12-30-2015

Genomewide association study of liver abscess in beef cattle1,2

J. W. Keele
USDA-ARS, john.keele@ars.usda.gov

L. A. Kuehn
USDA-ARS, Larry.Kuehn@ars.usda.gov

T. G. McDaneld
USDA-ARS, tara.mcdaneld@ars.usda.gov

R. G. Tait
USDA-ARS

S. A. Jones
USDA-ARS, shuna.jones@ars.usda.gov

See next page for additional authors

Published December 30, 2015

Genomewide association study of liver abscess in beef cattle[1,2]

J. W. Keele,[3] L. A. Kuehn, T. G. McDaneld, R. G. Tait, S. A. Jones, B. N. Keel, and W. M. Snelling

USDA, ARS, U. S. Meat Animal Research Center, Clay Center, NE 68933

ABSTRACT: Fourteen percent of U.S. cattle slaughtered in 2011 had liver abscesses, resulting in reduced carcass weight, quality, and value. Liver abscesses can result from a common bacterial cause, *Fusobacterium necrophorum*, which inhabits rumen lesions caused by acidosis and subsequently escapes into the blood stream, is filtered by the liver, and causes abscesses in the liver. Our aim was to identify SNP associated with liver abscesses in beef cattle. We used lung samples as a DNA source because they have low economic value, they have abundant DNA, and we had unrestricted access to sample them. We collected 2,304 lung samples from a beef processing plant: 1,152 from animals with liver abscess and 1,152 from animals without liver abscess. Lung tissue from pairs of animals, 1 with abscesses and another without, were collected from near one another on the viscera table to ensure that pairs of phenotypically extreme animals came from the same lot. Within each phenotype (abscess or no abscess), cattle were pooled by slaughter sequence into 12 pools of 96 cattle for each phenotype for a total of 24 pools. The pools were constructed by equal volume of frozen lung tissue from each animal. The DNA needed to allelotype each pool was then extracted from pooled lung tissue and the BovineHD Bead Array (777,962 SNP) was run on all 24 pools. Total intensity (TI), an indicator of copy number variants, was the sum of intensities from red and green dyes. Pooling allele frequency (PAF) was red dye intensity divided TI. Total intensity and PAF were weighted by the inverse of their respective genomic covariance matrices computed over all SNP across the genome. A false discovery rate ≤ 5% was achieved for 15 SNP for PAF and 20 SNP for TI. Genes within 50 kbp from significant SNP were in diverse pathways including maintenance of pH homeostasis in the gastrointestinal tract, maintain immune defenses in the liver, migration of leukocytes from the blood into infected tissues, transport of glutamine into the kidney in response to acidosis to facilitate production of bicarbonate to increase pH, aggregate platelets to liver injury to facilitate liver repair, and facilitate axon guidance. Evidence from the 35 detected SNP associations combined with evidence of polygenic variation indicate that there is adequate genetic variation in incidence rate of liver abscesses, which could be exploited to select sires for reduced susceptibility to subacute acidosis and associated liver abscess.

Key words: acidosis, beef cattle, genomewide association study, liver abscess, rumenitis, subacute ruminal acidosis

J. Anim. Sci. 2016.94
doi:10.2527/jas2015-9887

INTRODUCTION

The rate of liver abscess in U.S. cattle has been estimated to be 13.7% harvested under 20 mo of age (McKeith et al., 2012) and 32% in cull dairy cows (Rezac et al., 2014b). Recently, the incidence of liver abscess is apparently increasing, especially in Holstein steers (Reinhardt and Hubbert, 2015). Liver abscesses were associated with a reduction in carcass weight and quality, which led to a decline in carcass value of between $20 and $80 USD per animal affected (Brown and Lawrence, 2010). Metabolic conditions

[1]The authors would like to recognize Sandra Nejezchleb, Steve Simcox, Tammy Sorensen, and Shanda Watts for technical assistance.

[2]Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The USDA is an equal opportunity provider and employer.

[3]Corresponding author: john.keele@ars.usda.gov

Received September 18, 2015.
Accepted November 3, 2015.

that predispose cattle to liver abscess and acute and subacute ruminal acidosis negatively impacted feed intake and growth in the feedlot (Owens et al., 1998; Rezac et al., 2014a), causing additional decreases in profit as a result of BW gain not realized. *Fusobacterium necrophorum* is commonly found in the rumen of cattle and invades rumen lesions caused by acidic conditions, enters the blood, is filtered by the liver, and infects the liver, leading to abscesses (Tadepalli et al., 2009). Ruminal acidosis and liver abscess will respond to artificial selection if they are affected by genetic variation. Ruminal acidosis depends, to some extent, on ruminal microbes, and there are differences in rumen microbial communities between ruminant species. The composition of cellulolytic bacteria in the rumen was different between cattle and bison (Varel and Dehority, 1989). Differences in the rumen microbiome were observed among steers selected to be extreme for feed efficiency (Myer et al., 2015); however, host differences between selected groups might not be genetic. In the mouse, host genetics influences gut microbiota, which, in turn, modulates host metabolism (McKnite et al., 2012). We hypothesize that host factors exist that affect susceptibility of cattle to subacute acidosis and subsequent liver abscess. Our objective was to associate SNP with liver abscess incidence in cattle in a beef processing plant.

MATERIALS AND METHODS

The animals sampled in this experiment did not originate from and were not under the control of the U.S. Meat Animal Research Center (**USMARC**). Therefore, animal procedures were not reviewed and approved by the USMARC Animal Care and Use Committee. Animals were humanely harvested in a federally inspected packing plant and samples were collected postmortem.

Lung Sampling

We sampled lungs as a source of DNA because they are a rich source of DNA; lungs are of low economic value, so we had unrestricted access to collect a sample from them; and it was relatively easy to identify the lung and liver for an animal based on position on the table. We sampled 2,304 lungs from a large central Nebraska commercial beef processing plant across 10 d over a 5-wk period with 2 sampling days per week. Fifteen to 20 lots (or pens) were processed per 8-h shift with an average of 150 animals in a lot. We sampled from the plant work shift from 0600 to 1500 h. Generally, there were multiple case/control pairs per lot sampled. Lungs were sampled from the moving viscera table when livers with different phenotypes (abscess vs. no abscess) were in close proximity. Requiring variation in liver abscess in close proximity on the viscera table before extreme samples were collected increased the probability that divergent liver phenotypes were matched by source and, by proxy, breed ancestry. Liver abscesses were identified by 1 veterinarian (i.e., coauthor S.A. Jones) with extensive experience in liver pathology. Livers were scored as abscessed if they had moderate or severe abscesses and were condemned for abscesses by the Food Safety and Inspection Service (**FSIS**) inspectors. Livers were scored as not having an abscess (control) if there were no visible abscesses and they were not condemned by the FSIS inspector.

