



RESEARCH ARTICLE

Identification of Expression QTLs Targeting Candidate Genes for Residual Feed Intake in Dairy Cattle Using Systems Genomics

Salleh MS^{1,2}, Mazzoni G², Nielsen MO¹, Løvendahl P³ and Kadarmideen HN^{2,4*}

¹Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

²Department of Bio and Health Informatics, Technical University of Denmark, Lyngby, Denmark

³Department of Molecular Biology and Genetics-Center for Quantitative Genetics and Genomics, Aarhus University, AU Foulum, Tjele, Denmark

⁴Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark

*Corresponding author: Kadarmideen HN, Department of Applied Mathematics and Computer Science, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark, E-mail: hajak@dtu.dk



Abstract

Background: Residual feed intake (RFI) is the difference between actual and predicted feed intake and an important factor determining feed efficiency (FE). Recently, 170 candidate genes were associated with RFI, but no expression quantitative trait loci (eQTL) mapping has hitherto been performed on FE related genes in dairy cows. In this study, an integrative systems genetics approach was applied to map eQTLs in Holstein and Jersey cows fed two different diets to improve identification of candidate genes for FE.

Methods: Liver RNA-seq transcriptomics data from nine Holstein and ten Jersey cows that had been fed control (C) or high concentrate (HC) diets were integrated with genomic data (from 777k BovineHD Illumina BeadChip) by using the Matrix eQTL R package. A total of 170 previously identified candidate genes for FE (89 differentially expressed genes (DEGs) between high and low RFI groups and 81 hub genes (HG) in a group of co-expressed genes) were used in the data integration analysis.

Results: From the 241,542 SNPs used in the analysis, we identified 20 significant (FDR < 0.05) local-eQTLs targeting seven candidate genes and 16 significant (FDR < 0.05) local-eQTLs targeting five candidate genes related to RFI for the C and HC diet group analysis, respectively, in a breed-specific way.

Conclusions: Interestingly, Holstein and Jersey cows appear to rely on different strategies (lipid and cholesterol metabolism versus immune and inflammatory function) to achieve low RFI. The eQTLs overlapped with QTLs previously associated with FE trait (e.g. dry matter intake, longevity,

body weight gain and net merit). The eQTLs and biological pathways identified in this study improve our understanding of the complex biological and genetic mechanisms that determine FE traits in dairy cattle. The identified eQTLs/genetic variants can potentially be used in new genomic selection methods that include biological/functional information on SNPs.

Keywords

eQTL, RNA-seq, Genotype, Data integration, Systems genomics, Feed efficiency, Residual feed intake

Abbreviations

ANOVA: Analysis of Variance; C: Low Concentrate (Control); DCRC: Danish Cattle Research Centre; DEGs: Differentially Expressed Genes; EDTA: Ethylenediaminetetraacetic Acid; eQTL: Expression Quantitative Trait Loci; FDR: False Discovery Rate; FE: Feed Efficiency; HC: High Concentrate; HG: Hub Genes; Mb: Mega Base; QTL: Quantitative Trait Loci; RFI: Residual Feed Intake; RNA: Ribonucleic Acid; RNA-seq: RNA Sequencing; SNPs: Single Nucleotide Polymorphisms; WGCNA: Weighted Gene Co-expression Analysis

Background

Feed intake and the conversion of absorbed nutrients into milk components are major determinants of feed efficiency (FE) in dairy cattle and hence production economics. FE is a complex trait that is influenced by several genetic and environmental factors, which in an

interactive way control feed intake, nutrient partitioning and metabolic adaptation to lactation in different body tissues as well as milk synthesis and immune function. In dairy cattle, the use of FE for breeding purposes is quite complicated, since recording of individual feed intake is difficult in group fed cows. It is therefore desirable to be able to predict the genetic contributions to this trait to be able to select the most feed efficient cows for breeding purposes.

To date, transcriptomics has given precise and reliable results that identify candidate genes related to phenotypes of interest [1]. Although gene expressions associated with FE related genes have been studied for a long time, also in cattle [2-4], genetic markers are more easily accessible and not affected by environmental factors in contrast to gene expression data.

However, in some cows, the actual feed intake deviate from the predicted by their genetic heritage, even when they are exposed to similar environmental conditions. The term residual feed intake (RFI) describes this deviation and is calculated as the difference between the actual measured and the predicted feed intake of the cow [5]. Among groups of high and low RFI cattle, we have recently identified several candidate genes that predict the RFI in Danish dairy cattle [6].

Therefore, in this present study we focused on genetic markers for RFI in an attempt to improve the prediction of genetic merit for FE, which is needed to be able to use this type of determinants in practice.

