Identification of single nucleotide polymorphisms in genes involved in digestive and metabolic processes associated with feed efficiency and performance traits in beef cattle

M. K. Abo-Ismail, M. J. Kelly, E. J. Squires, K. C. Swanson, S. Bauck and S. P. Miller

J ANIM SCI 2013, 91:2512-2529.
doi: 10.2527/jas.2012-5756 originally published online March 18, 2013

Identification of single nucleotide polymorphisms in genes involved in digestive and metabolic processes associated with feed efficiency and performance traits in beef cattle[1,2]

M. K. Abo-Ismail,*† M. J. Kelly,*‡ E. J. Squires,* K. C. Swanson,*§ S. Bauck,# and S. P. Miller*[3]

*Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, Ontario, Canada, N1G 2W0; †Department of Animal and Poultry Science, Damanhour University, Damanhour, Egypt; ‡Queensland Alliance for Agriculture and Food Innovation University of Queensland, St Lucia, QLD 4072, Australia; §Animal Sciences Department, North Dakota State University, Fargo 58108-6050; and #GeneSeek, 4665 Innovation Drive, Suite 120, Lincoln, NE 68521

ABSTRACT: Discovery of genetic mutations that have a significant association with economically important traits would benefit beef cattle breeders. Objectives were to identify with an in silico approach new SNP in 8 genes involved in digestive function and metabolic processes and to examine the associations between the identified SNP and feed efficiency and performance traits. The association between SNP and daily DMI, ADG, midpoint metabolic weight (MMWT), residual feed intake (RFI), and feed conversion ratio (FCR; the ratio of average daily DMI to ADG) was tested in discovery and validation populations using a univariate mixed-inheritance animal model fitted in ASReml. Substitution effect of the T allele of SNP rs41256901 in protease, serine, 2 (trypsin 2; *PRSS2*) was associated with FCR (–0.293 ± 0.08 kg DMI kg^{-1} BW gain; $P < 0.001$) and RFI (–0.199 ± 0.08 kg; $P < 0.01$) and although not significant in the validation population, the phase of association remained. In the cholecystokinin B receptor (*CCKBR*) gene, genotypes in rs42670351 were associated with RFI ($P < 0.05$) whereas genotypes in rs42670352 were associated with RFI ($P = 0.002$) and DMI ($P < 0.05$). Substitution of the G allele in rs42670352 was associated with DMI (–0.236 ± 0.12 kg; $P = 0.055$) and RFI (–0.175 ± 0.09 kg; $P = 0.05$). Substitution of the G allele of SNP rs42670353 was associated with ADG (0.043 ± 0.02 kg/d; $P < 0.01$) and FCR (0.114 ± 0.05 kg BW gain kg^{-1} DMI; $P < 0.05$). In the validation dataset, SNP rs42670352 in gene *CCKBR* was significant for RFI and DMI and had the same phase of associations; SNP rs42670353 was significantly associated with FCR with same phase of association and the C allele in SNP rs42670351 was validated as decreasing DMI, RFI, and FCR. Substituting the G allele of SNP rs42670352 in CCKBR2 was associated with decreasing DMI and RFI in the validation study. New SNP were reported in genes *PRSS2* and *CCKBR*, being associated with feed efficiency and performance traits in beef cattle. The association between these SNP with fertility, carcass, and meat quality traits must still be tested.

Key words: beef cattle, candidate genes, feed efficiency, single nucleotide polymorphism

J. Anim. Sci. 2013.91:2512–2529
doi:10.2527/jas2012-5756

INTRODUCTION

A small improvement in feed efficiency would have a significant influence on the profitability of the beef production system (Herd et al., 2003). Residual feed intake (**RFI**) is one of the acceptable traits for improving feed efficiency in feedlot cattle (Wulfhorst et al., 2010). Estimates of the genetic variation in feed efficiency (Archer et al., 1999; Arthur et al., 2001; Schenkel et al., 2004) indicate that RFI is moderately heritable providing an opportunity for selection although the difficulty of recording feed intake has been reported as a major

[1]The authors thank the staff at the Elora Beef Research station, Gordon Vandervoort, and Tim Caldwell for their assistance with data collection and analysis.

[2]Financial support was provided by Ontario Cattlemen's Association, Agriculture Adaptation Council (FIP program), Beef Cattle Research Council, and Ontario Ministry of Agriculture Food and Rural Affairs.

[3]Corresponding author: miller@uoguelph.ca

Received August 15, 2012.
Accepted March 7, 2013.

limitation (Arthur et al., 2001) toward implementation of selection for improved RFI. Accordingly, other criteria to evaluate feed efficiency such as DNA markers have been considered (Barendse et al., 2007; Nkrumah et al., 2007; Sherman et al., 2008a,b, 2010).

Herd et al. (2004) and Richardson and Herd (2004) proposed that processes such as digestion, body composition development, metabolism, including biological processes such as ion pumping, proton leakage, and protein turnover, activity, and thermoregulation contribute to the variation in RFI. There is evidence suggesting that inadequate production of specific digestive enzymes could be responsible for limitation in digestive efficiency. Genes such as pancreatic α amylase (***AMY2B***) is known as a primary enzyme responsible for starch digestion in cattle fed high-concentration diets (Swanson et al., 2002). In pigs, the concentration of the pancreatic trypsin enzyme was positively associated with ADG and negatively associated with feed conversion ratio (**FCR**: the ratio of average daily feed intake to ADG; Van den Borne et al., 2007). Therefore, cationic trypsin (*LOC780933*), pancreatic anionic trypsinogen or protease, serine, 2 (**trypsin 2**; *PRSS2*; Le Huerou et al., 1990), and pancreatic trypsin inhibitor (***PTI***; Ascenzi et al., 2003) are potential candidate genes. In human, polymorphisms in *CT* activation peptides were associated with pancreatitis (Chen and Ferec, 2000; Teich and Mossner, 2008; Kereszturi et al., 2009). Cholecystokinin (**CCK**) B receptor (***CCKBR***) regulates effect of CCK (Le Meuth et al., 1993; Le Dréan et al., 1999; Rehfeld et al., 2007). Uncoupling protein 2 (*UCP2*) provided a link between mitochondrial respiration and feed efficiency in beef cattle (Kolath et al., 2006; Sherman et al., 2008b) and obesity and insulin secretion (Zhang et al., 2001). Pyruvate carboxylase (***PC***) is one of the key enzymes playing a potential role in gluconeogenesis (in liver and kidney), lipogenesis (in adipose tissue and lactating mammary gland), and insulin signaling pathway (in pancreatic islets; Jitrapakdee and Wallace 1999; Greenfield et al., 2000; Velez and Donkin, 2005; Haga et al., 2008). The adenosine triphosphatase (ATPase), H+ transporting, lysosomal 56/58 kDa, V1 subunit B2 (***ATP6V1B2***) is involved in transmembrane transport, hydrolase activity, proteolysis, and generation of precursor metabolites and energy (Jefferies et al., 2008; Appendix 1). In addition, other isoform of Vacuolar-type H+-ATPase is involved in the regulation of insulin secretion from pancreatic β-cells (Sun-Wada et al., 2006). Genes involved in these processes (Appendix 1) are good candidates for improving feed efficiency.

Identifying new SNP in these candidate genes significantly associated with RFI would be beneficial. Therefore, objectives were to discover new SNP in 8 genes involved in digestive and metabolic processes and examine the relationships between these SNP and feed efficiency and performance traits and validate these associations in a more recent group of cattle.

MATERIALS AND METHODS

Discovery Study

Phenotypic Data Collection. The study was approved from The University of Guelph Animal Care Committee based on the recommendations outlined in the Canadian Council on Animal Care (1993) guidelines. Animals were born in 1 of 3 University of Guelph cooperative herds, the University of Guelph Elora Beef Research Centre (**EBRC**), University of Guelph New Liskeard Agriculture Research Station, and the Agriculture Agri-Food Canada Kapuskasing Experimental Farm, or purchased from commercial sources. Calves were weaned at approximately 200 d of age and were involved in various postweaning trials at the EBRC with different nutritional treatments over time from 1998 to 2007. Phenotypes were collected from an average of 660 crossbred animals, heifers (40), steers (363), and bulls (257). Average breed contributions were Angus (**AN**; 41%), Simmental (**SM**; 24%), Piedmontese (**PI**; 11%), Charolais (**CH**; 8%), Gelbvieh (**GV**; 4%), and Limousin (**LIM**; 1%) determined by pedigree information on the ancestors. Body weight was recorded a number of times over the course of trials, with most trials recording BW at least every 4 wk. The ADG of the animals was calculated as a linear regression coefficient of their BW on the actual 112 d of measurement using nlme package from R software (Pinheiro et al., 2011). The R^2 for all growth curves averaged from 0.85 to 0.99. Midpoint metabolic weight (**MMWT**) was calculated as the midpoint BW to the power 0.75. Daily DMI data were acquired by 2 automated feeding systems: Calan-gate (American Calan, Northwood, NH; Ferris et al., 2007) and Insentec (Insentec, Marknesse, the Netherlands; Chapinal et al., 2007) systems where DMI data was filtered to exclude outlier records or days due to mechanical problems. The DMI was calculated for each animal as total DMI divided by number of days during the test period. The RFI was calculated from the difference between the average of the actual daily DMI and the expected DMI of the animal (Koch et al., 1963; Arthur et al., 2001). Expected DMI was determined through the regression coefficients estimated from the data using a multiple phenotypic regression model as follows:

$$y_{ijk}=\mu+\beta_1\text{ADG}+\beta_2\text{MMWT}+\text{Sex}_i+\text{TTY}_j+e_{ijk}, \quad [1]$$

in which y_{ijk} is the DMI for animal k during the feeding period, μ is the overall mean, β_1 is the regression on ADG as determined through a linear regression of BW on days of trial as described above, β_2 is the coefficient of the linear regression on MMWT, Sex_i is the effect of the *i*th sex, **TTY**

Table 1. Descriptive statistics in feedlot beef cattle for performance and feed efficiency in the training and validation datasets

	Mean		SD		Minimum		Maximum	
Trait[1]	Training	Validation	Training	Validation	Training	Validation	Training	Validation
ADG, kg d^{-1}	1.81	1.70	0.38	0.39	0.60	0.71	3.29	3.30
MMWT, kg	103.30	92.39	14.94	11.70	68.32	53.25	157.70	128.10
DMI, kg d^{-1}	9.89	9.81	1.60	1.76	5.38	4.18	15.64	15.54
RFI, kg d^{-1}	–0.12	–0.07	0.89	1.13	–5.62	–3.70	3.84	3.35
FCR, kg gain kg^{-1} DM	5.68	6.09	1.52	1.87	2.68	3.11	18.74	16.76

[1]MMWT = midpoint metabolic weight; RFI = residual feed intake; FCR = feed conversion ratio.

is the effect of *j*th treatment × trial × year (38 levels), and e_{ijk} is the residual random effect associated with animal *k* and is the resulting RFI used in further analyses. The descriptive statistics of the traits are given in Table 1.