Sample Preparation and Pooling

Sample preparation was similar to Keele et al. (2015). Five- to 10-g lung samples from the viscera of animals with liver abscesses (case) or no liver abscess (control) were placed on ice immediately after collection, frozen (–20°C) within 10 h after collection, and stored frozen for 5 to 14 mo until processing for pooling. One 1.2 mm diameter × 5 mm length cylindrical core per animal was taken from a partially thawed lung sample for each of 96 animals of the same phenotype (abscess or no abscess) and combined in a single tube to create a pool. Twelve pools were constructed for each phenotype (abscess or no abscess) for a total of 24 pools.

The combination of the pooling and sampling processes ensured that abscess and no-abscess pools were matched by proximity in the processing line. Samples were labeled by collection order within phenotype (abscess or no abscess). The first 96 abscessed animals made up the first abscess pool, the second 96 animals made up the second pool, and so on. Likewise, a similar process was used to make the no-abscess pools. Deoxyribonucleic acid was isolated from the pooled lung sample by a standard salt extraction procedure. In a pilot project, we compared this procedure of pooling lung samples before extraction with pooling of DNA after extraction of individual lung samples. We found that pool construction variation was not affected by the method used to construct pools (data not shown).

Deoxyribonucleic acid from pools was sent to Neogen Corporation (Lincoln, NE) for analysis with the Illumina BovineHD Bead Array (777,962 SNP; Illumina, Inc., San Diego, CA).

Statistical Analyses

Two variables were analyzed: total intensity (**TI**) of red and green dyes (Steemers et al., 2006) and pooling allele frequency (**PAF**). Pooling allele frequency (Peiris et al., 2011) was computed as red intensity divided by TI. Pooling allele frequency is an estimate of the frequency for the A allele. Total intensity estimates copy number

variants (**CNV**). Copy number variants are indicated when there are concordant increases or decreases in the $\log_2$ of the TI divided by the reference TI for individual animals for multiple neighboring SNP (Wu et al., 2015). In the current study, we assume that an association of TI with liver abscesses in pools might indicate an association with CNV; however, the CNV is most likely just a marker in linkage disequilibrium with causative variant, which is not necessarily a CNV. We are not requiring concordant changes in TI among multiple neighboring SNP, so TI might also reflect differences in hybridization efficiency because of cryptic SNP in the 5′ region of the 50-bp probe adjacent to the target SNP (Steemers et al., 2006) instead of CNV. Population stratification and outliers for TI and PAF were visualized using a neighbor-joining tree constructed from Euclidean distances among pools computed across all SNP using the ape package and the dist() function of R (R Core Team, 2014).

Initially, SNP were eliminated from the analysis if there were missing data for any of the pools. For the analysis of PAF (but not TI), SNP were eliminated if minor PAF was less than 1%: 1 – PAF if PAF > 0.5 or PAF otherwise. For TI, 3,702 SNP had missing data for red or green intensity, which we eliminated from further analysis, leaving 774,260 SNP. For PAF, there were 6 additional SNP with missing data because TI was 0, which is the denominator when computing PAF; hence, PAF was undefined for these SNP. Consequently, for PAF, 3,708 SNP were eliminated from the analysis because of missing data and 154 were eliminated because of minor allele frequency less than 1%, leaving 774,100 SNP for analysis.

Effects of phenotype (abscess or no abscess) on individual SNP PAF were estimated and *P*-values computed based on mixed model methods similar to McDaneld et al. (2014). Briefly, the average covariance matrix (**A**) among pools was computed across all SNP to account for population stratification and technical errors common to all SNP on an individual pooled sample on a SNP array. The average covariance matrix for PAF used the same methodology as Keele et al. (2015). The average covariance matrix for TI was computed as the covariance among standardized deviations of TI from the SNP mean divided by the SNP SD. Deviating TI from the SNP mean removed SNP variation from the estimated covariance and standardizing (dividing by) with the SNP SD ensured that SNP were given equal weight in computing the covariance. Single nucleotide polymorphisms with 0 SD were removed from further analysis ($n = 1$). The *F* test with numerator and denominator df of 1 and 22 (24 pools – 1 df for the difference between pools – 1 df for the mean), respectively, were used to compute *P*-values using the pf() function in the R software package. The SNP were considered significant if they achieved a false discovery rate (**FDR**) ≤ 5% (Benjamini and Hochberg, 1995). The list of *P*-values was interrogated for evidence of population stratification using the lambda value from genomic control and quantile–quantile plots (Devlin and Roeder, 1999; Zhao, 2007, 2015). Genes within 50 kbp from significant SNP were identified using BioMart (Smedley et al., 2015), and genes were associated with Kyoto Encyclopedia of Genes and Genomes (**KEGG**; Kanehisa and Goto, 2000; Kanehisa et al., 2014) pathways using The Database for Annotation, Visualization and Integrated Discovery (**DAVID**) version 6.7 (Huang et al., 2009).

The effect of genetic variance on quantile–quantile plots was evaluated using simulation based on mixtures of 99.5% central and 0.5% noncentral χ^2 distributions. Three different *P*-value lists were simulated. For each list, 774,072 (99.5%) were simulated from a central χ^2 distribution with 1 df and 3,890 (0.5%) were drawn from a noncentral χ^2 distribution with 1 df and noncentrality parameter equal to 2.5, 5, or 10, with the noncentrality parameter held constant within the list. The *P*-values were computed as the integral of the central χ^2 distribution from the simulated value to infinity using pchisq() in R. Quantile–quantile plots were computed using gap (Zhao, 2007, 2015).

RESULTS

Population Structure

Population structure was characterized using neighbor-joining trees (Fig. 1 and 2). For both PAF and TI, the trees did not have a central node with equal-length branches to individual pools (star configuration). We used mixed model methods (see Materials and Methods) to account for population structure at the expense of power (Segura et al., 2012). If our trees had been a star configuration, we would most likely have identified more SNP because of increased power. The mixed model method is expected to reduce bias (false positives) at the expense of reduced power (Segura et al., 2012). The trees indicate substantial clustering of pools with a quite different tree for PAF compared with TI (Fig. 1 and 2). The same 2 control pools were divergent for both PAF and TI. Our method gives these divergent control pools less emphasis in the analysis, which is true regardless of the reason for divergence. Certainly, if these pools were not divergent, the experiment would have had greater power, and we would likely have detected more associated SNP if the pools were not divergent (Segura et al., 2012). The tree structure is ambiguous as to whether the outliers are technical problems or genetic outliers. The fact that the divergent pools are similar to each other argues against