Integration of transcriptomics and genomics data can be used to identify potential causal genetic variants that affect particular phenotypes. This approach is known as Genetical Genomics or Integrative Genomics [7]. The identified regions are called expression Quantitative Trait Loci (eQTL). In other words, an eQTL is a region in a particular locus that influences or controls the differences of expressions of causal genes [8-11]. The expression profile is an intermediate biological space

between the phenotype and the genome. Therefore, eQTL analysis can identify interesting genetic variants even with a low sample size [12,13]. The identification of genomic regions influencing the expressions of the candidate genes could give a better perspective to use the information in animal selection as well as provide a better explanation about the way genomic regions control traits of interest.

A few studies have been conducted to identify genomic regions determining FE traits in beef cattle, chicken and other livestock species [14-17]. However, no eQTL mapping has hitherto been performed on FE related genes in dairy cows.

In this study, we performed an eQTL mapping analysis on candidate genes for the RFI trait. The hypothesis of the integrative genomics analysis is that SNPs associated with the expression of candidate genes are involved or in linkage with genomic regions regulating their expression. Therefore, the objective of this study was to identify eQTL regions together with their functional annotations associated with the RFI trait in two breeds of dairy cattle (Danish Holstein and Danish Jersey) fed two different diets and to present an eQTL mapping of candidate genes for RFI using matrix eQTL analysis, as well as characterize the SNPs by comparing our findings with previously annotated QTLs. The eQTL identified in this study could be important candidate genetic markers defining actual FE in dairy cattle, and our study suggests that there are differential traits relating to RFI in Danish Holsteins as compared to Jerseys.

Materials and Methods

Experimental animals, RFI characteristics and experimental design

The present study is based on biological samples obtained from nine Holstein and ten Jersey cows, housed at the Danish Cattle Research Centre (DCRC), Aarhus University, Denmark. The animals were part of a large

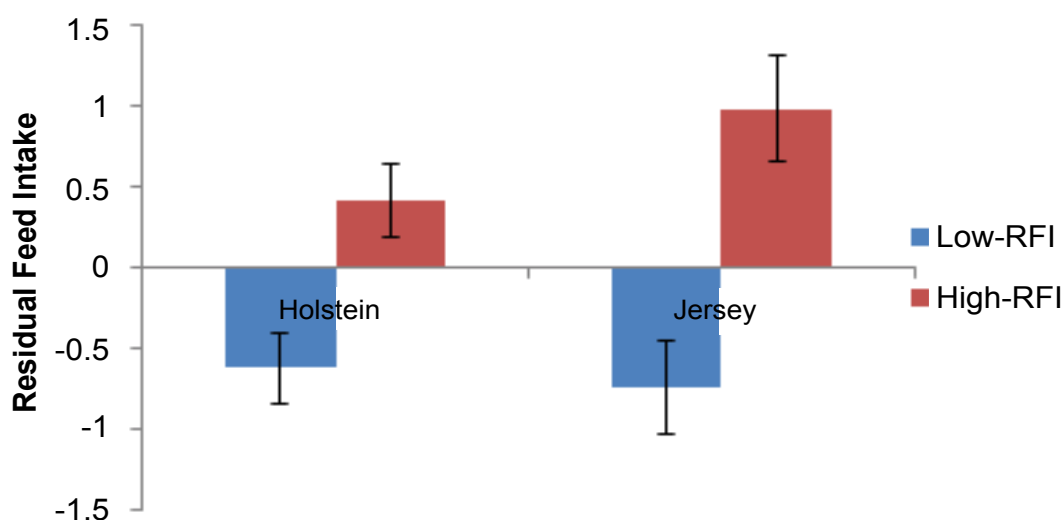


Figure 1: Mean \pm SE of Residual Feed Intake (RFI) value for the Holstein and Jersey cows used in the present experiment [6].

er experiment, where FE was determined in 200 dairy cows distributed on the two breeds [18]. The details about the animal's background and the overall experimental design of the larger trial can be found in Salleh, et al. and Li, et al. [6,18].

The experimental cows used in the present study were selected based on individually recorded RFI of cows from the larger study. A total of four Holstein cows with very high and five with very low RFI, and five Jersey cows with very high and five with very low RFI were selected, and their deviation from the average recorded RFI is shown in Figure 1. The experimental cows underwent two periods of feeding trials low concentrate (control (C)) and high concentrate (HC) diet. The two dietary exposures were separated by a conditioning period of 14 - 26 days. The details of the ration composition for both diet can be found in Salleh, et al. [6].

Biological samples

Liver biopsies (approximately 20 mg) were collected from each cow at the end of each feeding trial, RNA was extracted and sequenced. The details of the samples collection and processing were described in Salleh, et al. [6].

Blood samples were used for the DNA genotyping procedure. Ten milliliters blood samples were collected from the 19 cows using Ethylenediaminetetraacetic acid (EDTA) coated blood tubes. The blood samples were stored at -20 °C pending genomic DNA isolation and genotyping. The DNA was isolated and genotyped by Neogen GeneSeek® (Lincoln, NE, USA) using 777k BovineHD BeadChip (Illumina, Inc., San Diego, CA, USA).