Single Nucleotide Polymorphism Discovery. An in silico study was conducted to discover SNP within genes *AMY2B*, *LOC780933*, *PRSS2*, *PTI*, and *CCKBR*, *UCP2*, *PC*, and *ATP6V1B2* using the available expressed sequence tags (**EST**) or whole genome shotgun (**WGS**) traces in GenBank (Benson et al., 2005). The SNP were discovered in the candidate genes in silico in 3 main steps. The first step was to obtain the EST or WGS required for the alignment as follows: a) The reference sequence (cDNA) in FASTA format was acquired from the GenBank in the National Center for Biotechnology Information (NCBI), b) The reference sequence was aligned with the cow sequence using the basic local alignment search tool (**BLAST**; Zhang et al., 2000) at http://blast.ncbi.nlm.nih.gov, c) the Traces-EST (ftp://ftp.ncbi.nih.gov/repository/dbEST/) or WGS databases (Benson et al., 2012) and the MegaBlast program (Morgulis et al., 2008) using the default parameters were used, and d) traces were selected and acquired from the Trace archive including standard chromatography files (**SCF**).

The second step was the SNP identification process. In this step, a DNA sequence assembly software called Sequencher (Gene Codes Corporation, Ann Arbor, MI) was used to align the acquired sequences with the reference sequence. Using this software, SNP were detected based on the nucleotide sequences and attached standard chromatography files. The SNP that lead to a change in the sequence of AA were also detected.

The third step was to determine the position of SNP within the gene sequence. Briefly, the whole gene sequence was acquired from NCBI in FASTA format. Then it was aligned with the reference (cDNA) and the EST using Sequencher software. The flanking sequences of the SNP were obtained for genotyping purposes. In total, 39 SNP from 8 genes were selected from the in silico results. Eighteen of the SNP have not been previously reported in the public domain (Table 2) whereas 21 SNP have been reported in the SNP database in NCBI (Table 3).

Animal Genotyping. Genomic DNA was extracted from tissue or blood samples (Sambrook et al., 1989; Rudi et al., 1997; Caldarelli-Stefano et al., 1999). Prepared DNA samples were sent to Merial Ltd. (Lincoln, NE) for genotyping using a commercial platform for high-throughput SNP genotyping and an allele-specific primer extension on a microarray (Pastinen et al., 2000; Makridakis and Reichardt, 2001). In total, 993 animals were genotyped for 39 SNP for the discovery population.

Statistical Analysis. Allele frequencies were calculated for each SNP on all genotyped animals. The Hardy-Weinberg equilibrium (**HWE**) was tested using the likelihood ratio test (G-test), described by Lynch and Walsh (1998):

$$G = -2\sum_{i=1}^{n}\sum_{j>1}^{n} N_{ij} \ln\left(\tilde{N}_{ij} / N_{ij}\right), \quad [2]$$

in which N_{ij} and $\tilde{N}_{ij}$ are the observed and the expected number of genotype g_{ij}. The extent of linkage disequilibrium (**LD**) between pairs of SNP was calculated using the haploxt program from the Graphical Overview of Linkage Disequilibrium (**GOLD**) package (Abecasis and Cookson,

Table 2. Gene name, chromosome number (*Bos taurus* autosome (BTA), GenBank Entrez Gene Identifier, SNP name, SNP position, nucleotide change, and functional consequences for SNP discovered *in silico*

Gene	BTA: gene ID	SNP name	SNP position[1]	AA change[2]	SNP
AMY2B	3:539383	AMY2B1	39936868	Ser/Asn	G/A
		AMY2B2	39935769	Arg/Arg	C/T
		AMY2B4	39891301	Thr/Ala	A/G
LOC780933	4:780933	CT1	106362449	Ala/Ala	C/T
		CT2	106361029	Pro/Pro	C/T
		CT3	106360880	Ser/Cys	C/G
		CT4	106359896	Ser/Phe	C/T
		CT5	106359889	Ser/Ser	T/C
		CT6	106359796	Ala/Ala	C/T
PRSS2	4: 282603	TRYP81	106888934	Ala/Ala	A/G
PTI	13:404172	PTI2	74947682	Val/Ile	C/T
		PTI3	74947651	Ala/Val	G/A
		PTI4	74944796	Pro/Ser	C/T
		PTI6	74944705	Arg/Lys	G/A
UCP2	15:281562	UCP22	54197781	Ala/Ala	T/C
		UCP23	54197685	Ala/Ala	A/G
		UCP24	54197451	Tyr/Tyr	C/T

[1]The SNP position is based on *Bos taurus* UMD 3.1, Genome Build 37.3 (Zimin et al., 2009).

[2]The effect of mutation (SNP) on the AA sequence.

Table 3. Gene name, GenBank Entrez Gene Identifier, chromosome number, SNP name, SNP position, nucleotide change, and functional consequences for SNP reported in the National Center for Biotechnology Information

Gene	BTA and gene bank ID[1]	SNP name	Accession number	SNP position[2]	AA change[3]	SNP
AMY2B	3:539383	AMY2B6	rs42312301	39931232	Asp/Asn	G/A
PRSS2	4: 282603	TRYP82	rs41256900	106888943	Ser/Ser	C/T
		TRYP83	rs41256901	106890553	Ser/Phe	T/C
ATP6V1B2	8:338082	ATPase1	rs43563470	67810773	Asp/Asp	C/T
		ATPase2	rs43562811	67823611	3′ UTR	C/T
		ATPase3	rs43562810	67823802	3′ UTR	C/T
		ATPase4	rs43562809	67824091	3′ UTR	A/G
PTI	13:404172	PTI1	rs43024409	74947703	Met/Leu	A/T
		PTI5	rs41257167	74944767	Ile/Met	T/G
		PTI8	rs43024345[4]	74943702	3′ near gene	T/C
CCKBR	15:281665	CCKBR1	rs42670351	47386394	Arg/Arg	A/C
		CCKBR2	rs42670352	47385604	Ala/Ala	G/T
		CCKBR3	rs42670353	47385334	Phe/Phe	C/T
UCP2	15:281562	UCP21	rs41255549	54199080	Ala/Ala	G/T
		UCP25	rs41774217	54196971	Cys/Cys	A/G
PC	29:338471	PC1	rs42194938	45602034	Intronic	A/G
		PC2	rs42194937	45601239	Intronic	G/T
		PC3	rs42195008	45529368	Ile/Ile	A/G
		PC4	rs42197374	45510553	Val/Ile	A/G
		PC5	rs42197375	45510113	Tyr/Tyr	C/T
		PC6	rs42197376	45508443	3′ UTR	A/G

[1]BTA = *Bos taurus* autosome.

[2]The SNP position is from *Bos taurus* UMD_3.1, Genome Build 37.3 (Zimin et al., 2009).

[3]The effect of mutation (SNP) on AA sequence. UTR = the untranslated region.

[4]This SNP was merged to rs41257168 SNP.

2000). The file of marker haplotypes was prepared using fastPHASE 1.1 (Scheet and Stephens, 2006). Haploview software (Barrett et al., 2005) was used to graphically view the extent of LD, assign the haplotype blocks (i.e., SNP with high LD, $D' > 0.77$) based on the 4-gamete rule (Wang et al., 2002), and identify the haplotype-tagging SNP using the TAGGER algorithm (De Bakker et al., 2005).

Genotype Analysis. Associations of the genotypes for each SNP at a time with the traits were evaluated by genetic analysis using ASReml (Gilmour et al., 2009). An animal model was fitted as follows:

$$y_{ijklm} = \mu + Gj + Sex_k + TTY_l + \beta_1 AET + \beta_2 AN = \beta_3 CH + \beta_4 LIM + \beta_5 SM + \beta_6 PI + \beta_7 GV + \beta_8 HET + a_m + e_{ijklm}, \quad [3]$$

in which y_{jklm} is the trait measured in the mth animal of the kth sex and the lth treatment trial-year group, μ is the overall mean for the trait, G_j is the fixed effect of the jth genotype for the SNP considered, Sex_k is the fixed effect of the kth sex of mth animal, TTY_l is the fixed effect of the lth treatment trial-year group, β_1 is the regression coefficient of the linear regression on age at the end of test period (**AET**) of the mth animal, β_2 to β_7 are the regression coefficients of the linear regressions on the proportion of AN, CH, LIM, SM, PI, and GV breeds in the mth animal, β_8 is the regression coefficient of the linear regression on the percent of heterozygosity (**HET**) of mth animal, a_m is the random additive genetic (polygenic) effect of the mth animal, and e_{jklm} is the residual random effect associated with the mth animal record. Assumptions for this model are a_m: $a \sim N(0, \mathbf{A}\sigma^2_a)$ in which **A** is the relationship matrix and σ^2_a is the additive genetic variance and e_{jklm}: $e \sim (0, \mathbf{I}\sigma^2_e)$ in which **I** is the identity matrix and σ^2_e is the error variance. The expectations are $E(a_m) = 0$ and $E(e_{jklm}) = 0$ and the variances are $Var(a_m) = \sigma^2_a$ and $Var(e_{jklm}) = \sigma^2_e$. The $\mathbf{A}\sigma^2_a$ is the covariance matrix of the vector of animal additive genetic effects and the relationship matrix (**A**) is assumed to be complete back to the base population.

For the significance level used to assess the results, an overall value of $P < 0.05$ (α) was used. A modified Bonferroni correction was used [$\alpha/(N)^{1/2}$; Mantel, 1980] to adjust for multihypotheses testing for controlling type I errors where N is the number of SNP multiplied by the number of traits. Therefore, the modified Bonferroni-corrected significance level in the discovery population is 0.0043 [$0.005/(27 \times 5)^{1/2}$] at $\alpha = 0.05$.

Allele Substitution Effect Model. This model included the same effects as the genotypic model except that the genotypic effect was replaced with an allele substitution effect, which is estimated by

Table 4. Genotypic and minor allele frequencies and the Hardy-Weinberg equilibrium for SNP in the discovery population

Gene	SNP	Genotype frequency			MAF[1]	G[2]
AMY2B	rs42312301	GG (0.000)	AG (0.999)	AA (0.001)	0.50	1,306.80[3]
ATP6V1B2	rs43563470	CC (0.853)	CT (0.139)	TT (0.007)	0.08	0.33
	rs43562811	TT (0.592)	CT (0.356)	CC (0.052)	0.23	0.02
	rs43562810	TT (0.885)	CT (0.113)	CC (0.002)	0.06	0.61
	rs43562809	AA (0.008)	AG (0.144)	GG (0.847)	0.08	0.54
CCKBR	rs42670351	AA (0.598)	AC (0.350)	CC (0.051)	0.23	0
	rs42670352	TT (0.957)	GT (0.004)	GG (0.039)	0.04	280.85[3]
	rs42670353	TT (0.297)	CT (0.481)	CC (0.221)	0.46	0.93
LOC780933	CT2	CC (0.002)	CT (0.999)	TT (0.002)	0.50	1,301.40[3]
	CT5	CC (0.133)	CT (0.867)	TT (0.001)	0.43	714.55[3]
PC	rs42194938	AA (0.522)	AG (0.405)	GG (0.073)	0.28	0.24
	rs42194937	GG (0.944)	GT (0.055)	TT (0.001)	0.03	0.06
	rs42195008	GG (0.942)	AG (0.057)	AA (0.001)	0.03	0.04
	rs42197374	GG (0.571)	GA (0.000)	AA (0.429)	0.43	9.56[3]
	rs42197375	TT (0.733)	CT (0.248)	CC (0.019)	0.14	0.14
	rs42197376	GG (0.943)	AG (0.056)	AA (0.001)	0.03	0.04
PTI	rs43024409	AA (0.000)	AT (0.531)	TT (0.469)	0.27	188.04[3]
	PTI2	CC (0.187)	CT (0.812)	TT (0.001)	0.41	571.12[3]
	PTI3	GG (0.747)	AG (0.253)	AA (0.000)	0.13	34.86[3]
	rs41257167	GG (0.007)	GT (0.993)	TT (0.000)	0.50	1,246.63[3]
PRSS2	TRYP81	GG (0.451)	AG (0.462)	AA (0.087)	0.32	4.03[3]
	rs41256901	CC (0.679)	CT (0.321)	TT (0.000)	0.16	58.41[3]
UCP2	rs41255549	TT (0.497)	GT (0.413)	GG (0.089)	0.30	0.07
	UCP22	TT (0.603)	CT (0.352)	CC (0.045)	0.22	0.48
	UCP23	AA (0.600)	AG (0.356)	GG (0.044)	0.22	1.01
	UCP24	CC (0.902)	CT (0.097)	TT (0.001)	0.05	1.06
	rs41774217	GG (0.971)	AG (0.029)	AA (0.000)	0.01	0.41

[1]MAF = minor allele frequency.