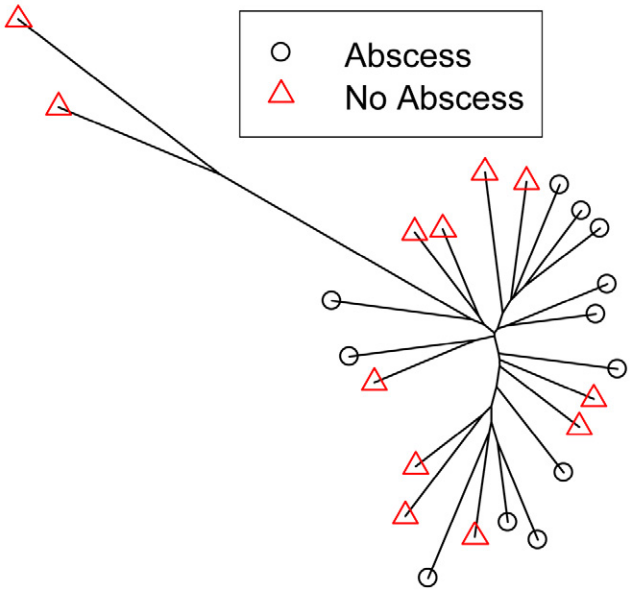

Figure 1. Neighbor-joining tree (unrooted) constructed from Euclidean distances among pooling allele frequencies for the 24 pools.

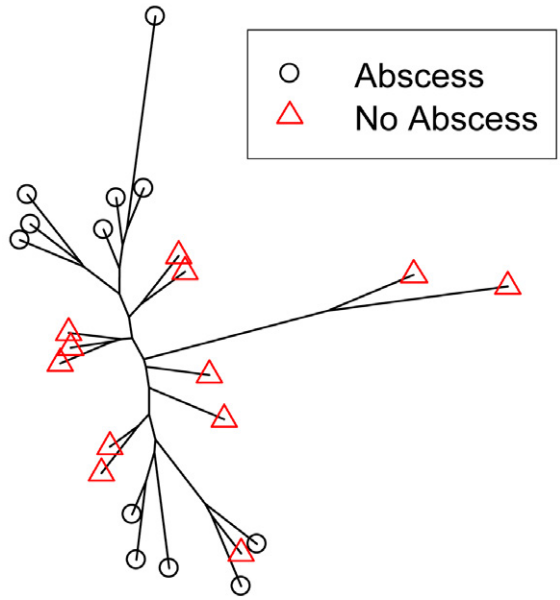

Figure 2. Neighbor-joining tree (unrooted) constructed from Euclidean distances among total intensity (sum or intensity for red and green dyes) for the 24 pools.

a technical problem and supports the possibility that they are genetically divergent. However, the divergent control pools were not near their corresponding case pools, which were sampled near them on the viscera table, which suggests a technical problem. The PAF tree is similar to a star configuration (nearly equal branch lengths from a central node) with the exception of the 2 divergent pools (Fig. 1). In contrast, the TI tree shows the liver abscess pools primarily at opposite ends of a fairly long axial branch with most of the control pools distributed between the 2 extremes (Fig. 2). One exception is a control pool clustered closely with 2 abscess pools near the bottom of the tree.

Single Nucleotide Polymorphism Association

Genomic control lambda (Devlin and Roeder, 1999; Zhao, 2007, 2015) estimates the discrepancy of *P*-value distributions from the expected uniform distribution under the null QTL model. Lambda values near 1 support the null QTL model and population homogeneity, values greater than 1 indicate population stratification or true QTL, and values less than 1 indicate *P*-value deflation, which is usually the result of overadjustment for population stratification. Genomic control lambda was 0.9 for PAF and 0.8 for TI, indicating that mixed model methods slightly overcorrected for population stratification, resulting in a deflated distribution of *P*-values relative to the distribution expected when there is no true association. If the lambda value was closer to 1, we would have expected to observe a list of SNP with FDR ≤ 5% longer than 35 with everything else being equal. However, we observed substantial curvature in the quantile–quantile plots of $-\log_{10}$(*P*-value; Fig. 3 [PAF] and Fig. 4 [TI]) for liver abscess, which is consistent with there being genetic variation in susceptibility to liver abscesses. Simulation results indicate that the amount of curvature or change in slope for quantile–quantile plots is related to the amount of genetic variation in the trait or sample size (Fig. 5). In our simulations, we varied genetic variation by changing the noncentrality parameter for the χ^2 distribution. The number of SNP achieving a FDR of 5% for association with liver abscess was 15 for PAF (Table 1) and 20 for TI (Table 2). We observed no intersection between the PAF and TI lists, which is to be expected for 2 reasons. First, both SNP lists are relatively short, so the probability of overlap is quite low given the lack of extreme *P*-values dramatically lower than the significance threshold, indicating modest power even for the most significant SNP. Second, significant TI associations are most likely the result of non-Mendelian variation, which would not necessarily be associated with Mendelian variation predominant with PAF.

Functional Analysis

Chromosomal regions encompassing 50 kbp on either side of 35 significant SNP were submitted to BioMart (Smedley et al., 2015), which identified 49 genes. Of these, 39 were found in the DAVID bovine gene list (Huang et al., 2009) and 15 were mapped to KEGG pathways (Kanehisa and Goto, 2000; Kanehisa et al., 2014; Supplementary Table S1 [see the online ver-

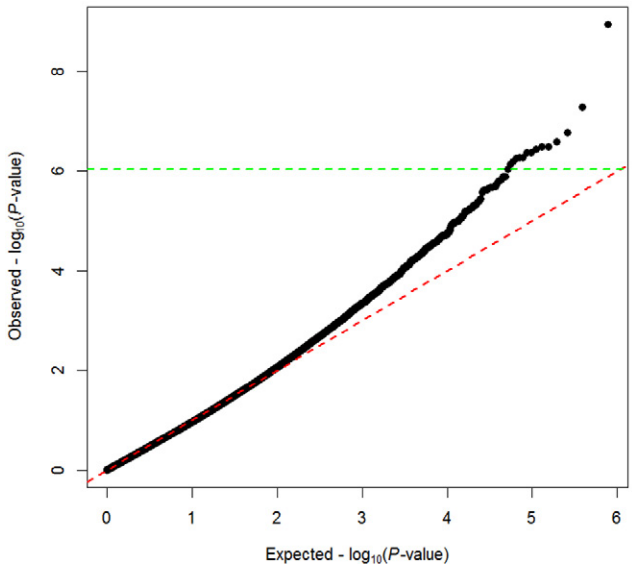

Figure 3. Quantile–quantile plot for $-\log_{10}(P\text{-values})$ for pooling allele frequency. The red dashed line indicates where observed and expected values are equivalent. The green dashed line indicates the threshold where *P*-values achieve a false discovery rate of 5%.