Gene expressions data, genotype data and data pre-processing

Briefly, the RNA-seq data were pre-processed and processed to find candidate gene through differential expression analysis and weighted gene co-expression network analysis (WGCNA). RNA-seq analysis was performed as previously described in Salleh, et al. [6]. Briefly, the RNA raw reads were pre-processed using FastQC version 0.11.3 [19]. The reads were aligned to the Bo-

vine reference genome release 82 using STAR aligner [20]. After the alignment, quality control of the mapped reads was done using Qualimap version 2.0 [21]. Then the HTSeq-count tool was used to compute the gene expression counts [22]. The DEGs analyses were done using DESeq2 package [23] and weighted gene co-expression analyses using WGCNA package [24]. Hub genes were selected from the top significant modules that have significant association with RFI trait and having more than 80% module membership. The RNA-seq data for the present study is available in

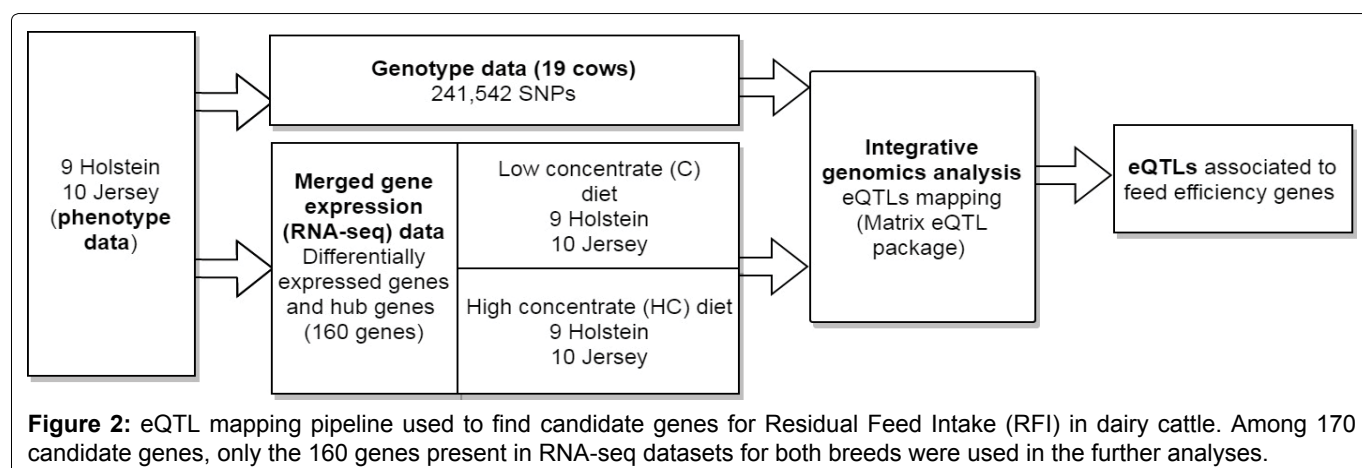
<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE92398>.

Subsequently, the two gene expression datasets of Holsteins and Jerseys were preprocessed independently to filter low counts genes in each breed. Next, the two datasets were merged by keeping only genes present in both RNA-seq datasets.

We performed two separate analyses as replicates, one using the expression profile of the cows on C diet and another one using the expression profile of the cows on the HC diet. A summary of the eQTL mapping pipeline is presented in Figure 2. The eQTL mapping analysis was performed on 170 candidate genes for RFI (Supplementary Table 1, Supplementary Table 2, Supplementary Table 3 and Supplementary Table 4) that were identified in two previous studies based on the same RNA-seq data. The total list of candidate genes included 89 DEGs between cows with high and low RFI [6] and 81 hub genes in groups of co-expressed genes associated with RFI identified by using a weighted gene co-expression network analysis (WGCNA) (unpublished).

Among the 170 candidate genes, 160 survived after the filtering step in both datasets and were used in the rest of the analyses. The numbers of candidate genes that survived after filtration were the same in the two separate analyses, which were performed for cows fed low as compared to high concentrate diets. The log2 transformation of the gene count matrix was used in the eQTL mapping.

The genotype data was filtered by Hardy Weinberg



Equilibrium ($HWE < 0.0001$), Minor Allele Frequency ($MAF < 0.15$), and missing genotype rates ($mind > 0.1$). The genotype data were also pruned to remove SNPs in strong linkage disequilibrium [25]. The preprocessing was performed using PLINK 1.90 beta software [26]. A total of 536,420 SNPs was removed after the filtration procedure. The remaining 241,542 SNPs were used for the rest of the analysis.

Integrative genomics analysis (eQTL mapping)

The theoretical aspects of eQTL mapping and applications of findings in animal sciences are well described in the literature, including our previous studies [9-11]. The eQTL mapping was performed by fitting an analysis of variance (ANOVA) model to test both additive and dominant effects. The Matrix eQTL v2.1.1 [27] package in R software was used to identify the local and distant-eQTL associated to the RFI trait. We included the breed and the lactation number as covariates in the model.