[2]G = the G-test statistic.

[3]Not in agreement with the Hardy-Weinberg equilibrium

regressing the phenotype on the number of copies of a given allele (0, 1, or 2) using ASReml.

Validation Study

Tissue or blood samples from 1,032 animals born subsequent to the animals used in the discovery population were prepared and sent to Molecular Supercentre Laboratory Services, University of Guelph, Guelph, Canada, for genomic DNA extraction. Then prepared DNA samples were sent to GeneSeek, Inc., (Lincoln, NE) for genotyping using Illumina Infinium BeadChip with single-base extension assay (Steemers et al., 2006). A total of 1,032 animals were genotyped for validation using 17 out of 39 SNP from the discovery population. The reason for reducing the number of SNP (17) is related to cost and efficiency as some of these SNP were highly linked to each other.

A quality control (**QC**) procedure was conducted using the GenABEL package (Aulchenko et al., 2007) in R software. The SNP and animals with a low call rate (<90%) were excluded from the analysis. Animals with an estimated high frequency of SNP identical by state ≥ 0.95 were excluded. The SNP with a minor allele frequency <1% (e.g., UCP25, AMY2B6, and CT2 SNP) were excluded from the analysis. Animals with high autosomal heterozygosity 0.446 [false discovery rate < 0.05] were also excluded. Phenotypes not within the mean ± 3 SD were excluded. Contemporary groups (TTY levels) that had fewer than 3 animals were excluded. The QC procedure resulted in 14 SNP and 726 animals being used for further analyses.

The association analysis was performed with a univariate animal model fitting the allele substitution or genotypic effect using ASReml. The model included the same fixed systematic effects as previously stated as well as the fixed effect of herd by year of birth. The modified Bonferroni-corrected significance level in the validation population is 0.0058 $[0.05/(15 \times 5)^{1/2}]$ at $\alpha = 0.05$.

Table 5. Genotypic and minor allele frequencies (MAF) for SNP in the validation population

Gene ID[1]	SNP name	Accession number	Genotype frequency			MAF
539383	AMY2B6	rs42312301	AA (0.000)	AG (0.004)	GG (0.996)	0.002
338082	ATPase1	rs43563470	TT (0.004)	TC (0.133)	CC (0.863)	0.07
338082	ATPase2	rs43562811	TT (0.554)	TC (0.382)	CC (0.063)	0.255
338082	ATPase4	rs43562809	AA (0.006)	AG (0.137)	GG (0.858)	0.074
281665	CCKBR1	rs42670351	AA (0.624)	AC (0.319)	CC (0.056)	0.216
281665	CCKRB2	rs42670352	TT (0.629)	TG (0.317)	GG (0.055)	0.213
281665	CCKBR3	rs42670353	TT (0.329)	TC (0.496)	CC (0.176)	0.423
780933	CT2	in silico	TT (0.999)	TC (0.000)	CC (0.001)	0.001
338471	PC1	rs42194938	AA (0.062)	AG (0.347)	GG (0.591)	0.235
338471	PC3	rs42195008	AA (0.002)	AG (0.04)	GG (0.958)	0.022
338471	PC4	rs42197374	AA (0.447)	AG (0.435)	GG (0.119)	0.336
338471	PC5	rs42197375	TT (0.775)	TC (0.211)	CC (0.014)	0.119
338471	PC6	rs42197376	AA (0.003)	AG (0.038)	GG (0.959)	0.022
404172	PTI1	rs43024409	TT (0.131)	AT (0.457)	AA (0.412)	0.36
282603	TRYP81	in silico	AA (0.101)	AG (0.437)	GG (0.462)	0.32
282603	TRYP83	rs41256901	TT (0.000)	TC (0.316)	CC (0.684)	0.158
281562	UCP25	rs41774217	AA (0.001)	AG (0.013)	GG (0.986)	0.007

[1]Gene ID = Entrez Gene Identifier.

RESULTS

Discovery Population

In Silico Study. The genotyping success rate ranged from 95 to 98% except for SNP rs42197374 and rs41256900, which had success rates of 0.7 and 0%, respectively. Genotyping results showed that 8 of the 18 SNP discovered using Sequencher have both alleles in the genotyped discovery population (Table 4) whereas the remaining putative SNP were fixed (i.e., only 1 allele was present in the genotyped population). The SNP genotyped for the validation study are summarized in Table 5.

The population was tested for HWE using a G-test where the G value has a distribution that approximates to χ^2 with df equal to the number of genotypes minus the number of alleles. The total number of genotyped animals, the allelic frequencies, and the G value for each SNP are reported in Table 4. The SNP rs42312301, rs42670352, CT2, CT5, rs42197374, rs43024409, PTI2, rs41257167, and rs41256901 were not in HWE.

The values of LD (r^2) for each marker pair on a given chromosome are presented in Table 6. Values ranged from 0.0 to 0.99. The LD between SNP within *CCKBR* and *UCP2* was less than 0.10 whereas r^2 was 0.24 between gene *LOC780933* (SNP CT2) and gene *PRSS2* (SNP TRYP81). In addition, the extent of LD was high, ranging from 0.26 to 0.97, between the SNP pairs in *ATP6V1B2*. The extent of LD between SNP pairs is presented graphically in Fig. 1.

Association Analysis

Protease Serine 2. In gene *PRSS2*, substitution with the T allele of SNP rs41256901 was associated with a decrease of 0.184 kg in DMI (Table 7; $P = 0.084$), a decrease of 0.298 kg DMI/kg gain in FCR (Table 7; $P < 0.001$), and a decrease of 0.199 kg in RFI (Table 7; $P < 0.01$). The SNP rs41256901 was significantly associated with FCR where genotype CC had greater FCR (5.1%) than CT genotype ($P < 0.001$).

Cholecystokinin B Receptor. Genotypes in SNP rs42670351 were significantly associated with RFI (Table 8; $P < 0.05$). Substitution to the G allele of SNP rs42670352 tended to be associated with a 0.236 kg decrease in DMI (Table 8; $P = 0.055$) and a decrease of 0.175 kg in RFI (Table 8; $P = 0.053$). Genotypes in rs42670352 were significantly associated with DMI (Table 8; $P = 0.033$) and RFI (Table 8; $P = 0.002$). Substitution to the G allele of SNP rs42670353 was significantly associated with a 0.043 kg increase in ADG (Table 8; $P = 0.006$) and a 0.114 kg gain kg^{-1} DMI decrease in FCR (Table 8; $P = 0.033$).

Uncoupling Protein 2. Substitution to the T allele of SNP UCP24 was slightly associated with a 0.076 kg decrease in ADG (Table 9; $P = 0.054$). The SNP UCP24 did not show a significant relationship with feed efficiency.

Validation Study

The SNP within *PRSS2* were not associated with feed efficiency and performance traits. Four SNP in *CCKBR* were evaluated in the validation data set. The C allele of SNP CCKBR1 (rs42670351) decreased DMI by 0.235 kg (Table 9; $P = 0.00084$), decreased

Table 6. The extent of linkage disequilibrium (r^2) between pairs of SNP within the same chromosome (Chr) in the discovery population

BTA[1]	SNP_1[2]	SNP_2[2]	r^2	BTA	SNP_1	SNP_2	r^2
15	rs42670353	rs42670352	0.048	8	rs43563470	rs43562811	0.294
15	rs42670353	rs42670351	0.307	8	rs43563470	rs43562810	0.662
15	rs42670353	rs41774217	0.009	8	rs43563470	rs43562809	0.965
15	rs42670353	UCP24	0.006	8	rs43562811	rs43562810	0.215
15	rs42670353	UCP23	0.065	8	rs43562811	rs43562809	0.304
15	rs42670353	UCP22	0.064	8	rs43562810	rs43562809	0.706
15	rs42670353	rs41255549	0.014	13	rs41257167	PTI3	0.129
15	rs42670352	rs42670351	0.145	13	rs41257167	PTI2	0.116
15	rs42670352	rs41774217	0.004	13	rs41257167	rs43024409	0.157
15	rs42670352	UCP24	0.002	13	PTI3	PTI2	0.191
15	rs42670352	UCP23	0	13	PTI3	rs43024409	0.004
15	rs42670352	UCP22	0	13	PTI2	rs43024409	0.087
15	rs42670352	rs41255549	0.003	4	CT2	CT5	0.761
15	rs42670351	rs41774217	0.034	4	CT2	TRYP81	0.243
15	rs42670351	UCP24	0	4	CT2	rs41256901	0
15	rs42670351	UCP23	0.032	4	CT5	TRYP81	0.185
15	rs42670351	UCP22	0.032	4	CT5	rs41256901	0
15	rs42670351	rs41255549	0.094	4	TRYP81	rs41256901	0.055
15	rs41774217	UCP24	0.001	29	rs42197376	rs42197375	0.186
15	rs41774217	UCP23	0.004	29	rs42197376	rs42195008	0.982
15	rs41774217	UCP22	0.004	29	rs42197376	rs42194937	0.964
15	rs41774217	rs41255549	0.002	29	rs42197376	rs42194938	0.077
15	UCP24	UCP23	0.182	29	rs42197375	rs42195008	0.182
15	UCP24	UCP22	0.183	29	rs42197375	rs42194937	0.179
15	UCP24	rs41255549	0.021	29	rs42197375	rs42194938	0.003
15	UCP23	UCP22	0.991	29	rs42195008	rs42194937	0.982
15	UCP23	rs41255549	0.115	29	rs42195008	rs42194938	0.079
15	UCP22	rs41255549	0.117	29	rs42194937	rs42194938	0.078

[1]BTA = *Bos taurus* autosome.

[2]SNP_1 = single nucleotide polymorphism in locus 1; SNP_2 = single nucleotide polymorphism in locus 2.

RFI by 0.164 kg (Table 9; $P = 0.0315$), and decreased DMI by 0.117 (kg BW gain kg^{-1} DMI) in FCR (Table 9; $P = 0.059$). In addition, genotypes in rs42670351 were associated with DMI, RFI, and FCR (Table 10; $P = 0.009$, $P = 0.014$, and $P = 0.085$, respectively).

The G allele of rs42670352 was associated with a 0.222 kg decrease in DMI (Table 9; $P = 0.0116$), a 0.139 kg decrease in RFI (Table 9; $P = 0.066$), and a 0.099 kg BW gain kg^{-1} DMI decrease in FCR (Table 9; $P = 0.106$). Genotypes in rs42670352 were significantly associated with DMI and RFI (Table 10; $P < 0.05$).

The G allele of SNP CCKBR3 (rs42670353) was associated with a 0.251 kg decrease in DMI (Table 9; $P = 0.0008$), a 0.159 kg decrease in RFI (Table 9; $P = 0.0135$), and a 0.125 kg gain kg^{-1} DMI decrease in FCR (Table 9; $P = 0.0168$). Genotypes in rs42670353 were significantly associated with DMI, RFI, and FCR (Table 10; $P = 0.002$, $P = 0.005$, $P = 0.048$, respectively).