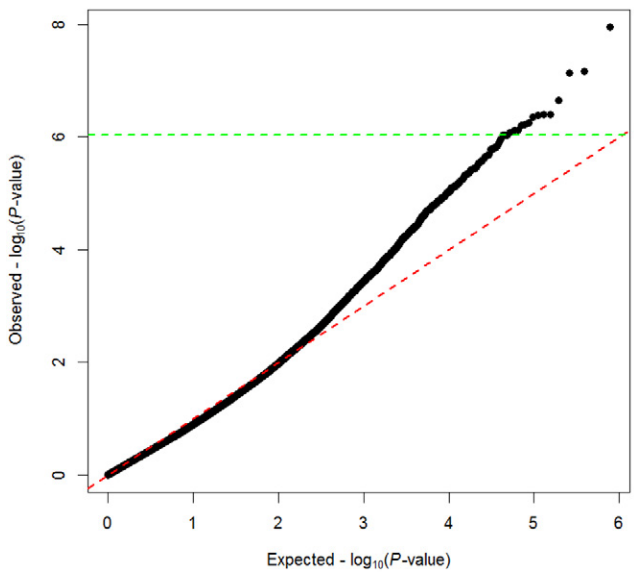

Figure 4. Quantile–quantile plot for $-\log_{10}(P\text{-values})$ for total intensity or the sum of intensities for red and green dyes. The red dashed line indicates where observed and expected values are equivalent. The green dashed line indicates the threshold where *P*-values achieve a false discovery rate of 5%.

sion of the article at http://journalofanimalscience.org]). No KEGG pathways were enriched with genes from our list. Identified pathways were involved in sensing and regulating pH in the gastrointestinal tract; repair of liver damage; neurological development and function; axon guidance; gap junction; taste; smell; taste transduction; calcium signaling, nucleotide-binding oligomerization domain (NOD)-like receptor, Notch, chemokine, adipocytokine, Hedgehog, and insulin signaling; antigen processing; spliceosome; endocytosis; and phagocytosis.

DISCUSSION

Gene Functions

Thirty-five SNP on 17 autosomal chromosomes and the X chromosome (sex distribution in pools is unknown) were associated with liver abscess with a FDR of 5%. The expected number of false positives out of the 35 positives is 1.75 (0.05 × 35). Two *P*-value lists were interrogated, 1 for PAF and 1 for TI, so we divide 0.05 by 2 to evaluate the distribution of false positives and estimate the upper confidence limit at the 97.5 percentile. Assuming a Poisson distribution with a lambda value of 0.05 times the number of positives for the upper 97.5 percentile [1 – (0.05/2)] of the false positive distribution would be 3 for both PAF and TI, indicating that at least 29 [35 – (2 × 3)] of the 35 significant associations are true positives with high probability. Hence, the expected value of true positives is 33.25 with a lower 95% confidence limit of 29. An alternative and equally valid approach would be to combine the *P*-value lists and perform 1 FDR analysis. This approach yields 34 positive SNP with an expected value of 32.3 true positives and a lower 95% confidence limit of 30, so both approaches for estimating the 95% confidence interval of number of true positives yields similar inferences for these data.

Curvilinear quantile–quantile plots of *P*-values (Fig. 1, 2, and 3) indicate substantial polygenic variation in liver abscess, which was not declared significant with the number of pools and pool sizes used in this experiment. Our results (evidence of both detected SNP effects and cryptic polygenic variation) indicate that liver abscess is partially under genetic control and would be expected to respond to selection of sires based on genetic evaluation using phenotyped relatives and/or SNP.

There is little doubt that an animal that develops a liver abscess will likely experience decreased productivity; hence, one might expect that the beef industry's current efforts to select for postweaning BW gain in the seed stock sector might be decreasing the incidence of liver abscess as a correlated response. However, segmentation of the beef industry into seed stock and commercial sectors with selection occurring only in the seed stock sector may be preventing this desirable correlated outcome. Conversely, selection for growth of bulls in the seed stock sector could increase liver abscesses of finishing steers and heifers in the commercial sector. When we select for growth, we increase appetite and feed intake. Cattle with higher intakes on a feedlot diet will be more susceptible to acidosis. Most selection pressure happens at the seed stock level in growing bulls and replacement heifers, and these animals are fed and managed to main-

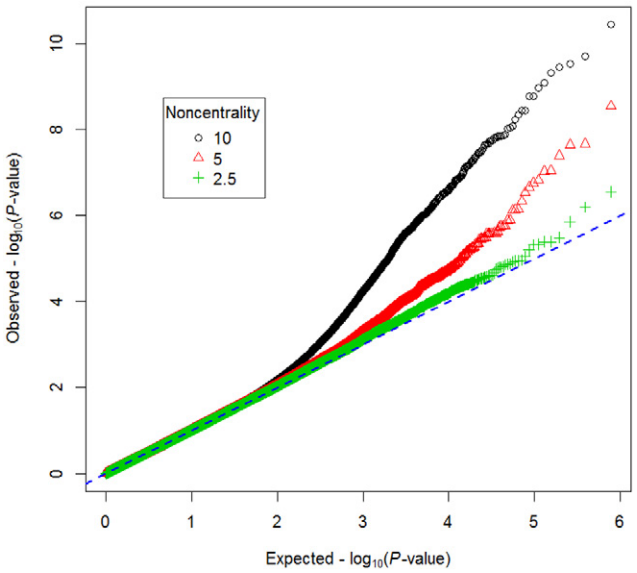

Figure 5. Quantile–quantile plot for simulated $-\log_{10}(P\text{-values})$ based on mixtures of 99.5% central and 0.5% noncentral χ^2 distributions with 1 df. The *P*-values are the area in the upper tail of the central χ^2 distribution above the simulated χ^2 values. The blue dashed line indicates where observed and expected values are equivalent. The legend indicates the values used for noncentrality for 0.5% of the SNP; for the remaining 99.5% of the SNP, noncentrality was 0.

tain good health. So we do not "push" the replacement bulls and heifers in a seed stock setting to the same extent as feedlot animals on a higher energy diet. So in the selection candidates, selection for increased growth is possible, without the negative feedback from acidosis and liver abscesses. It creates the argument that data from feedlots under typical commercial feeding programs (as was done in this experiment) is required to train predictions for genomic selection of seed stock, instead of data from replacement bulls and heifers, which is the majority of data at the moment. Steers with a greater genetic propensity for growth and the correlated trait of intake will be more susceptible to liver abscesses. This concept is supported by evidence in this study that polymorphisms in olfactory receptor genes are associated with liver abscess and the possibility that olfactory receptors might influence intake via sensitivity (or the lack there of) to dietary components (see Discussion).