The local-and distant-eQTLs analyses were performed separately. The analysis of local eQTL was performed on SNPs that were located at less than 1 Mb distance from the start or end position of the gene of interest, while distant-eQTLs analysis was performed on SNPs located at a distance of more than 1 Mb on the same chromosome and on SNPs in other chromosomes. The SNPs were mapped onto the *Bos taurus* genome UMD 3.1. The information about gene locations were retrieved from Ensembl database for *Bos taurus* v82. P-values were adjusted using the false discovery rate (FDR) procedure for multiple comparison corrections [28]. SNPs were considered significant with FDR lower than 0.05.

Comparison of the eQTL with the Animal Genome cattle QTLdb

The significant eQTLs identified in this study were further compared to the Animal Genome cattle QTLdb

database [29]. From the cattle QTL database we filtered out long QTL regions and more than one flank markers. In total, 94,322 SNPs were used in the comparison. The SNPs information was obtained from 337 studies, 63 breeds, 366 traits of 6 trait types. The flanking regions of 500 kb around the eQTL identified in our studies were compared against the cattle QTLdb. The QTLs overlapping for at least one nucleotide were considered as a match.

Results and Discussion

The eQTL mapping analysis allows identification of SNPs associated with the expression level of specific genes. The hypothesis of this analysis is that the eQTL or eSNPs are in linkage with regulatory regions or region that encode for transcription factors responsible for the control of the expression of the targeted gene [10]. In the present study, we have analyzed candidate genes associated with the RFI trait in dairy cattle. Despite the small sample size, we identified several loci significantly related to the expression of the candidate genes. In addition, since the study focused on genes significantly associated with RFI, and the eQTL analysis was done on animals with widely different RFI, either very low or very high, our study had enough power to detect biologically meaningful expression variants.

Different strategies for obtaining significant eQTLs associated to RFI candidate genes by using C versus HC dataset

In the expression profile of cows fed the control diet, we identified 20 local-eQTLs SNPs or cis-eQTLs SNP ($FDR < 0.05$) associated with the expression of seven genes (*BDH2*, *CHRNE*, *ELOVL6*, *GIMAP4*, *FDXR*, *CXCL9* and *CD52*) (Table 1). However, there was no significant distant-eQTL (trans-eQTLs) associated with the candidate genes in the analysis performed among cows fed the control diet. On the other hand, among cows fed

Table 1: Top significant local-eQTLs targeting candidate genes for cows fed the control (C) diet.

rsID	Gene.name	Gene type	FDR	Position	Freq	Genotype	Gene position
rs133674837	<i>BDH2</i>	DE	7.35E-04	6:23051485	6/8/5	CC/CA/AA	6:23047057-23077431
rs109975461	<i>CHRNE</i>	DE	4.23E-02	19:26981374	8/6/5	AA/AG/GG	19:27118517-27123114
rs109947248	<i>CHRNE</i>	DE	4.23E-02	19:26994134	8/6/5	AA/AG/GG	19:27118517-27123114
rs109341116	<i>CHRNE</i>	DE	4.23E-02	19:27098154	8/6/5	AA/AG/GG	19:27118517-27123114
rs110896981	<i>CHRNE</i>	DE	4.23E-02	19:27192150	10/4/5	AA/AG/GG	19:27118517-27123114
rs43318602	<i>ELOVL6</i>	DE	4.23E-02	6:16656338	15/2/2	GG/GA/AA	6:16376642-16510240
rs43318545	<i>ELOVL6</i>	DE	4.23E-02	6:16678359	15/2/2	GG/GA/AA	6:16376642-16510240
rs43317462	<i>ELOVL6</i>	DE	4.23E-02	6:16731878	9/8/2	GG/GA/AA	6:16376642-16510240
rs110036492	<i>ELOVL6</i>	DE	4.23E-02	6:16738741	8/8/2	GG/GA/AA	6:16376642-16510240
rs43315610	<i>ELOVL6</i>	DE	4.23E-02	6:16755625	6/11/2	GG/GA/AA	6:16376642-16510240
rs43316358	<i>ELOVL6</i>	DE	4.23E-02	6:16761983	6/11/2	GG/GA/AA	6:16376642-16510240
rs43317449	<i>ELOVL6</i>	DE	4.23E-02	6:16725243	9/8/2	GG/GA/AA	6:16376642-16510240
rs109963253	<i>GIMAP4</i>	DE	3.02E-02	4:113638587	8/9/2	GG/GA/AA	4:113866800-113874303
rs134589272	<i>FDXR</i>	DE	4.23E-02	19: 56624163	10/8/1	GG/GA/AA	19: 57164031-57175524

rs ID = reference SNP cluster ID; Gene Name = name of the targeted genes; Gene type = type of candidate gene: Hub genes (HG) or differentially expressed genes (DE); FDR = False Discovery Rate of the association between SNP and gene expression, Position = genome position of the SNP, Freq = frequency of each genotypes (A1A1/A1A2/A2A2) in the cows analysed, Genotype = genotype at the SNP locus (A1A1/A1A2/A2A2), Gene position = genome position of the targeted gene.