In gene *PTI*, substitution to the T allele of SNP PTI1 (rs43024409) was associated with a 0.973 kg^{75} increase in MMWT (Table 9; $P = 0.023$). The C allele of SNP ATPase2 (rs43562811) was associated with a 0.037 kg increase in ADG (Table 9; $P = 0.028$), a 0.941 kg^{75} increase in MMWT (Table 9; $P = 0.038$), and a 0.139 BW kg gain kg^{-1} DMI decrease in FCR (Table 9; $P = 0.022$). Genotypes in SNP ATPase2 (rs43562811) were significantly associated with MMWT (Table 10; $P = 0.037$).

DISCUSSION

Three SNP, AMY2B6 (rs42312301), TRYP83 (rs41256901), and UCP21 (rs41255549), were present in the GenBank. The GenBank had information for only SNP TRYP83 and UCP21. The remaining 22 SNP reported in the GenBank were identified using the *Bos taurus* Assembly SNP Discovery method where little information were available for these SNP. The SNP PTI8 (rs43024345) was reported in the GenBank and was not segregating in the current genotyped population. It is common that SNP in a database such as GenBank can be segregating in 1 particular population but not in another (Kitts and Sherry, 2007). The percentage of segregating SNP in the current population was 44.4% (8/18) based on SNP resulting

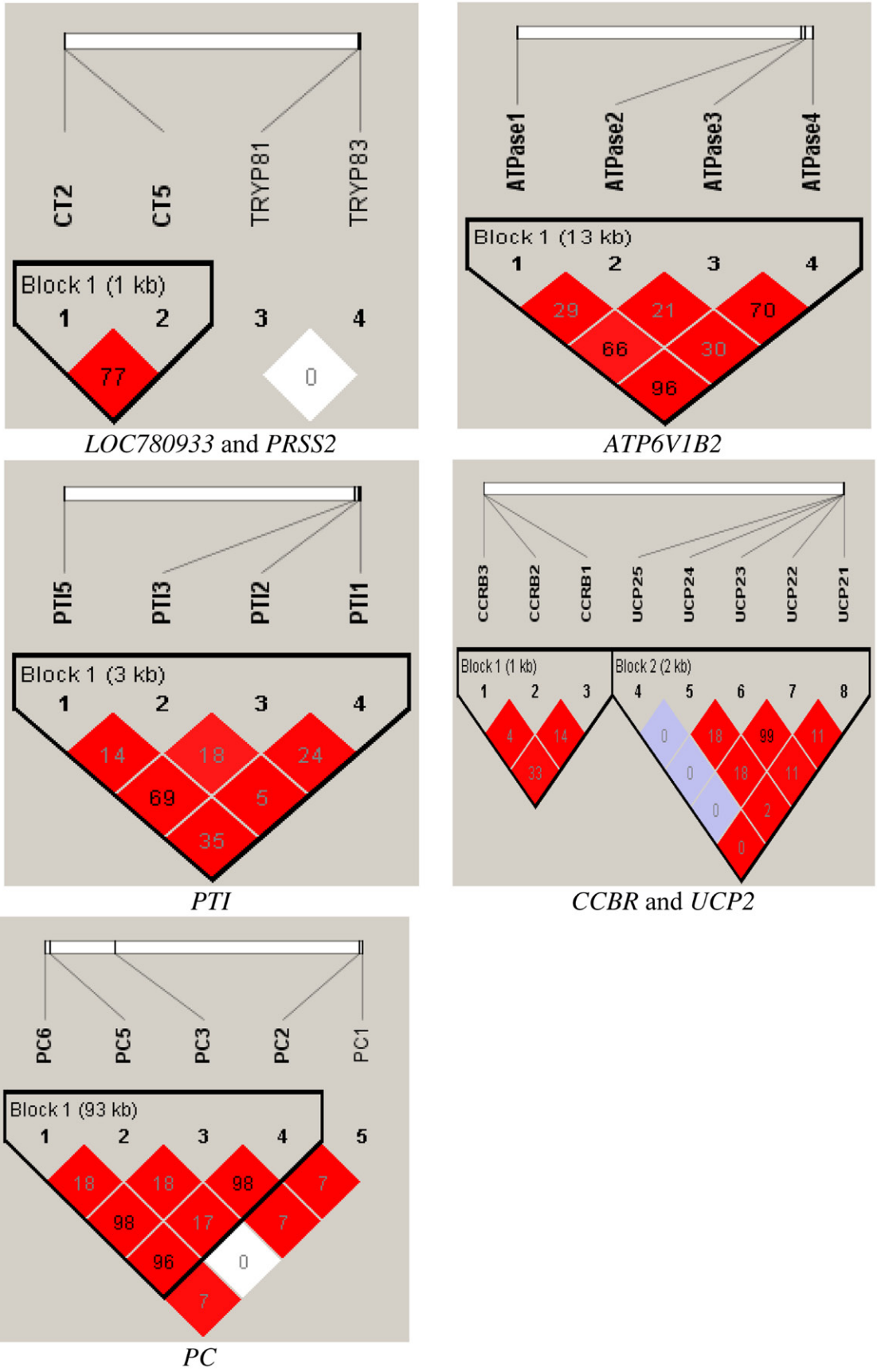

Figure 1. The extent of linkage disequilibrium (r^2) between pairs of SNP and haplotypes block structure in the candidate genes using Haploview. Cationic trypsin (*LOC780933*), protease serine 2 (trypsin 2; *PRSS2*), pancreatic trypsin inhibitor (*PTI*), cholecystokinin B receptor (*CCKBR*), uncoupling protein 2 (*UCP2*), pyruvate carboxylase (*PC*), and *ATP6V1B2* genes using Haploview. Linkage disequilibrium between each SNP pair is illustrated in a square where the number on the square represents the r^2 value between the 2 SNP corresponding to the cell. The empty square refers to $r^2 = 1$. Thick lines (black triangles) specify haplotype blocks where the size of the block is written in parentheses. See online version for figure in color.

Table 7. Estimates of allele substitution effects and genotypic effects (least squares means) of SNP in gene protease serine 2 (trypsin 2; *PRSS2*) in the discovery population

		Allele substitution effect				Genotype as fixed effect		
						LSM[3] ± SE	LSM ± SE	LSM ± SE
Trait[1]	SNP name	*P*-value	RA[2]	Estimate ± SE	*P*-value	C/C (G/G)	C/T (A/G)	T/T (A/A)
ADG	rs41256901	0.128	T	0.038 ± 0.02	0.128	1.659 ± 0.048	1.697 ± 0.048	
	TRYP81	0.595	A	0.01 ± 0.02	0.737	1.668 ± 0.048	1.668 ± 0.048	1.698 ± 0.048
DMI	rs41256901	0.084†	T	–0.184 ± 0.11	0.084†	9.596 ± 0.20	9.412 ± 0.20	
	TRYP81	0.189	A	0.101 ± 0.08	0.295	9.464 ± 0.20	9.621 ± 0.20	9.597 ± 0.20
FCR	rs41256901‡	<0.001**	T	–0.293 ± 0.08	<0.001**	6.201 ± 0.16	5.909 ± 0.16	
	TRYP81	0.332	A	0.059 ± 0.06	0.421	6.104 ± 0.149	6.21 ± 0.149	6.163 ± 0.149
MMWT	rs41256901	0.627	T	–0.331 ± 0.69	0.627	102.4 ± 1.331	102.069 ± 1.33	
	TRYP81	0.279	A	0.549 ± 0.5	0.058†	101.599 ± 1.25	103.071 ± 1.25	101.536 ± 1.25
RFI	rs41256901	0.010*	T	–0.199 ± 0.08	0.010*	–0.141 ± 0.146	–0.34 ± 0.146	
	TRYP81	0.478	A	0.039 ± 0.06	0.766	–0.216 ± 0.147	–0.186 ± 0.147	–0.126 ± 0.147

†Tended to affect the trait before the modified Bonferroni adjustment for multiple testing ($P < 0.10$).

‡Significant after the modified Bonferroni adjustment for multiple testing ($P < 0.05$).

*Significant effect before the modified Bonferroni adjustment for multiple testing ($P < 0.05$).

**Significant effect before the modified Bonferroni adjustment for multiple testing ($P < 0.01$).

[1]ADG = average daily gain (kg d^{-1}); DMI = daily dry matter intake (kg d^{-1}); FCR = feed conversion ratio (kg gain kg^{-1} DM); MMWT = midpoint metabolic weight (kg); RFI = residual feed intake (kg d^{-1}).

[2]RA = substitution allele.

[3]LSM = least squares mean.

from Sequencher with a minimum match percentage of 85% and minimum overlap of 20 bases. This proportion of segregated SNP was in agreement with Weckx et al. (2005) who estimated the false positives percentage (i.e., the percentage of SNP found fixed after genotyping) using different sequence-variation programs (PolyPhred (Nickerson et al., 1997), PolyBayes (Marth et al., 1999), and novoSNP (Weckx et al., 2005)). Weckx et al. (2005) reported that the percentages of false positives were 15.4, 51.5, and 86.2% for PolyPhred, PolyBayes, and novoSNP, respectively, at the greatest level of quality cutoff. However, they found the false positive rates at the lowest level of quality cutoff were greater at 94.9, 66.7, and 92.6% for PolyPhred, PolyBayes, and novoSNP, respectively. The results from using Sequencher, PolyPhred, PolyBayes, and novoSNP programs during SNP discovery indicated there was a high rate of false positives due to the direct relationship between the false positive rate and the quality of the sequence traces, particularly the background noise (Picoult-Newberg et al., 1999; Cox et al., 2001; Weckx et al., 2004). Nonetheless, the in silico approach provides cost-effective SNP detection in spite of the high rate of false positives, particularly with the advent of overwhelming results (millions of EST or reads) obtained from next generation sequencing stored in the public domain at Sequence Read Archive (at http://www.ncbi.nlm.nih.gov/Traces/sra/ from NCBI, http://www.ebi.ac.uk/ena/, or http://trace.ddbj.nig.ac.jp/dra/index_e.shtml; Shumway et al., 2010; Leinonen et al., 2011). The in silico approach provides a lower cost option with fewer lab resources required compared with direct sequencing of a particular gene from pooled DNA samples for SNP discovery. In addition, the probability of discovering SNP using the in silico approach may be greater than direct sequencing as a result of accumulation of new sequences over time in the public domain as well as these sequences might be from different populations increasing the possibility of finding new SNP.

Markers deviating from HWE indicate problems with genotyping or population stratification. Because of the missing class of genotypes within some SNP, the association analysis results must be viewed with caution. Nonetheless, some deviation from HWE indicates a potential association between a particular marker and the trait of interest (Wigginton et al., 2005). The functional mutation might have a rare allele that can be missing in some breeds or in populations within a breed (Goddard, 2009). Furthermore, the genetic markers that are linked to the QTL with large effects within a particular breed contribute to the composite or cross (Piyasatian et al., 2006). Bansal et al. (2010) discussed many reasons for considering rare variants as a source of variation. Therefore, in the current study, associations were tested for rare variants or for SNP that are not in HWE (minor allele frequency of less than 10%) as these might be informative or provide promising results that could be considered in crossbred or multibreed populations. However, the obtained significant associations should be validated in other populations.

