. Sinauer Associates Inc.

Welkie, D. G., D. M. Stevenson, and P. J. Weimer. 2010. ARISA analysis of ruminal bacterial community dynamics in lactating dairy cows during the feeding cycle. Anaerobe 16:94–100. https://doi.org/10.1016/j.anaerobe.2009.07.002.

Wiegand, S., M. Jogler, and C. Jogler. 2018. On the maverick *Planctomycetes*. FEMS Microbiol. Rev. 42:739–760. https://doi.org/10.1093/femsre/fuy029.

Wiggans, G. R., J. B. Cole, S. M. Hubbard, and T. S. Sonstegard. 2017. Genomic selection in dairy cattle: The USDA experience. Annu. Rev. Anim. Biosci. 5:309–327. https://doi.org/10.1146/annurev-animal-021815-111422.

Xu, X., Q. Yan, Y. Wang, and X. Dong. 2017. NTN4 is associated with breast cancer metastasis via regulation of EMT-related biomarkers. Oncol. Rep. 37:449–457. https://doi.org/10.3892/or.2016.5239.

ORCIDS

H. M. Golder https://orcid.org/0000-0003-1298-3107
J. Thomson https://orcid.org/0000-0003-1921-0975
A. H. Smith https://orcid.org/0000-0003-4108-5613
E. Block https://orcid.org/0000-0003-2588-3236
I. J. Lean https://orcid.org/0000-0002-1045-7907