Table 2: Top significant local-eQTLs targeting candidate genes for cows fed the high concentrate (HC) diet.

rsID	Gene.name	Gene type	FDR	Position	Freq	Genotype	Gene position
rs135948495	<i>UHRF1</i>	DE	6.72E-03	7:20322296	5/9/5	CC/CA/AA	7:20436670-20469912
rs134849198	<i>UHRF1</i>	DE	6.72E-03	7:20327318	5/9/5	CC/CA/AA	7:20436670- 20469912
rs137012774	<i>UHRF1</i>	DE	6.72E-03	7:20336175	5/9/5	CC/CA/AA	7:20436670- 20469912
rs133674837	<i>BDH2</i>	DE	4.81E-03	6:23051485	6/8/5	CC/CA/AA	6:23047057- 23077431
rs109739833	<i>HSD17B4</i>	DE	1.14E-02	7:35653128	7/10/2	AA/AG/GG	7:35662599- 35763653
rs110212970	<i>GIMAP4</i>	DE	7.88E-04	4:113608223	8/9/2	GG/GA/AA	4:113866800- 113874303
rs109963253	<i>GIMAP4</i>	DE	7.88E-04	4:113638587	8/9/2	GG/GA/AA	4:113866800- 113874303

rs ID = reference SNP cluster ID; Gene Name = name of the targeted genes; Gene type = type of candidate gene: Hub genes (HG) or differentially expressed genes (DE); FDR = False Discovery Rate of the association between SNP and gene expression, Position = genome position of the SNP, Freq = frequency of each genotypes (A1A1/A1A2/A2A2) in the cows analysed, Genotype = genotype at the SNP locus (A1A1/A1A2/A2A2), Gene position = genome position of the targeted gene.

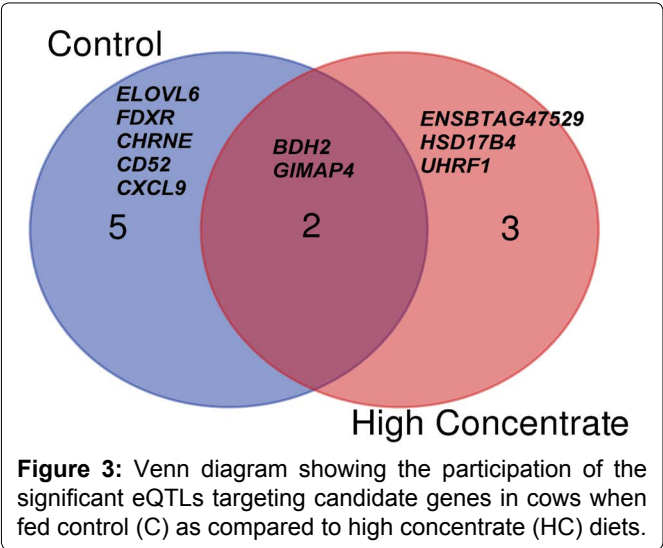


Figure 3: Venn diagram showing the participation of the significant eQTLs targeting candidate genes in cows when fed control (C) as compared to high concentrate (HC) diets.

the high concentrate diet, we identified 16 local eQTLs SNPs (FDR < 0.05) associated with the expression of five genes (*UHRF1*, *BDH2*, *HSD17B4*, *GIMAP4* and *ENSBTAG00000047529*) (Table 2) and 2891 distant-eQTLs associated with the expression of 45 genes. Among the local-eQTL, genes that were in common in both diet groups were the *BDH2* and *GIMAP4* genes. Figure 3 shows the significant eQTLs targeting candidate genes. A complete list of the significant distant-eQTLs, including the chromosomal position and annotation of the SNPs in the HC dataset analysis is presented in Supplementary Table 5.