Table 8. Estimates of allele substitution effects and genotypic effects (least squares means) of SNP in gene cholecystokinin B receptor (*CCKBR*) in the discovery population

		Allele substitution effect				Genotype as fixed effect		
						LSM[3] ± SE	LSM ± SE	LSM ± SE
Trait[1]	SNP name	*P*-value	RA[2]	Estimate ± SE	*P*-value	C/C (G/G)	C/T (A/G)	T/T (A/A)
ADG	rs42670353	0.006**	T	–0.043 ± 0.02	0.019*	1.73 ± 0.05	1.67 ± 0.05	1.64 ± 0.05
	rs42670351	0.135	C	0.03 ± 0.02	0.112	1.66 ± 0.05	1.70 ± 0.05	1.65 ± 0.05
	rs42670352	0.427	G	–0.024 ± 0.03	0.620	1.63 ± 0.05	1.56 ± 0.05	1.67 ± 0.05
DMI	rs42670352	0.055†	G	–0.236 ± 0.12	0.033*	9.14 ± 0.20	8.28 ± 0.20	9.52 ± 0.20
	rs42670351	0.467	C	–0.06 ± 0.08	0.069†	9.08 ± 0.20	10.00 ± 0.20	9.53 ± 0.20
	rs42670353	0.563	T	–0.038 ± 0.07	0.828	9.61 ± 0.20	9.55 ± 0.20	9.53 ± 0.20
FCR	rs42670353	0.033*	T	0.114 ± 0.05	0.055†	5.92 ± 0.16	6.12 ± 0.16	6.16 ± 0.16
	rs42670351	0.194	C	–0.087 ± 0.07	0.405	6.00 ± 0.15	6.14 ± 0.151	6.21 ± 0.15
	rs42670352	0.673	G	–0.041 ± 0.1	0.727	6.12 ± 0.17	5.82 ± 0.17	6.17 ± 0.17
MMWT	rs42670353	0.229	T	–0.523 ± 0.43	0.306	103.44 ± 1.33	102.28 ± 1.33	102.3 ± 1.33
	rs42670351	0.391	C	–0.467 ± 0.54	0.689	101.42 ± 1.24	102.01 ± 1.24	102.45 ± 1.24
	rs42670352	0.490	G	–0.55 ± 0.8	0.396	100.33 ± 1.36	105.66 ± 1.36	101.82 ± 1.36
RFI	rs42670352	0.053	G	–0.175 ± 0.09	0.002‡	–0.42 ± 0.15	–1.61 ± 0.15	–0.18 ± 0.15
	rs42670353	0.280	T	0.052 ± 0.05	0.434	–0.30 ± 0.15	–0.19 ± 0.15	–0.19 ± 0.15
	rs42670351	0.355	C	–0.056 ± 0.06	0.028*	–0.60 ± 0.15	–0.16 ± 0.15	–0.21 ± 0.15

†Tended to affect the trait before the modified Bonferroni adjustment for multiple testing ($P < 0.10$)

‡Significant after the modified Bonferroni adjustment for multiple testing ($P < 0.05$)

*Significant effect before the modified Bonferroni adjustment for multiple testing ($P < 0.05$)

**Significant effect before the modified Bonferroni adjustment for multiple testing ($P < 0.01$)

[1]ADG = average daily gain (kg d^{-1}); DMI = daily dry matter intake (kg d^{-1}); FCR = feed conversion ratio (kg gain kg^{-1} DM); MMWT = midpoint metabolic weight (kg); RFI = residual feed intake (kg d^{-1}).

[2]RA = substitution allele.

[3]LSM = least squares mean.

The extent of LD in the current study was measured using r^2 as it is less dependent on allele frequencies or affected by small sample size (Ardlie et al., 2002; McRae et al., 2002; Khatkar et al., 2008). Generally, the magnitude of LD between SNP pairs was greater in some genes than in others as it is due to SNP density (i.e., the relationship between distance and r^2 was high; Khatkar et al., 2008). The magnitude of LD present between SNP pairs in genes *CCKBR* and *UCP2* was less than 0.1, which is expected as they were up to 7.1 Mbp apart (Sargolzaei et al., 2008). The LD between SNP CT2 in gene *LOC780933* and SNP TRYP81 in gene *PRSS2* was 0.243 as the distance between these 2 SNP is 0.75 Mbp, indicating they may capture some of the same effects. The SNP ATPase4 (rs43562809) is sufficient to capture 100% of the genetic variation explained by SNP ATPase1 in gene *ATP6V1B2* using haplotype tagging from the Haploview analysis. In addition, SNP PC2 can capture 100% of the variance explained by SNP PC3 and PC6. Consequently, not all SNP would be selected for genotyping in the validation population to reduce costs.

Gene *PRSS2* (SNP rs41256901) was significantly associated with RFI and FCR and suggestively associated with DMI. The identified significant associations were not found in the validation study. These associations are in agreement with the significant relationship between feed efficiency and digestive function reported by Richardson and Herd (2004) where the digestibility of feed accounted for 10 to 14% of the variation in feed efficiency. Pancreatic enzymes may be partially responsible for the variation in digestive efficiency between animals (Swanson et al., 2004). Conversely, there was no significant relationship between either performance or feed efficiency and the concentration of the pancreatic trypsin enzyme in feedlot cattle (Mader et al., 2009). Trypsinogen can be an activator of proteinase-activated receptor 2 (**PAR-2**), which is highly expressed in digestive organs, such as the pancreas and intestine, and stimulates many biological processes, such as cell proliferation (Ossovskaya and Bunnett, 2004). In mice, downregulation of trypsinogen was associated with growth retardation in α1, 6-fucosyltransferase-knockout mice (Li et al., 2006).

Results from the current association analysis indicate that there were significant associations between gene *CCKBR*, represented in SNP rs42670351 and rs42670352, and RFI, DMI, ADG, and FCR. Substitution to the C allele in SNP rs42670351 was associated with decreasing DMI, RFI, and FCR in the validation population. Also, the G allele of SNP CCKBR2 (rs42670352) was validated to be associated with decreasing DMI and RFI. In addition, substitution to the G allele of SNP CCKBR3

Table 9. Estimates of allele substitution effects on feed efficiency traits in the validation study

BTA: gene ID[1]	Trait[2]	SNP name	Ref_SNP	*n*[3]	MA[4]	MAF[5]	Estimate ± SE	*P*-value[6]
8: 338082	MMWT	ATPase4	rs43562809	726	A	0.073	1.292 ± 0.76	0.0897 d
8: 338082	FCR	ATPase2	rs43562811	726	C	0.252	–0.139 ± 0.06	0.0218 d
8: 338082	ADG	ATPase2	rs43562811	726	C	0.252	0.037 ± 0.02	0.0277
8: 338082	MMWT	ATPase2	rs43562811	726	C	0.252	0.941 ± 0.45	0.0378
13: 404172	MMWT	PTI1	rs43024409	698	T	0.352	0.973 ± 0.43	0.0234
13: 404172	ADG	PTI1	rs43024409	698	T	0.352	0.026 ± 0.02	0.0970
15: 281665†	DMI	CCRB3	rs42670353	725	G	0.420	–0.251 ± 0.07	0.0008 d†
15: 281665	RFI	CCRB3	rs42670353	725	G	0.420	–0.159 ± 0.06	0.0135
15: 281665	FCR	CCRB3	rs42670353	725	G	0.420	–0.125 ± 0.05	0.0168 sd*
15: 281665	DMI	CCRB2	rs42670352	725	G	0.217	–0.222 ± 0.09	0.0116 sd*
15: 281665	RFI	CCRB2	rs42670352	725	G	0.217	–0.139 ± 0.08	0.0658 sd*
15: 281665	FCR	CCRB2	rs42670352	725	G	0.217	–0.099 ± 0.06	0.106 d
15: 281665†	DMI	CCRB1	rs42670351	700	C	0.221	–0.235 ± 0.09	0.0084 d
15: 281665	RFI	CCRB1	rs42670351	700	C	0.221	–0.164 ± 0.08	0.0315 d*
15: 281665	FCR	CCRB1	rs42670351	700	C	0.221	–0.117 ± 0.06	0.0589 d

†Significant after the modified Bonferroni adjustment for multiple testing ($P < 0.05$) in the validation population.

[1]BTA = *Bos taurus* autosome; gene ID = Entrez Gene Identifier.

[2]ADG = average daily gain (kg d^{-1}); DMI = daily dry matter intake (kg d^{-1}); FCR = feed conversion ratio (kg gain kg^{-1} DM); MMWT = midpoint metabolic weight (kg); RFI = residual feed intake (kg d^{-1}).

[3]*n* = the number of records used in the analyses.

[4]MA = minor allele.

[5]MAF = minor allele frequency.

[6]d = The same direction but was not significant in the discovery population; d† = the same direction but showed a trend ($P = 0.069$) in the discovery population using the genotypic model; sd* = the same direction and significance was found for both the discovery and validation populations; d* = the same direction and significance was found for both the discovery and validation populations using the genotypic model.

Table 10. The genotypic analysis for SNP affecting feed efficiency traits in the validation population

BTA[1]	Gene ID[2]	SNP name	Ref. SNP[3]	MA[4]	Trait[5]	*P*-value	MAF[6]	LSM[7] ± SE C/C (GG)	LSM ± SE C/T (A/G)	LSM ± SE T/T (A/A)
8	338082	ATPase2	rs43562811	C	ADG	0.058	0.255	1.76 ± 0.03	1.70 ± 0.03	1.67 ± 0.03
8	338082	ATPase2	rs43562811	C	FCR	0.067	0.255	6.59 ± 0.12	6.75 ± 0.12	6.89 ± 0.12
8	338082	ATPase2	rs43562811	C	MMWT	0.037	0.255	106.33 ± 0.87	104.09 ± 0.87	103.53 ± 0.87
13	404172	PTI1	rs43024409	T	MMWT	0.082	0.36	105.03 ± 0.88	104.41 ± 0.88	103.31 ± 0.88
15	281665	CCKBR1	rs42670351	C	DMI	0.009	0.216	10.22 ± 0.16	10.69 ± 0.16	10.87 ± 0.16
15	281665	CCKBR1	rs42670351	C	FCR	0.085	0.216	6.58 ± 0.12	6.66 ± 0.12	6.82 ± 0.12
15	281665	CCKBR1	rs42670351	C	RFI	0.014	0.216	–0.18 ± 0.13	0.04 ± 0.13	0.25 ± 0.13
15	281665	CCKBR2	rs42670352	G	DMI	0.011	0.213	10.22 ± 0.16	10.689 ± 0.16	10.85 ± 0.16
15	281665	CCKBR2	rs42670352	G	FCR	0.107	0.213	6.63 ± 0.12	6.72 ± 0.12	6.86 ± 0.12
15	281665	CCKBR2	rs42670352	G	RFI	0.024	0.213	–0.16 ± 0.13	0.085 ± 0.13	0.26 ± 0.13
15	281665	CCKBR3	rs42670353	G	DMI	0.002	0.423	10.48 ± 0.16	10.70 ± 0.16	10.98 ± 0.16
15	281665	CCKBR3	rs42670353	G	FCR	0.048	0.423	6.68 ± 0.1178	6.80 ± 0.1178	6.94 ± 0.1178
15	281665	CCKBR3	rs42670353	G	RFI	0.005	0.423	0.0003 ± 0.13	0.13 ± 0.13	0.38 ± 0.13
29	338471	PC5	rs42197375	C	MMWT	0.062	0.119	108.98 ± 0.87	103.986 ± 0.87	104.56 ± 0.87

[1]BTA = *Bos taurus* autosome.

[2]Gene ID = Entrez Gene Identifier.

[3]Ref. SNP = SNP reference number at the National Center for Biotechnology Information.