Functions of 15 genes near significant SNP encompassed many diverse pathways and biological functions (Supplementary Table S1; see the online version of the article at http://journalofanimalscience.org), which is not surprising given the complexity of the phenotype. Liver abscess is known to be impacted by acute and subacute acidosis, which results from low pH caused by rapid metabolism of soluble carbohydrate in the rumen and the overgrowth of a few species of ruminal microbes that are tolerant of low pH. *Fusobacterium necrophorum* is a major bacterial cause of liver abscess (Tadepalli et al., 2009). *Fusobacterium necrophorum* is a common bacterium within the rumen; however, under conditions of low rumen pH, it invades rumen abscesses, gains access to the circulatory system, is filtered by the liver, and causes infection and, ultimately, abscesses in the liver. The liver is well defended by immune cells; however, when many bacteria escape the rumen, they overwhelm these defenses.

In addition to a soluble carbohydrate diet, fluctuation in meal sizes, varying daily feed intake, and small particle size in the diet can predispose the animal to bouts of acidosis and liver abscess; hence, genetic defects in appetite control or satiety signals may also affect liver abscess incidence.

Two genes near significant SNP, acid-sensing (proton-gated) ion channel 2 (*ASIC2*) and thromboxane A2 receptor (*TBXA2R*), have functions related to acidosis and/or liver function and both are abundantly expressed along the gastrointestinal tract of cattle based on the Bovine Gene Atlas (Harhay et al., 2010). First, *ASIC2* (or *ACCN1*) is one of the acid sensing ion channel proteins expressed throughout the gastrointestinal tract and the liver to detect unhealthy deviations in acidity and maintain pH homoestasis (Holzer, 2015). Second, *TBXA2R* is part of a signaling pathway that aggregates platelets at the site of liver injury and facilitates liver repair (Minamino et al., 2015).

One of the defenses in the liver against invading bacteria includes phagocytosis by leukocytes and Kupffer cells. Nearby a significant SNP is a gene, phosphatidylinositol-4-phosphate 5-kinase, type I, γ (*PIP5K1C*), that codes for a protein involved in Fc γ R-mediated phagocytosis (Supplementary Table S1; see the online version of the article at http://journalofanimalscience.org). Other pathways that interact with Fc γ R-mediated phagocytosis are calcium signaling, regulation of actin cytoskeleton, endocytosis, and lysosome, all of which contain genes near our significant SNP. Other genes near significant SNP are in the leukocyte transendothelial migration pathway, which is the process by which leukocytes get from the blood stream into tissues such as the liver and rumen wall.

Acidosis has been reported to trigger expression of solute carrier family 38, member 3 (*SLC38A3*) to produce SNAT3 in the kidney, which transports glutamine from the blood into the kidney, resulting in the production of bicarbonate and normalization of pH (Balkrishna et al., 2014). The Bovine Gene Atlas did not show large amounts of *SLC38A3* in the kidney or other gastrointestinal tissues, which might be a reflection that the animals sequenced to generate the expression tags were not undergoing acidosis so their expression would not be expected to be high.

The olfactory receptors such as *OR8J1* (Table 1) may be important for the animal sensing the food that

Table 1. Associations between SNP pooling allele frequency[1] (PAF) and liver abscesses achieving false discovery rate of 5%

			PAF						Gene[4]	
BTA	Position, bp	SNP[2]	Abscess	No abscess	MSE[3]	*t*-distribution	*P*-value	Gene	Starting position, bp	Ending position, bp
1	90,138,593	rs43248252	0.19	0.12	0.053	6.98	5.27×10^{-7}	XR_804373.1	90,120,141	90,121,253
1	117,247,146	rs137456530	0.16	0.11	0.029	7.50	1.70×10^{-7}	*LOC782298*	117,222,990	117,223,951
2	47,326,757	rs135856611	0.40	0.30	0.098	7.21	3.19×10^{-7}	*KIF5C*	47,311,871	47,472,004
2	107,508,288	rs43315181	0.40	0.26	0.200	6.96	5.54×10^{-7}	*PRKAG3*	107,507,841	107,517,485
7	21,558,133	rs136477409	0.23	0.14	0.083	7.07	4.30×10^{-7}	*TBXA2R*	21,556,155	21,569,886
8	24,743,994	rs133269746	0.16	0.09	0.055	6.84	7.16×10^{-7}	*SLC24A2*	24,487,108	24,779,293
15	47,296,349	rs109184225	0.51	0.37	0.184	7.16	3.58×10^{-7}	*LOC786995*	47,310,391	47,313,308
15	57,340,890	rs41776932	0.50	0.31	0.286	8.06	5.25×10^{-8}	*MYO7A*	57,315,857	57,419,715
15	80,515,623	**rs134787048**	0.15	0.09	0.016	10.03	1.15×10^{-9}	*OR8J1*	80,515,014	80,515,961
16	11,424,152	rs42601618	0.28	0.21	0.042	7.30	2.58×10^{-7}	*LOC104970159*	12,355,176	12,378,333
18	48,806,574	rs41885703	0.25	0.40	0.206	–7.07	4.29×10^{-7}	*LGALS4*	48,803,147	48,810,095
21	68,600,417	rs109014822	0.35	0.25	0.107	6.89	6.35×10^{-7}	*HSP90AA1*	68,595,870	68,601,237
22	50,708,448	**rs42019238**	0.18	0.12	0.045	6.97	5.42×10^{-7}	*SLC38A3*	50,702,769	50,717,523
X	16,094,115	rs135306367	0.31	0.19	0.164	6.74	9.03×10^{-7}	*MBNL3*	16,104,214	16,224,696
X	148,694,034	rs134648412	0.38	0.24	0.205	7.20	3.26×10^{-7}	*LOC505052*	148,690,204	148,703,791

[1]Pooling allele frequency is red dye intensity divided by the sum of red and green dye intensities.

[2]SNP shown in bold are missense (nonsynonymous) mutations; all others are intron or intergenic mutations.

[3]MSE = mean squared error.

[4]Positions are on the UMD3.1 build (https://ccb.jhu.edu/bos_taurus_assembly.shtml, accessed Oct. 9, 2015).

it is eating and whether or not consuming more food is compatible with maintaining pH homeostasis in the gastrointestinal tract. Genetic mutations in olfactory receptors might compromise the animal's ability to sense what they are eating or adjust their food consumption and selectivity accordingly.