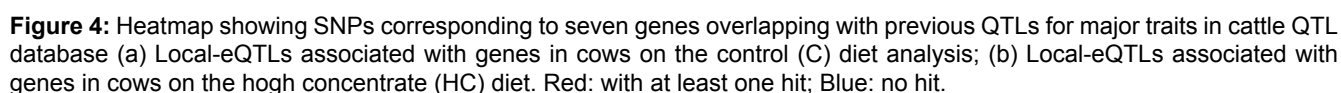
In Holstein, cows fed with C diet dataset, we detected eQTLs associated to the three genes *BDH2*, *CHRNE* and *ELOVL6*, whereas for cows fed the HC diet dataset, the eQTLs associated to two other genes, *UHRF1* and *HSD17B4*. In Jersey cows, two DEGs (*GIMAP4* and *FDXR*) and two HG's (*CXCL9* and *CD52*) belonging to a group of co-expressed genes associated with RFI, when they were fed the C diet, whereas only the *GIMAP4* gene was detected as significant local-eQTL, when they were fed the HC diet. However, the HG's (*CXCL9* and *CD52*) were not further analyzed, since only one animal with rare allele at these loci were present in the dataset. Supplementary Figure 1 and Supplementary Figure 2 present boxplots of genotypes and their correlation with gene expressions for the top seven significant local-eQTLs

corresponding to seven candidate genes (*BDH2*, *CHRNE*, *ELOVL6*, *GIMAP4*, *FDXR*, *UHRF1* and *HSD17B4*) for the RFI trait in the present study.

Previous studies showed that identification of eQTLs and genomic regions would give additional information towards the identification of causal variants [30]. Hence, the eQTLs that were identified as associated to the RFI trait in the present study would provide additional information for the development of biomarkers.

The first top eQTL with a significant relationship between the gene expression and genotype is rs133674837, which is associated to the *BDH2* gene (Supplementary Figure 1a), and as mentioned the association was found to be significant in the two separate analyses for cows when fed the C diet as well as when fed the HC diet. The expression of the *BDH2* gene was previously identified to be upregulated in high FE cows [6]. All low RFI (high FE) Holstein cows (n = 5) had homozygous (AA) genotype at this locus, while 80% of the high RFI (low FE) Holstein cow had heterozygous genotype (CA). *BDH2* encodes for the enzyme 3-Hydroxybutyrate Dehydrogenase 2, which is responsible for degradation of 3-hydroxybutyrate-a ketone body derived partly from rumen fermentation and partly from incomplete oxidation of fatty acids in the liver [31,32]. The *BDH2* gene in the liver has been observed to be downregulated in animals, when ketogenesis occurred (mice and pigs) [31,33]. This happens, for example, during feed restriction or fasting of animals, and mRNA expression of *BDH2* gene has been shown to be lower in such animals compared to normal feeding animals [33]. In the present study, the hepatic *BDH2* gene expression was downregulated in high RFI (low FE) animals. The positive association between a homozygous (AA) genotype and upregulation of the *BDH2* gene in low RFI (high FE) Holstein cows shows that this locus might influence the RFI trait. However, in Jersey cows, 80% (n = 4) of low RFI (high FE) cows were homozygous (CC) at this allele. Hence, specifically for Holstein cows, a homozygous (AA) genotype is expected to favor low RFI and hence high FE.

GIMAP4 gene is another gene that has been detected as significantly associated with the eQTLs listed in Table 1 and Table 2 in both analysis (i.e. when cows were fed



[35–38]. A possible explanation is that animals with poorer immune function are more prone to develop infectious diseases like e.g. mastitis, and this can reduce milk production, induce fever-associated increases in metabolism, and hence increased energy expenditures per kg of produced milk, which subsequently reduces FE [39].

rs109975461, which is associated with the *CHRNA* gene, was also a significant eQTL. At this locus, all high RFI (low FE) Holstein cows had a homozygous (GG) genotype, whereas 80% of the low RFI (high FE) cows (n = 4) had a heterozygous (AG) genotype (Supplementary Figure 1c). In other words, high feed efficient cows that had a high expression of the *CHRNA* gene also had the heterozygote genotype. However, in the Jersey group, there was no association to be seen for this *CHRNA* gene. The *CHRNA* gene encodes for the acetylcholine recep-

tors in mature mammalian neuromuscular junctions. In general, this gene was never discussed before in relation with FE traits. Acetylcholine has been reported to influence hepatocyte glucose metabolism in rodents via actions on muscarinic receptors [40], but whether this is also the case in ruminants is not clear. Perhaps, more importantly, acetylcholine plays a critical role in the complex regulation of hypothalamic neuronal activity that influences feed intake [41], and in dairy cows, feed intake is a major factor limiting milk production in high-yielding dairy cows in early lactation [42].

Another interesting candidate identified in the analyses of Holstein cows on the C diet was the *ELOVL6* gene. In our study, the top SNP targeting *ELOVL6* gene was rs43315610. The *ELOVL6* gene has previously been discovered as an important gene that influences FE in beef cattle and pigs [43,44]. *ELOVL6*, which is also known as elongation of very long chain fatty acids protein 6, is part of the pathway of de novo fatty acid synthesis [45]. The lower expression of this gene in low RFI Holstein cows might be associated with low rates of de novo synthesis of fatty acids, as it has previously been described in pigs [44,46], and de novo synthesized fatty acids constitute up to 50% of fatty acids in milk on a molecular weight basis [47]. This gene has also been associated to long chain fatty acid synthesis in beef cattle [48]. Previously, the expression of *ELOVL6* was found differentially expressed in liver, adipose tissue and muscle [48]. In another study on QTL mapping for RFI in Holstein calves, it was found that another gene involved in fatty acid metabolism, *FABP4* gene were significantly associated with the top SNPs significantly associated RFI across three stages of age [49]. This gene is encoded for fatty acid binding protein which suggests that fatty acid synthesis and metabolism may be important parts of the RFI trait. In Jersey cattle, we did not observe any relation between RFI genotype and the *ELOVL6* gene expression. We found that 80% ($n = 4$) of the low RFI (high FE) Holstein cows had a heterozygous (GA) genotype, while 20% ($n = 1$) were homozygous (GG) (Supplementary Figure 1d). Therefore, a heterozygous genotype is expected to favor high FE.