[4]MA = minor allele.

[5]ADG = average daily gain (kg d^{-1}); DMI = daily dry matter intake (kg d^{-1}); FCR = feed conversion ratio (kg gain kg^{-1} DM); MMWT = midpoint metabolic weight (kg); RFI = residual feed intake (kg d^{-1})

[6]MAF = minor allele frequency.

[7]LSM = least squares mean.

(rs42670353) was validated to be associated with decreasing FCR and found to be significantly associated with decreasing DMI and RFI. The identified significant SNP in gene *CCKBR* were synonymous, so they might be in LD with the functional mutation. Recently, a missense mutation (rs133526822) was identified using a whole-genome sequencing method (Kawahara-Miki et al., 2011) where SNP rs133526822 is located between SNP rs42670351 and rs42670352. Therefore, further investigation of SNP rs133526822 is required as it might be the functional mutation responsible for the reported significant association. In pigs, Houston et al. (2006, 2008) reported an association between polymorphisms in the 5′-untranslated region of the porcine CCK type A receptor gene with feed intake and growth. Gene *CCKBR* is expressed in gastric parietal cells, the brain, and smooth muscle (Huppi et al., 1995; Wank, 1995). The significant associations with feed efficiency and performance found in the current study are consistent with the functions of gene *CCKBR*. Gene *CCKBR* is predominant in the hypothalamus and is also expressed in the vagus nerve stem complex, so it plays a very important role as a mediator in the satiety effect of CCK (Dufresne et al., 2006), affecting feed intake and efficiency. Gene *CCKBR* is the predominant CCK receptor subtype for the veal and weaned calves (Le Meuth et al., 1993). As in calves, CCKBR are predominant in the pancreas of pigs (Philippe et al., 1997). Pancreatic enzyme secretion was mediated by CCKBR under stimulation by the physiological levels of CCK and gastrin (Le Dréan et al., 1999). Also, pancreatic growth and secretion were regulated by CCKBR particularly after weaning (Le Meuth et al., 1993). Therefore, association between polymorphisms in gene CCKBR might be associated with pancreas growth or secretion, suggesting further study to test these biological relationships.

Conclusion

The in silico study was an effective method for SNP discovery in candidate genes. New SNP were reported in genes *PRSS2* and *CCKBR* that have an association with feed efficiency and performance traits in these data. The SNP rs42670352 in *CCKBR* was significantly associated with RFI and DMI in the discovery and validation populations and had the same phase of associations. In addition, SNP rs42670353 in *CCKBR* was significantly associated with FCR in the discovery population with same phase of association in the validation populations. Investigating the biological mechanisms underpinning these discoveries by studying gene expression (RNA and protein abundance) will also increase our understanding of the underlying biology of these SNP.

LITERATURE CITED

Abecasis, R. G., and W. O. C. Cookson. 2000. GOLD – Graphical overview of linkage disequilibrium. Bioinformatics 16:182–183.

Archer, J. A., E. C. Richardson, R. M. Herd, and P. F. Arthur. 1999. Potential for selection to improve efficiency of feed use in beef cattle: A review. Austr. J. Agric. Res. 50:147–161.

Ardlie, K. G., L. Kruglyak, and M. Seielstad. 2002. Patterns of linkage disequilibrium in the human genome. Nat. Rev. Genet. 3:299–309.

Arthur, P. F., G. Renand, and D. Krauss. 2001. Genetic and phenotypic relationships among different measures of growth and feed efficiency in young charolais bulls. Livest. Prod. Sci. 68:131–139.

Ascenzi, P., A. Bocedi, M. Bolognesi, A. Spallarossa, M. Coletta, R. De Cristofaro, and E. Menegatti. 2003. The bovine basic pancreatic trypsin inhibitor (kunitz inhibitor): A milestone protein. Curr. Protein Peptide Sci. 4:231–251.

Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. van Duijn. 2007. GenABEL: An R package for genome-wide association analysis. Bioinformatics 23:1294–1296.

Bansal, V., O. Libiger, A. Torkamani, and N. J. Schork. 2010. Statistical analysis strategies for association studies involving rare variants. Nat. Rev. Genet. 11:773–785.

Barendse, W., A. Reverter, R. J. Bunch, B. E. Harrison, W. Barris, and M. B. Thomas. 2007. A validated whole-genome association study of efficient food conversion in cattle. Genetics 176:1893–1905.

Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265.

Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler. 2005. GenBank. Nucleic Acids Res. 33:D34–D38.

Benson, D. A., I. Karsch-Mizrachi, K. Clark, D. J. Lipman, J. Ostell and E. W. Sayers. 2012. GenBank. Nucleic Acids Research, 40: 48–53.

Caldarelli-Stefano, R., L. Vago, S. Bonetto, M. Nebuloni, and G. Costanzi. 1999. Use of magnetic beads for tissue DNA extraction and IS6110 Mycobacterium tuberculosis PCR. Mol. Pathol. 52:158–160.

Chapinal, N., D. M. Veira, D. M. Weary, and M. A. G. von Keyserlingk. 2007. Technical note: Validation of a system for monitoring individual feeding and drinking behavior and intake in group-housed cattle. J. Dairy Sci. 90:5732–5736.

Chen, J. M., and C. Ferec. 2000. Molecular basis of hereditary pancreatitis. Eur. J. Hum. Genet. 8:473–479.

Cox, D. G., C. Boillot, and F. Canzian. 2001. Data mining: Efficiency of using sequence databases for polymorphism discovery. Hum. Mutat. 17:141–150.

De Bakker, P. I., R. Yelensly, I. Pe'er, S. B. Gabriel, M. J. Daly, and D. Altshuler. 2005. Efficiency and power in genetic association studies. Nat. Genet. 37:1217–1223.

Dufresne, M., C. Seva, and D. Fourmy. 2006. Cholecystokinin and gastrin receptors. Physiol. Rev. 86:805–847.

Ferris, C. P., T. W. J. Keady, F. J. Gordon, and D. J. Kilpatrick. 2007. Comparison of a Calan gate and a conventional feed barrier system for dairy cows: Feed intake and behaviour. Ir. J. Agric. Food Res. 45:149–156.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. ASReml user guide release 3.0. VSN International Ltd., Hemel Hempstead, UK.

Goddard, M. 2009. How can we best use DNA data in selection of cattle? In: Proceedings of the Beef Improvement Federation 41st Annual Research Symposium. April 30–May 3, 2009, Sacramento, CA. p. 81-91.

Greenfield, R. B., M. J. Cecava, and S. S. Donkin. 2000. Changes in mRNA expression for gluconeogenic enzymes in liver of dairy cattle during the transition to lactation. J. Dairy Sci. 83:1228–1236.

Haga, S., S. Fujimoto, T. Yonezawa, K. Yoshioka, H. Shingu, Y. Kobayashi, T. Takahashi, Y. Otani, K. Katoh, and Y. Obara. 2008. Changes in hepatic key enzymes of dairy calves in early weaning production systems. J. Dairy Sci. 91:3156–3164.

Herd, R. M., J. A. Archer, and P. F. Arthur. 2003. Reducing the cost of beef production through genetic improvement in residual feed intake: Opportunity and challenges to application. J. Anim. Sci. 81:9–17.

Herd, R. M., V. H. Oddy, and E. C. Richardson. 2004. Biological basis for variation in residual feed intake in beef cattle. 1. Review of potential mechanisms. Aust. J. Exp. Agric. 44(4–5):423–430.

Houston, R. D., C. S. Haley, A. L. Archibald, N. D. Cameron, G. S. Plastow, and K. A. Rance. 2006. A polymorphism in the 5'-untranslated region of the porcine cholecystokinin type a receptor gene affects feed intake and growth. Genetics 174:1555–1563.

Houston, R. D., K. A. Rance, E. Sutcliffe, A. L. Archibald, and C. S. Haley. 2008. The cholecystokinin type A receptor g.179A > G polymorphism affects feeding rate. Anim. Genet. 39:187–188.

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

Huppi, K., D. Siwarski, J. R. Pisegna, and S. Wank. 1995. Chromosomal localization of the gastric and brain receptors for cholecystokinin (CCKAR and CCKBR) in human and mouse. Genomics 25:727–729.

Jefferies, K. C., D. J. Cipriano, and M. Forgac. 2008. Function, structure and regulation of the vacuolar (H+)-ATPases. Arch. Biochem. Biophys. 476:33–42.

Jitrapakdee, S., and J. C. Wallace. 1999. Structure, function and regulation of pyruvate carboxylase. Biochem. J. 340:1–16.

Kawahara-Miki, R., K. Tsuda, Y. Shiwa, Y. Arai-Kichise, T. Matsumoto, Y. Kanesaki, S. Oda, S. Ebihara, S. Yajima, H. Yoshikawa, and T. Kono. 2011. Whole-genome resequencing shows numerous genes with nonsynonymous SNPs in the Japanese native cattle Kuchinoshima-Ushi. BMC Genomics 12:103.

Kereszturi, E., R. Szmola, Z. Kukor, P. Simon, F. U. Weiss, M. M. Lerch, and M. Sahin-Tóth. 2009. Hereditary pancreatitis caused by mutation-induced misfolding of human cationic trypsinogen: A novel disease mechanism. Hum. Mutat. 30:575–582.

Khatkar, M. S., F. W. Nicholas, A. R. Collins, K. R. Zenger, J. A. L. Cavanagh, W. Barris, R. D. Schnabel, J. F. Taylor, and H. W. Raadsma. 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics 9:187.

Kitts, A., and S. Sherry. 2007. The single nucleotide polymorphism database (dbSNP) of nucleotide sequence variation. In: J. McEntyre and J. Ostell, editors, The NCBI handbook. Bethesda (MD): National Center for Biotechnology Information (US); 2002-. Chapter 5. Available from: http://www.ncbi.nlm.nih.gov/books/NBK21088/.

Klukas, C., and F. Schreiber. 2007. Dynamic exploration and editing of KEGG pathway diagrams. Bioinformatics 23:344–350.

Koch, R. M., L. A. Swiger, D. Chambers, and K. E. Gregory. 1963. Efficiency of feed use in beef cattle. J. Anim. Sci. 22:486–494.

Kolath, W. H., M. S. Kerley, J. W. Golden, and D. H. Keisler. 2006. The relationship between mitochondrial function and residual feed intake in angus steers. J. Anim. Sci. 84:861–865.

Le Dréan, G., I. Le Huerou-Luron, M. Gestin, C. Desbois, V. Rome, C. Bernard, M. Dufresne, L. Moroder, D. Gully, J. A. Chayvialle, D. Fourmy, and P. Guilloteau. 1999. Exogenous CCK and gastrin stimulate pancreatic exocrine secretion via CCK-A but also via CCK-B/gastrin receptors in the calf. Pflugers Arch. 438:86–93.

Le Huerou, I., C. Wicker, P. Guilloteau, R. Toullec, and A. Puigserver. 1990. Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic trypsinogen. structural identity within the trypsin family. Eur. J. Biochem. 193:767–773.

Le Meuth, V., V. Philouze-Rome, I. Le Huerou-Luron, M. Formal, N. Vaysse, C. Gespach, P. Guilloteau, and D. Fourmy. 1993. Differential expression of A- and B-subtypes of cholecystokinin/gastrin receptors in the developing calf pancreas. Endocrinology 133:1182–1191.

Leinonen, R., H. Sugawara, and M. Shumway. 2011. The sequence read archive. Nucleic Acids Res. 39:D19–21.