Distribution of P-*Values*

When genetic variation in a trait is polygenic or influenced by a large number of loci, we expect the *P*-values produced by a genomewide association study to follow a mixture distribution. As a result, 1 kernel or component distribution corresponding to SNP with 0 effect would follow a uniform distribution. The other kernel corresponding to SNP in linkage disequilibrium with loci affecting the trait would have a nonuniform distribution with greater density for smaller *P*-values. A quantile–quantile plot of $-\log_{10}$(*P*-values) for a polygenic trait is expected to show curvature with the slope transitioning from 1 to a larger value. As the genetic variation or sample size for the experiment increases, the curvature becomes sharper and the slope transition becomes more rapid as demonstrated by our simulations presented in Fig. 5. Higher noncentrality, which is a proxy for genetic variance, sample size, and allele frequency, results in increased power; noncentrality is proportional to genetic variance (or QTL effect squared), inversely proportional to sample size, and higher for intermediate allele frequencies than for alleles that are nearly fixed at 0 or 1. Our simulation results demonstrate that quantile–quantile plots are fairly sensitive to modest amounts of genetic variation relative to sample size even though the power to detect this polygenic variation at the genomewide level would be very small, which explains why quantile–quantile plots can have substantial curvature even though few SNP are detected at 5% FDR. We illuminate this by comparing the power for detecting a SNP at the 5% FDR level with and without adjusting for multiple testing. Using a nominal Type I error rate of $6.46 \times 10^{-8} = 0.05/774{,}260$ to adjust for testing all 774,260 SNP results in very small power values of 6.5×10^{-5}, 7.63×10^{-4}, and 1.24×10^{-4} for noncentrality values of 2.5, 5, and 10, respectively. On the other hand, using a nominal Type I error rate of 0.05 without adjusting for multiple testing results in power values of 0.35, 0.61, and 0.89 for noncentrality values of 2.5, 5, and 10.

Liver abscess is a polygenic trait with a relatively small number of loci with modest effects that are detected in this experiment. Also, curvature in the quantile–quantile plots (Fig. 3 and 4) indicates that there are a large number of loci that are individually undetectable but would likely be detectable with larger experiments.

A feasible and cost-effective genetic evaluation system to reduce incidence of acidosis can be developed from functionally focused genetic variants in seed stock cattle that can be identified from differences in PAF among pools of commercial cattle with different phenotypes (liver abscess or no liver abscess). Reduced

Table 2. Associations between SNP total intensity[1] and liver abscesses achieving false discover rate of 5%

			Total intensity						Gene[3]	
BTA	Position, bp	SNP	Abscess	No abscess	MSE[2]	*t*-distribution	*P*-value	Gene	Start position, bp	Ending position, bp
1	58,732,548	rs133447892	1.20	0.72	0.0367	7.06	4.44×10^{-7}	*SIDT1*	58,661,940	58,775,515
4	100,365,566	rs133850673	1.10	0.86	0.0107	6.73	9.20×10^{-7}	*FAM180A*	100,331,402	100,345,439
5	35,186,437	rs110025204	1.14	0.70	0.0341	6.76	8.49×10^{-7}	*ANO6*	35,056,795	35,291,725
6	57,219,374	rs110926019	1.00	0.69	0.0168	6.81	7.70×10^{-7}	*ARAP2*	57,068,266	57,266,996
6	115,311,236	rs109917213	0.83	0.50	0.0164	7.10	4.00×10^{-7}	*C1QTNF7*	115,275,797	115,404,779
8	65,656,612	rs135987354	1.02	0.69	0.0162	7.11	3.94×10^{-7}	*INVS*	65,597,389	65,734,344
9	37,391,299	rs41969895	1.00	0.71	0.0122	7.37	2.26×10^{-7}	*LOC104969531*	37,375,035	37,533,536
10	42,312,864	rs41971514	1.31	0.97	0.0198	6.72	9.37×10^{-7}	*XR_809266.1*	42,129,616	42,269,782
10	66,121,449	rs109783238	1.08	0.71	0.0245	6.66	1.08×10^{-7}	*DDHD1*	65,905,385	65,971,532
10	78,223,140	rs109144272	1.79	1.29	0.0314	7.93	6.83×10^{-8}	*FUT8*	77,831,015	78,228,253
10	89,888,962	rs109375524	1.08	1.34	0.0127	–6.59	1.26×10^{-6}	*ALKBH1*	89,981,032	89,914,287
10	91,321,589	rs42649299	1.41	1.08	0.0110	8.82	1.13×10^{-8}	*XR_236247.1*	91,252,775	91,255,151
11	63,947,037	rs109236099	1.11	0.67	0.0248	7.90	7.27×10^{-8}	*XR_809813.1*	63,778,736	63,779,503
11	65,909,126	rs110946939	1.63	1.23	0.0269	6.79	8.09×10^{-7}	*ETAA1*	65,813,280	65,827,092
12	19,743,828	rs135470573	1.72	1.24	0.0387	6.73	9.23×10^{-7}	*LOC101904762*	19,670,646	19,686,926
15	72,437,026	rs110134302	1.89	1.42	0.0339	7.09	4.14×10^{-7}	*LRRC4C*	71,243,139	73,680,560
15	72,743,769	rs110308201	0.78	0.45	0.0183	6.81	7.59×10^{-7}	*XR_811673.1*	72,741,075	72,744,934
18	2,231,562	rs109874485	1.28	0.93	0.0197	6.92	6.00×10^{-7}	*FA2H*	2,182,821	2,237,316
18	15,749,436	rs136217287	1.58	1.13	0.0336	6.90	6.27×10^{-7}	*ITFG1*	15,625,548	15,928,962
19	17,433,304	rs134572895	1.71	1.27	0.0308	6.96	5.53×10^{-7}	*ASIC2*	16,353,062	17,563,070

[1]Total intensity is the sum of red and green dye intensity.

[2]MSE = mean squared error.

[3]Positions are given on the UMD3.1 build (https://ccb.jhu.edu/bos_taurus_assembly.shtml; accessed Oct. 9, 2015).

ruminal acidosis would be associated with increased performance, improved carcass quality, reduced cost, and increased revenue. Phenotypic data (liver abscess incidence) used to predict genetic evaluations would need to be from feedlot cattle fed high-concentrate diets, because seed stock cattle are managed so they are less likely to express the phenotype. Commercial feedlot cattle harvested in packing plants are not likely to be close first- or second-degree relatives to seed stock candidates being evaluated, so genomic evaluations based on unweighted genotypes from standard SNP genotyping platforms may not be accurate. Predictive genomic relationships would need to be computed using genotypes of SNP targeted to chromosomal regions containing genes coding for proteins that functionally affect liver abscess or SNP in tight linkage disequilibrium with these genes. It might be possible to impute genotypes for functional loci from standard genotyping platforms depending on allele frequencies and linkage disequilibrium. Maybe a selection of SNP from the emerging functional array would work, but additional research needs to be done to identify the best collection of SNP. It seems that the less functional the SNP platform, the less genetic distance there can be between the seed stock candidates and the commercial pools used to accurately predict their merit. Research is needed to identify functionally targeted SNP, those SNP that are in tight linkage disequilibrium with variation underlying liver abscesses.

Costs of genetic evaluation could be reduced if it is possible to accurately estimate the genetic relationship between the functional regions of the seed stock candidates and the analogous functional regions of commercial feedlot cattle with pooled samples.