When cows were fed the C diet, rs134589272 was identified as an eQTL, which corresponded to the *FDXR* gene in Jersey cows. All ($n = 5$) low RFI (high FE) Jersey cows were heterozygous (GA) and had high expression of this gene, while 80% ($n = 4$) of the high RFI (low FE) Jersey cows were homozygous (GG) corresponding to a lower expression of the *FDXR* gene (Supplementary Figure 1e). For Holstein cows, RFI was not related neither to this eQTL nor to the genotype for the *FDXR* gene. Functions of the *FDXR* gene are related to cholesterol metabolism [6], which is an important feature of e.g. membrane synthesis, which is important for formation of the milk fat globule membrane covering secreted milk fat.

In addition, another four eQTLs associated to the gene *HSD17B4* and *UHRF1* expression were found as significant in the analysis for cows fed the HC diet (Supplementary Figure 2a and Supplementary Figure 2b). Interestingly, these two genes were also previously found associated with the FE trait. The *HSD17B4* gene encodes for a major enzyme involved in peroxisomal β -oxidation, and it was found to be upregulated in abdominal fat of low growth chicken [50], and this appears to be in line with the present study, where the *HSD17B4* gene expression was upregulated in the high RFI Holstein cows. *UHRF1* encodes for Ubiquitin like With PHD and Ring Finger Domains 1 (<http://www.genecards.org/cgi-bin/carddisp.pl?gene=UHRF1>), which is an essential regulator of DNA methylation. Several studies have identified that Ubiquitin family genes were significantly associated with RFI traits in *Bos taurus* [51,52]. In the present study *UHRF1* gene was found significantly downregulated in high RFI Holstein cows.

Overlapping genomic regions for FE trait in the QTL database

In order to gain more information regarding the eQTLs that we discovered in the present study, we compared the results of the SNPs locations with the previously reported QTLs and variants from GWAS study from the Animal genome cattle QTL database. We identified several overlaps of our eQTLs with QTLs from previous studies. The QTLs overlapping with our eQTLs were associated with a different type of traits (Figure 4a and Figure 4b).

The eQTLs which associated to the expression of *ELOVL6* and *FDXR* genes are the most overlapped with many traits. Only the *GIMAP4* gene was previously associated to production traits, such as RFI, rump width, metabolic body weight, body weight gain, body weight (yearling), body weight, body depth, average daily gain as well as average daily feed intake [53,54]. However, the same region contains QTLs for other traits, such as reproduction, milk, meat and carcass, health and exterior association traits [55,56].

The fact that all these associations with different type of traits were found within this 1Mb region, shows that this must be a significant region with control points for several targeting genes. The eQTLs identified are close to the QTL for production traits and for FE traits. At the same time, this confirms that the candidate genes which associated to FE trait in our findings were also closely associated to several production traits. However, the association with other important traits can be a sign of double association between reproduction and production traits, which were well discussed elsewhere [57]. Thus, the uses of genomic region information need to be tested and validated in a different and a larger population before further usage in any genomic selection procedures can be implemented.

Conclusion

To bridge the gap between genotype and phenotype, we attempted in this study to identify DEGs and HG's among previously identified candidate genes for the FE trait. The identified local-eQTLs provide additional evidence of the involvement of some of previously identified candidate genes in RFI determination, and our study provides new information on possible regulatory and causative genetic variants that can be used in genomics-based selection for FE in dairy cows. We identified eQTL associated to the expression of seven genes (*BDH2*, *CHRNE*, *ELOVL6*, *GIMAP4*, *UHRF1*, *HSD17B4* and *FDXR*) that appear to be involved in metabolic pathways related to RFI and hence feed efficiency. The eQTLs overlapped with QTLs previously associated with FE trait (e.g. dry matter intake, longevity, body weight gain and net merit). Interestingly, Holstein and Jersey cows appear to rely on different strategies to achieve low RFI, and this was associated to cholesterol and lipid metabolism related pathways in Holstein cows, but to immune and inflammatory related functions in Jersey cows. Thus, our findings suggest that the identified eQTLs can be used as potential biomarkers for feed efficiency and used to predict feed efficiency level. The genomic region around the identified SNP markers could be included in genomics/genetic-based selection in Holstein and Jersey. However, before applying this new knowledge in genetic testing or in commercial applications, the results must be validated in a larger population, and it must be further analyzed if pleiotropic effects of eQTLs also include adverse disease traits.