Li, W., T. Nakagawa, N. Koyama, X. Wang, J. Jin, Y. Mizuno-Horikawa, J. Gu, E. Miyoshi, I. Kato, K. Honke, N. Taniguchi, and A. Kondo. 2006. Down-regulation of trypsinogen expression is associated with growth retardation in {alpha}1,6-fucosyltransferase-deficient mice: Attenuation of proteinase-activated receptor 2 activity. Glycobiology 16:1007–1019.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. 1st ed. Sinauer Associates, Inc., Sunderland, MA.

Mader, C. J., Y. R. Montanholi, Y. J. Wang, S. P. Miller, I. B. Mandell, B. W. McBride, and K. C. Swanson. 2009. Relationships among measures of growth performance and efficiency with carcass traits, visceral organ mass, and pancreatic digestive enzymes in feedlot cattle. J. Anim. Sci. 87:1548–1557.

Makridakis, N. M., and J. K. V. Reichardt. 2001. Multiplex automated primer extension analysis: Simultaneous genotyping of several polymorphisms. BioTechniques 31:1374–1380.

Mantel, N. 1980. Assessing laboratory evidence for neoplastic activity. Biometrics 36:381–399.

Marth, G.T., I. Korf, M. D. Yandell, R. T. Yeh, Z. Gu, H. Zakeri, N. O. Stitziel, L. Hillier, P. Y. Kwok, and W. R. Gish. 1999. A general approach to single-nucleotide polymorphism discovery. Nat. Genet. 23: 452–456.

McRae, A. F., J. C. McEwan, K. G. Dodds, T. Wilson, A. M. Crawford, and J. Slate. 2002. Linkage disequilibrium in domestic sheep. Genetics 160:1113–1122.

Morgulis, A., G. Coulouris, Y. Raytselis, T. L. Madden, R. Agarwala and A. A. Schaffer. 2008. Database indexing for production MegaBLAST searches. Bioinformatics, 24:1757–1764.

Nkrumah, J. D., E. L. Sherman, C. Li, E. Marques, D. H. Crews Jr., R. Bartusiak, B. Murdoch, Z. Wang, J. A. Basarab, and S. S. Moore. 2007. Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. J. Anim. Sci. 85:3170–3181.

Nickerson, D.A., V.O. Tobe and S. L. Taylor. 1997. PolyPhred: Automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res. 25: 2745–2751.

Ossovskaya, V. S., and N. W. Bunnett. 2004. Protease-activated receptors: Contribution to physiology and disease. Physiol. Rev. 84(2):579–621.

Pastinen, T., M. Raitio, K. Lindroos, P. Tainola, L. Peltonen, and A. C. Syvanen. 2000. A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. Genome Res. 10:1031–1042.

Philippe, C., E. F. Lhoste, M. Dufresne, L. Moroder, T. Corring, and D. Fourmy. 1997. Pharmacological and biochemical evidence for the simultaneous expression of CCKB/gastrin and CCKA receptors in the pig pancreas. Br. J. Pharmacol. 120:447–454.

Picoult-Newberg, L., T. E. Ideker, M. G. Pohl, S. L. Taylor, M. A. Donaldson, D. A. Nickerson, and M. Boyce-Jacino. 1999. Mining SNPs from EST databases. Genome Res. 9:167–174.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and the R Development Core Team. 2011. nlme: Linear and nonlinear mixed effects models. R package version 3.1-102.

Piyasatian, N., L. R. Totir, R. L. Fernando, and J. C. M. Dekkers. 2006. QTL detection and marker-assisted composite line development. J. Anim. Sci. 84:134–135.

Rehfeld, J. F., L. Friis-Hansen, J. P. Goetze, and T. V. Hansen. 2007. The biology of cholecystokinin and gastrin peptides. Curr. Top. Med. Chem. 7:1154–1165.

Richardson, E. C., and R. M. Herd. 2004. Biological basis for variation in residual feed intake in beef cattle. 2. Synthesis of results following divergent selection. Aust. J. Exp. Agric. 44:431–440.

Rudi, K., M. Kroken, O. J. Dahlberg, A. Deggerdal, K. S. Jakobsen, and F. Larsen.1997. Rapid, universal method to isolate PCR-ready DNA using magnetic beads. Biotechniques 22:506–511.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual. 2nd ed. Cold Spring Harbour Lab Press, New York, NY.

Sargolzaei, M., F. S. Schenkel, G. B. Jansen, and L. R. Schaeffer. 2008. Extent of linkage disequilibrium in Holstein cattle in North America. J. Dairy Sci. 91:2106–2117.

Scheet, P., and M. Stephens. 2006. A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. Am. J. Hum. Genet. 78:629–644.

Schenkel, F. S., S. P. Miller, and J. W. Wilton. 2004. Genetic parameters and breed differences for feed efficiency, growth, and body composition traits of young beef bulls. Can. J. Anim. Sci. 84:177–185.

Sherman, E. L., J. D. Nkrumah, and S. S. Moore. 2010. Whole genome single nucleotide polymorphism associations with feed intake and feed efficiency in beef cattle. J. Anim. Sci. 88:16–22.

Sherman, E. L., J. D. Nkrumah, B. M. Murdoch, C. Li, Z. Wang, A. Fu, and S. S. Moore. 2008b. Polymorphisms and haplotypes in the bovine neuropeptide Y, growth hormone receptor, ghrelin, insulin-like growth factor 2, and uncoupling proteins 2 and 3 genes and their associations with measures of growth, performance, feed efficiency, and carcass merit in beef cattle. J. Anim. Sci. 86:1–16.

Sherman, E. L., J. D. Nkrumah, B. M. Murdoch, and S. S. Moore. 2008a. Identification of polymorphisms influencing feed intake and efficiency in beef cattle. Anim. Genet. 39:225–231.

Shumway, M., C. Cochrane, and H. Sugawara. 2010. Archiving next generation sequencing data. Nucleic Acids Res. 38:D870–871.

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(1):31–33.

Sun-Wada, G., T. Toyomura, Y. Murata, A. Yamamoto, M. Futai, and Y. Wada. 2006. The a3 isoform of V-ATPase regulates insulin secretion from pancreatic β-cells. J. Cell Sci. 119:4531–4540.

Swanson, K. C., J. A. Benson, J. C. Matthews, and D. L. Harmon. 2004. Pancreatic exocrine secretion and plasma concentration of some gastrointestinal hormones in response to abomasal infusion of starch hydrolyzate and/or casein. J. Anim. Sci. 82:1781–1787.

Swanson, K. C., C. J. Richards, and D. L. Harmon. 2002. Influence of abomasal infusion of glucose or partially hydrolyzed starch on pancreatic exocrine secretion in beef steers. J. Anim. Sci. 80:1112–1116.

Teich, N., and J. Mossner. 2008. Hereditary chronic pancreatitis. Best Pract. Res. Clin. Gastroenterol. 22:115–130.

Van den Borne, J. J. G. C., B. R. Westrom, D. Kruszewska, J. A. M. Botermans, J. Svendsen, J. Woliński, and S. G. Pierzynowski. 2007. Exocrine pancreatic secretion in pigs fed sow's milk and milk replacer, and its relationship to growth performance. J. Anim. Sci. 85:404–412.

Velez, J. C., and S. S. Donkin. 2005. Feed restriction induces pyruvate carboxylase but not phosphoenolpyruvate carboxykinase in dairy cows. J. Dairy Sci. 88:2938–2948.

Wang, N., J. M. Akey, K. Zhang, R. Chakraborty, and L. Jin. 2002. Distribution of recombination crossovers and the origin of haplotype blocks: The interplay of population history, recombination, and mutation. Am. J. Hum. Genet. 71:1227–1234.

Wank, S. A. 1995. Cholecystokinin receptors. Am J Physiol., 269:628 – 646.

Weckx, S., P. De Rijk, C. Van Broeckhoven, and J. Del-Favero. 2004. SNPbox: Web-based high-throughput primer design from gene to genome. Nucleic Acids Res. 32:W170–172.

Weckx, S., J. Del-Favero, R. Rademakers, L. Claes, M. Cruts, P. De Jonghe, C. Van Broeckhoven, and P. De Rijk. 2005. novoSNP, a novel computational tool for sequence variation discovery. Genome Res. 15:436–442.

Wigginton, J. E., D. J. Cutler, and G. R. Abecasis. 2005. A note on exact tests of hardy-weinberg equilibrium. Am. J. Hum. Genet. 76:887–893.

Wulfhorst, J. D., J. K. Ahola, S. L. Kane, L. D. Keenan, and R. A. Hill. 2010. Factors affecting beef cattle producer perspectives on feed efficiency. J. Anim. Sci. 88:3749–3758.

Zhang, C., G. Baffy, P. Perret, S. Krauss, O. Peroni, D. Grujic, T. Hagen, A. J. Vidal-Puig, O. Boss, Y. Kim, X. X. Zheng, M. B. Wheeler, G. I. Shulman, C. B. Chan, and B. B. Lowell. 2001. Uncoupling protein-2 negatively regulates insulin secretion and is a major link between obesity, cell dysfunction, and type 2 diabetes. Cell 105:745–755.

Zhang, Z., S. Schwartz, L. Wagner, and W. Miller. 2000. A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7(1–2):203–214.

Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. P. Van Tassell, T. S. Sonstegard, G. Marcais, M. Roberts, P. Subramanian, J. A. Yorke, S. L. Salzberg. 2009. A whole-genome assembly of the domestic cow Bos taurus. Genome Biol 10:42.

APPENDIX 1. Biological mechanisms, molecular function, and pathways associated with the candidate genes in the study

Category[1]	Gene ontology term	*P*-value[2]	Genes
BP	GO:0007586~digestion	0.013025	282603, 780933
PATH	bta04080:Neuroactive ligand-receptor interaction	0.021155	282603, 780933, 281665
MF	GO:0004252~serine-type endopeptidase activity	0.062294	282603, 780933
MF	GO:0008236~serine-type peptidase activity	0.070418	282603, 780933
MF	GO:0017171~serine hydrolase activity	0.070923	282603, 780933
MF	GO:0004175~endopeptidase activity	0.156439	282603, 780933
MF	GO:0070011~peptidase activity, acting on L-amino acid peptides	0.201141	282603, 780933
BP	GO:0055085~transmembrane transport	0.202684	281562, 338082
MF	GO:0008233~peptidase activity	0.207363	282603, 780933
MF	GO:0016787~hydrolase activity	0.212711	282603, 780933, 338082
MF	GO:0005509~calcium ion binding	0.252329	282603, 780933
BP	GO:0006508~proteolysis	0.304393	282603, 780933
MF	GO:0003824~catalytic activity	0.343324	282603, 780933, 338082, 338471
MF	GO:0046872~metal ion binding	0.378607	282603, 780933, 338471
MF	GO:0043169~cation binding	0.384751	282603, 780933, 338471
MF	GO:0043167~ion binding	0.390012	282603, 780933, 338471
MF	GO:0005515~protein binding	0.704959	282603, 780933, 281665
MF	GO:0005488~binding	0.773127	282603, 780933, 281562, 281665, 338471
MF	GO:0004857~enzyme inhibitor activity	1	404172
MF	GO:0004866~endopeptidase inhibitor activity	1	404172
MF	GO:0004867~serine-type endopeptidase inhibitor activity	1	404172
BP	GO:0005996~monosaccharide metabolic process	1	338471
BP	GO:0006006~glucose metabolic process	1	338471
BP	GO:0006091~generation of precursor metabolites and energy	1	338082
BP	GO:0006094~gluconeogenesis	1	338471
BP	GO:0006119~oxidative phosphorylation	1	338082
BP	GO:0006163~purine nucleotide metabolic process	1	338082
BP	GO:0006754~ATP biosynthetic process	1	338082
BP	GO:0006793~phosphorus metabolic process	1	338082
BP	GO:0006811~ion transport	1	338082
BP	GO:0006812~cation transport	1	338082
BP	GO:0006818~hydrogen transport	1	338082
BP	GO:0006839~mitochondrial transport	1	281562
BP	GO:0006874~cellular calcium ion homeostasis	1	281665
BP	GO:0007166~cell surface receptor linked signal transduction	1	281665
BP	GO:0007186~G-protein coupled receptor protein signaling pathway	1	281665
BP	GO:0008284~positive regulation of cell proliferation	1	281665
BP	GO:0008610~lipid biosynthetic process	1	338471
BP	GO:0009057~macromolecule catabolic process	1	282603
BP	GO:0009141~nucleoside triphosphate metabolic process	1	338082
BP	GO:0009205~purine ribonucleoside triphosphate metabolic process	1	338082
BP	GO:0015985~energy coupled proton transport, down electrochemical gradient	1	338082
BP	GO:0015986~ATP synthesis coupled proton transport	1	338082
BP	GO:0015992~proton transport	1	338082
BP	GO:0016051~carbohydrate biosynthetic process	1	338471
BP	GO:0016310~phosphorylation	1	338082
BP	GO:0019318~hexose metabolic process	1	338471
BP	GO:0030003~cellular cation homeostasis	1	281665
BP	GO:0030163~protein catabolic process	1	282603
MF	GO:0030234~enzyme regulator activity	1	404172
BP	GO:0032963~collagen metabolic process	1	282603
BP	GO:0034220~ion transmembrane transport	1	338082
BP	GO:0034404~nucleobase, nucleoside and nucleotide biosynthetic process	1	338082
BP	GO:0034637~cellular carbohydrate biosynthetic process	1	338471
BP	GO:0034654~nucleobase, nucleoside, nucleotide and nucleic acid biosynthetic process	1	338082