Finally, it has long been known that selection response is proportional to additive genetic variation. Selection of phenotypic extremes (as was done with the current study) obscures our ability to estimate the amount of genetic variation accounted for by identified SNP. The genetic variation that is associated with the SNP we have identified is most likely incomplete.

Conclusions

Incidence of liver abscesses in feedlot cattle fed high levels of soluble carbohydrate is a heritable trait with a small number of loci (n = approximately 33) having moderate effects and a large number (over 1,000) with small effects. Genetic variation in incidence of liver abscess includes variation in genes that regulate gastrointestinal pH and liver repair following injury and affect function of immune cells. Acute and subacute ruminal acidosis would likely decrease in response to selection of bulls using genetic evaluations

based on associated SNP and liver abscess incidence in related feedlot cattle recorded at the packing plant.

LITERATURE CITED

Balkrishna, S., A. Bröer, S. M. Welford, M. Hatzoglou, and S. Bröer. 2014. Expression of glutamine transporter Slc38a3 (SNAT3) during acidosis is mediated by a different mechanism than tissue-specific expression. Cell. Physiol. Biochem. 33:1591–1606. doi:10.1159/000358722.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J.R. Stat. Soc. 57:289–300.

Brown, T. R., and T. E. Lawrence. 2010. Association of liver abnormalities with carcass grading performance and value. J. Anim. Sci. 88:4037–4043. doi:10.2527/jas.2010-3219.

Devlin, B., and K. Roeder. 1999. Genomic control for association studies. Biometrics 55:997–1004. doi:10.1111/j.0006-341X.1999.00997.x.

Harhay, G. P., T. P. Smith, L. J. Alexander, C. D. Haudenschild, J. W. Keele, L. K. Matukumalli, S. G. Schroeder, C. P. Van Tassell, C. R. Gresham, S. M. Bridges, S. C. Burgess, and T. S. Sonstegard. 2010. An atlas of bovine gene expression reveals novel distinctive tissue characteristics and evidence for improving genome annotation. Genome Biol. 11(10):R102. doi:10.1186/gb-2010-11-10-r102.

Holzer, P. 2015. Acid-sensing ion channels in gastrointestinal function. Neuropharmacology 94:72–79. doi:10.1016/j.neuropharm.2014.12.009.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4:44–57. doi:10.1038/nprot.2008.211.

Kanehisa, M., and S. Goto. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28:27–30.

Kanehisa, M., S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. 2014. Data, information, knowledge and principle: Back to metabolism in KEGG. Nucleic Acids Res. 42:D199–D205. doi:10.1093/nar/gkt1076.

Keele, J. W., L. A. Kuehn, T. G. McDaneld, R. G. Tait, S. A. Jones, T. P. Smith, S. D. Shackelford, D. A. King, T. L. Wheeler, A. K. Lindholm-Perry, and A. K. McNeel. 2015. Genomewide association study of lung lesions in cattle using sample pooling. J. Anim. Sci. 93(3):956–964. doi:10.2527/jas.2014-8492.

McDaneld, T. G., L. A. Kuehn, M. G. Thomas, W. M. Snelling, T. P. Smith, E. J. Pollak, J. B. Cole, and J. W. Keele. 2014. Genomewide association study of reproductive efficiency in female cattle. J. Anim. Sci. 92:1945–1957. doi:10.2527/jas.2012-6807.

McKeith, R. O., G. D. Gray, D. S. Hale, C. R. Kerth, D. B. Griffin, J. W. Savell, C. R. Raines, K. E. Belk, D. R. Woerner, J. D. Tatum, J. L. Igo, D. L. VanOverbeke, G. G. Mafi, T. E. Lawrence, R. J. Delmore Jr., L. M. Christensen, S. D. Shackelford, D. A. King, T. L. Wheeler, L. R. Meadows, and M. E. O'Connor. 2012. National Beef Quality Audit – 2011: Harvest-floor assessments of targeted characteristics that affect quality and value of cattle, carcasses, and byproducts. J. Anim. Sci. 90:5135–5142. doi:10.2527/jas.2012-5477.

McKnite, A. M., M. E. Perez-Munoz, L. Lu, E. G. Williams, S. Brewer, P. A. Andreux, J. W. Bastiaansen, X. Wang, S. D. Kachman, J. Auwerx, R. W. Williams, A. K. Benson, D. A. Peterson, and D. C. Ciobanu. 2012. Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits. PLoS ONE 7(6):e39191. doi:10.1371/journal.pone.0039191.

Minamino, T., Y. Ito, H. Ohkubo, Y. Shimuzu, K. Kojo, N. Nishizwa, H. Amano, S. Narumiya, W. Koizumi, and M. Majima. 2015. Adhesion of platelets through thromboxane A_2 receptor signaling facilitates liver repair during acute chemical-induced hepatotoxicity. Life Sci. 132:85–92. doi:10.1016/j.lfs.2015.03.015.

Myer, P. R., T. P. Smith, J. E. Wells, L. A. Kuehn, and H. C. Freetly. 2015. Rumen microbiome from steers differing in feed efficiency. PLoS ONE 10(6):e0129174. doi:10.1371/journal.pone.0129174.

Owens, F. N., D. S. Secrist, W. J. Hill, and D. R. Gill. 1998. Acidosis in cattle: A review. J. Anim. Sci. 76:275–286.

Peiris, B. L., J. Ralph, S. J. Lamont, and J. C. Dekkers. 2011. Predicting allele frequencies in DNA pools using high density SNP genotyping data. Anim. Genet. 42:113–116. doi:10.1111/j.1365-2052.2010.02077.x.

R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reinhardt, C. D., and M. E. Hubbert. 2015. Review: Control of liver abscesses in feedlot cattle: A review. Prof. Anim. Sci. 31:101–108. doi:10.15232/pas.2014-01364.

Rezac, D. J., D. U. Thomson, S. J. Bartle, J. B. Osterstock, F. L. Prouty, and C. D. Reinhardt. 2014a. Prevalence, severity, and relationships of lung lesions, liver abnormalities, and rumen health scores measured at slaughter in beef cattle. J. Anim. Sci. 92:2595–2602. doi:10.2527/jas.2013-7222.

Rezac, D. J., D. U. Thomson, M. G. Siemens, F. L. Prouty, C. D. Reinhardt, and S. J. Bartle. 2014b. A survey of gross pathologic conditions in cull cows at slaughter in the Great Lakes region of the United States. J. Dairy Sci. 97:4227–4235. doi:10.3168/jds.2013-7636.

Segura, V., B. J. Vilhjálmsson, A. Platt, A. Korte, Ü. Seren, Q. Long, and M. Nordborg. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat. Genet. 44(7):825–830. doi:10.1038/ng.2314.