Continued

APPENDIX 1. Continued

Category[1]	Gene ontology term	*P*-value[2]	Genes
BP	GO:0042127~regulation of cell proliferation	1	281665
BP	GO:0044257~cellular protein catabolic process	1	282603
BP	GO:0044259~multicellular organismal macromolecule metabolic process	1	282603
BP	GO:0044271~nitrogen compound biosynthetic process	1	338082
BP	GO:0046034~ATP metabolic process	1	338082
BP	GO:0046364~monosaccharide biosynthetic process	1	338471
BP	GO:0046907~intracellular transport	1	281562
BP	GO:0050801~ion homeostasis	1	281665
BP	GO:0051480~cytosolic calcium ion homeostasis	1	281665
BP	GO:0051603~proteolysis involved in cellular protein catabolic process	1	282603
BP	GO:0055065~metal ion homeostasis	1	281665
BP	GO:0055074~calcium ion homeostasis	1	281665
PATH	bta00620:Pyruvate metabolism	1	338471
PATH	bta00500:Starch and sucrose metabolism	1	539383
PATH	bta00190:Oxidative phosphorylation	1	338082
PATH	bta00020:Citrate cycle (TCA cycle)	1	338471
PATH	bta04020:Calcium signaling pathway	1	281665

[1]BP = biological process; PATH = KEGG biological pathway; MF = molecular function.

[2]*P*-value = the *P*-value produced by enrichment analysis using DAVID software (Huang et al., 2009).

APPENDIX 2. Reported QTL in the Animal QTL database overlapping with candidate genes

Gene name	QTL name[1]	QTL start position (bp)	QTL end position	Trait
AMY2B	QTL_10129	34002036	65739450	Milk fat yield (daughter deviation)
AMY2B	QTL_10683	29956341	43937012	Height (mature)
AMY2B	QTL_10684	23481347	43937012	BW (birth)
AMY2B	QTL_10685	23481347	43937012	BW (weaning)
AMY2B	QTL_10686	29956341	43937012	Height (yearling)
AMY2B	QTL_10687	23481347	43937012	Carcass weight
AMY2B	QTL_1326	17737279	42568866	BW (birth)
AMY2B	QTL_1351	37151994	48988111	Marbling score
AMY2B	QTL_2437	17033086	46894623	Milk protein yield
AMY2B	QTL_2442	20288493	43373658	Milk fat percentage
AMY2B	QTL_2443	18441472	63526272	Milk protein percentage
AMY2B	QTL_2490	0	43937012	Somatic Cell Count
AMY2B	QTL_2541	28000529	90939023	Marbling score
AMY2B	QTL_2584	22194674	57994437	Milk yield
AMY2B	QTL_2653	29300268	47891582	Milk yield
AMY2B	QTL_2654	29300268	47891582	Milk protein percentage
AMY2B	QTL_2657	29300268	48706434	Milk protein yield
ATP6V1B2	QTL_10831	62732685	70703039	Calving ease (dire*LOC780933*)
ATP6V1B2	QTL_10832	62732685	83158666	BW (birth)
ATP6V1B2	QTL_11442	46113400	92708598	Dystocia (maternal)
ATP6V1B2	QTL_11443	46113400	92708598	Stillbirth (maternal)
ATP6V1B2	QTL_1683	62735433	83360187	Somatic Cell Count
ATP6V1B2	QTL_1684	62735433	83360187	Stru*LOC780933*ural soundness (legs
ATP6V1B2	QTL_2497	43089984	83569900	Clinical mastitis
ATP6V1B2	QTL_2498	43089984	83569900	Somatic Cell Count
ATP6V1B2	QTL_3599	59682392	103895369	Foot angle
CCKBR	QTL_10993	32063789	47949409	Height (mature)
CCKBR	QTL_10994	32063789	47949409	Carcass weight
CCKBR	QTL_10995	32063789	47949409	Longissimus muscle area
CCKBR	QTL_12195	32881427	55518119	Liver percentage
CCKBR	QTL_1335	21718187	54967641	Kidney
CCKBR	QTL_1596	40000451	47942043	Udder attachment

Continued

APPENDIX 2. Continued

Gene name	QTL name[1]	QTL start position (bp)	QTL end position	Trait
CCKBR	QTL_1598	40000451	47942043	Stature
CCKBR	QTL_1601	40000451	47942043	Udder depth
CCKBR	QTL_1699	15363886	47946463	Rump angle
CCKBR	QTL_2678	40000451	47942043	Somatic cell score
LOC780933	QTL_10515	98051474	119280357	Parasites mean of natural logarithm
LOC780933	QTL_10716	95179745	107915671	BW (mature)
LOC780933	QTL_10717	81300315	107915671	Fat thickness at the 12th rib
LOC780933	QTL_10718	95179745	113949759	Scrotal circumference
LOC780933	QTL_4485	98051474	107905229	Postweaning average daily gain
LOC780933	QTL_5055	57599923	119280357	Milk fat percentage
PC	QTL_11297	37089419	48660164	Carcass weight
PC	QTL_11298	37089419	48660164	BW (yearling)
PC	QTL_11299	37089419	51080436	BW (birth)
PC	QTL_11301	37089419	48660164	BW (weaning)
PC	QTL_1343	34055421	46031383	Retail produ*LOC780933* yield
PC	QTL_1345	27433819	48199160	Tenderness score
PC	QTL_1373	36423729	46550490	Tenderness score
PC	QTL_1374	37521610	51800755	Tenderness score
PC	QTL_1380	32188341	45594750	BW at castration
PC	QTL_1664	35467866	46211463	Foot angle
PC	QTL_1665	35467866	46211463	Stru*LOC780933*ural soundness (legs
PC	QTL_1717	26859725	48660164	Teat placement
PC	QTL_1722	26859725	48660164	Twinning
PC	QTL_2593	18168758	46211940	Milk yield
PC	QTL_2612	18168758	46211940	Milk protein yield
PC	QTL_4506	37089419	48660164	305 d milk yield
PC	QTL_4651	35465705	46211940	Rump angle
PC	QTL_4851	34170672	51606269	Juiciness
PC	QTL_4852	34170672	51606269	Shear force
PC	QTL_4853	34170672	51606269	Tenderness score
PC	QTL_5371	1983481	45887319	Gestation length
PC	QTL_7153	34170672	51606269	Flight from feeder
PC	QTL_7154	34170672	51606269	Flight fm feeder
PC	QTL_7154	34170672	51606269	Flight from feeder
PTI	QTL_10946	50098085	77464923	Weaning weight-maternal milk
PTI	QTL_10947	60604089	77464923	Marbling score (EBV)
PTI	QTL_10948	60604089	77464923	Longissimus muscle area
PTI	QTL_10949	71832245	77464923	BW (weaning)
PTI	QTL_11446	69174400	81815003	Stillbirth (dire*LOC780933*)
PTI	QTL_1386	15506564	84433115	Teat length
PTI	QTL_1584	59740226	79842537	Udder attachment
PTI	QTL_1585	59740226	79842537	PTA type
PTI	QTL_1586	59740226	79842537	Udder height
PTI	QTL_1587	59740226	79842537	Udder width
PTI	QTL_1588	59740226	79842537	Udder depth
PTI	QTL_1589	59740226	79842537	Udder composite index
PTI	QTL_2670	59740226	79842537	Milk yield
PTI	QTL_2671	59740226	79842537	Milk protein yield
PTI	QTL_2775	69333703	77472469	Somatic cell score
PTI	QTL_3569	0	77464923	Heat intensity
PTI	QTL_5011	71832245	77464923	Interval to first estrus after calving
PRSS2	QTL_10515	98051474	119280357	Parasites mean of natural logarithm
PRSS2	QTL_10716	95179745	107915671	BW (mature)
PRSS2	QTL_10717	81300315	107915671	Fat thickness at the 12th rib
PRSS2	QTL_10718	95179745	113949759	Scrotal circumference

Continued

APPENDIX 2. Continued

Gene name	QTL name[1]	QTL start position (bp)	QTL end position	Trait
PRSS2	QTL_4485	98051474	107905229	Postweaning average daily gain
PRSS2	QTL_5055	57599923	119280357	Milk fat percentage
PRSS2	QTL_10515	98051474	119280357	Parasites mean of natural logarithm
PRSS2	QTL_10716	95179745	107915671	BW (mature)
PRSS2	QTL_10717	81300315	107915671	Fat thickness at the 12th rib
PRSS2	QTL_10718	95179745	113949759	Scrotal circumference
PRSS2	QTL_4485	98051474	107905229	Postweaning average daily gain
PRSS2	QTL_5055	57599923	119280357	Milk fat percentage
UCP2	QTL_10996	47949409	60494505	Weaning weight-maternal milk
UCP2	QTL_10997	51606842	60494505	BW (mature)
UCP2	QTL_10998	47949409	60494505	BW (weaning)
UCP2	QTL_10999	51606842	60641830	Marbling score (EBV)
UCP2	QTL_11001	51606842	79392371	Height (mature)
UCP2	QTL_12195	32881427	55518119	Liver percentage
UCP2	QTL_1335	21718187	54967641	Kidney
UCP2	QTL_1594	47946463	57050270	Body form composite index
UCP2	QTL_1595	47946463	57050270	Teat placement
UCP2	QTL_1597	47946463	57050270	PTA type
UCP2	QTL_1599	47946463	57050270	Thurl width
UCP2	QTL_1600	47946463	57050270	Udder cleft
UCP2	QTL_1602	47946463	57050270	Udder composite index
UCP2	QTL_5122	50892329	55864326	Abomasum displacement

[1]QTL = Quantitative trait loci.