Smedley, D., S. Haider, S. Durinck, L. Pandini, P. Provero, J. Allen, O. Arnaiz, M. H. Awedh, R. Baldock, G. Barbiera, P. Bardou, T. Beck, A. Blake, M. Bonierbale, A. J. Brookes, G. Bucci, I. Buetti, S. Burge, C. Cabau, J. W. Carlson, C. Chelala, C. Chrysostomou, D. Cittaro, O. Collin, R. Cordova, R. J. Cutts, E. Dassi, A. D. Genova, A. Djari, A. Esposito, H. Estrella, E. Eyras, J. Fernandez-Banet, S. Forbes, R. C. Free, T. Fujisawa, E. Gadaleta, J. M. Garcia-Manteiga, D. Goodstein, K. Gray, J. A. Guerra-Assunção, B. Haggarty, D. J. Han, B. W. Han, T. Harris, J. Harshbarger, R. K. Hastings, R. D. Hayes, C. Hoede, S. Hu, Z. L. Hu, L. Hutchins, Z. Kan, H. Kawaji, A. Keliet, A. Kerhornou, S. Kim, R. Kinsella, C. Klopp, L. Kong, D. Lawson, D. Lazarevic, J. H. Lee, T. Letellier, C. Y. Li, P. Lio, C. J. Liu, J. Luo, A. Maass, J. Mariette, T. Maurel, S. Merella, A. M. Mohamed, F. Moreews, I. Nabihoudine, N. Ndegwa, C. Noirot, C. Perez-Llamas, M. Primig, A. Quattrone, H. Quesneville, D. Rambaldi, J. Reecy, M. Riba, S. Rosanoff, A. A. Saddiq, E. Salas, O. Sallou, R. Shepherd, R. Simon, L. Sperling, W. Spooner, D. M. Staines, D. Steinbach, K. Stone, E. Stupka, J. W. Teague, A. Z. Dayem Ullah, J. Wang, D. Ware, M. Wong-Erasmus, K. Youens Clark, A. Zadissa, S. J. Zhang, and A. Kasprzyk. 2015. The BioMart community portal: An innovative alternative to large, centralized data repositories. Nucleic Acids Res. 43(W1):W589–W598. doi:10.1093/nar/gkv350.

Steemers, F. J., W. Chang, G. Lee, D. L. Barker, R. Shen, and K. L. Gunderson. 2006. Whole-genome genotyping with the single-base extension assay. Nat. Methods 3:31–33. doi:10.1038/nmeth842.

Tadepalli, S., S. K. Narayanan, G. C. Stewart, M. M. Chengappa, and T. G. Nagaraja. 2009. *Fusobacterium necrophorum*: A ruminal bacterium that invades liver to cause abscesses in cattle. Anaerobe 15:36–43. doi:10.1016/j.anaerobe.2008.05.005.

Varel, V. H., and B. A. Dehority. 1989. Ruminal cellulolytic bacteria and protozoa from bison, cattle-bison hybrids, and cattle fed three alfalfa-corn diets. Appl. Environ. Microbiol. 55:148–153.

Wu, Y., H. Fan, S. Jing, J. Xia, Y. Chen, L. Zhang, X. Gao, J. Li, H. Gao, and H. Ren. 2015. A genome-wide scan for copy number variations using high-density single nucleotide polymorphism array in Simmental cattle. Anim. Genet. 46(3):289–298. doi:10.1111/age.12288.

Zhao, J. H. 2007. gap: Genetic Analysis Package. J. Stat. Softw. 23(8):1–18. doi:10.18637/jss.v023.i08.

Zhao, J. H. (2015). gap: Genetic Analysis Package. R package version 1.1-16. http://cran.r-project.org/package=gap. (Accessed Oct. 9, 2015.)

Supplementary Table 1. KEGG[1] pathway assignments for genes.[2]

Chr.	Starting Position	Ending Position	Gene Symbol	Gene Name	KEGG Pathways
2	107,462,687	107,504,968	CYP27A1	cytochrome P450, family 27, subfamily A, polypeptide 1	Primary bile acid biosynthesis PPAR signaling pathway
2	107,509,452	107,516,981	PRKAG3	protein kinase, AMP-activated, gamma 3 non-catalytic subunit	Insulin signaling pathway Adipocytokine signaling pathway Hypertrophic cardiomyopathy (HCM)
2	107,544,683	107,556,681	WNT6	wingless-type MMTV integration site family, member 6	Wnt signaling pathway Hedgehog signaling pathway Melanogenesis Pathways in cancer Basal cell carcinoma
6	57,070,203	57,251,086	ARAP2	similar to centaurin delta 1	Endocytosis
7	21,514,211	21,535,516	PIP5K1C	phosphatidylinositol-4-phosphate 5-kinase, type I, gamma	Inositol phosphate metabolism Phosphatidylinositol signaling system Endocytosis Fc gamma R-mediated phagocytosis Regulation of actin cytoskeleton
7	21,561,179	21,569,885	TBXA2R	thromboxane A2 receptor	Calcium signaling pathway Neuroactive ligand-receptor interaction
10	89,756,991	89,852,261	SPTLC2	serine palmitoyltransferase, long chain base subunit 2	Ceramide Biosynthesis Metabolic pathways Sphingolipid metabolism Sphingolipid signaling pathway Sphingosine biosynthesis
10	89,923,193	89,958,906	SNW1	SNW domain containing 1	Spliceosome Epstein-Barr virus infection Notch signaling pathway Viral carcinogenesis
15	47,109,758	47,267,567	APBB1	amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)	Alzheimer's disease
15	47,267,703	47,272,165	SMPD1	sphingomyelin phosphodiesterase 1, acid lysosomal	Sphingolipid metabolism Lysosome
19	16,353,233	17,562,209	ACCN1	amiloride-sensitive cation channel 1, neuronal	Taste transduction
21	68,595,873	68,601,204	HSP90AA1	similar to 90-kDa heat shock protein alpha; heat shock 90kD protein 1, alpha	Antigen processing and presentation NOD-like receptor signaling pathway Progesterone-mediated oocyte maturation Pathways in cancer Prostate cancer
22	50,655,648	50,661,989	SEMA3B	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B	Axon guidance
22	50,670,852	50,691,007	GNAI2	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 2	Chemokine signaling pathway Axon guidance Tight junction Gap junction Leukocyte transendothelial migration Long-term depression Progesterone-mediated oocyte maturation Melanogenesis

22	50,733,157	50,756,457	SEMA3F	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3F	Axon guidance

[1]KEGG is a database of metabolic, signaling and regulatory pathways (Kanehisa and Goto, 2000; Kanehisa et al., 2014). [2]Annotation was performed using the DAVID gene annotation website (Huang et al., 2009).