Acknowledgements

The authors would like to thank the collaborator scientists, specifically Peter Lund, Johanna Höglund and Dana Olijhoek, for contributions to the present study. This project was funded by Milk Levy Foundation, Skejby Denmark. The PhD student (Suraya Mohamad Salleh) would also like to acknowledge the Ministry of Higher Education of Malaysia and Universiti Putra Malaysia for the PhD stipend as well as University of Copenhagen for the tuition fee waiver.

References

- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: A revolutionary tool for transcriptomics. *Nat Rev Genet* 10: 57-63.
- Weber KL, Welly BT, Van Eenennaam AL, Young AE, Porto-Neto LR, et al. (2016) Identification of gene networks for residual feed intake in angus cattle using genomic prediction and RNA-seq. *PLoS One* 11: e0152274.
- Alexandre PA, Kogelman LJ, Santana MH, Passarelli D, Pulz LH, et al. (2015) Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. *BMC Genomics* 16: 1.
- Paradis F, Yue S, Grant J, Stothard P, Basarab J, et al. (2015) Transcriptomic analysis by RNA sequencing reveals that hepatic interferon-induced genes may be associated with feed efficiency in beef heifers. *J Anim Sci* 93: 3331-3341.
- Koch RM, Swiger LA, Chambers D, Gregory KE (1963) Efficiency of Feed Use in Beef Cattle. *Journal of Animal Science* 22.
- Salleh M, Mazzoni G, Höglund J, Olijhoek D, Lund P, et al. (2017) RNA-Seq transcriptomics and pathway analyses reveal potential regulatory genes and molecular mechanisms in high-and low-residual feed intake in Nordic dairy cattle. *BMC Genomics* 18: 258.
- Le Mignon G, Désert C, Pitel F, Leroux S, Demeure O, et al. (2009) Using transcriptome profiling to characterize QTL regions on chicken chromosome 5. *BMC Genomics* 10: 575.
- Civelek M, Lusis AJ (2014) Systems genetics approaches to understand complex traits. *Nat Rev Genet* 15: 34.
- Kadarmideen HN, Reverter A (2007) Combined genetic, genomic and transcriptomic methods in the analysis of animal traits. *CABI review: Perspectives in Agriculture, Veterinary science, Nutrition and Natural resources* 2: 16.
- Kadarmideen HN, von Rohr P, Janss LL (2006) From genetical genomics to systems genetics: Potential applications in quantitative genomics and animal breeding. *Mamm Genome* 17: 548-564.
- Suravajhala P, Kogelman LJ, Kadarmideen HN (2016) Multi-omic data integration and analysis using systems genomics approaches: Methods and applications in animal production, health and welfare. *Genet Sel Evol* 48: 38.
- Kadarmideen H (2008) Genetical systems biology in livestock: Application to gonadotrophin releasing hormone and reproduction. *IET Syst Biol* 2: 423-441.
- Mazzoni G, Pedersen H, De Oliveira Junior G, Alexandre P, Razza E, et al. (2017) Application of integrative genomics and systems biology to conventional and in vitro reproductive traits in cattle.14.
- Yuan J, Wang K, Yi G, Ma M, Dou T, et al. (2015) Genome-wide association studies for feed intake and efficiency in two laying periods of chickens. *Genetics Selection Evolution* 47: 82.
- Seabury CM, Oldeschulte DL, Saatchi M, Beever JE, Decker JE, et al. (2017) Genome-wide association study for feed efficiency and growth traits in US beef cattle. *BMC genomics* 18: 386.
- Rolf M, Taylor J, Schnabel R, McKay S, McClure M, et al. (2012) Genome-wide association analysis for feed efficiency in Angus cattle. *Anim Genet* 43: 367-374.
- De Koning D, Cabrera C, Haley C (2007) Genetical genomics: Combining gene expression with marker genotypes in poultry. *Poult Sci* 86: 1501-1509.
- Li B, Fikse W, Lassen J, Lidauer M, Løvendahl P, et al. (2016) Genetic parameters for dry matter intake in primiparous Holstein, Nordic Red, and Jersey cows in the first half of lactation. *J Dairy Sci* 99: 7232-7239.
- Andrews S (2010) FastQC: A quality control tool for high throughput sequence data. In.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, et al. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15-21.
- Okonechnikov K, Conesa A, García-Alcalde F (2016) Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* 32: 292-294.
- Anders S, Pyl PT, Huber W (2014) HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166-169.

23. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550.
24. Langfelder P, Horvath S (2008) WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559.
25. Dudbridge F, Newcombe PJ (2015) Accuracy of gene scores when pruning markers by linkage disequilibrium. *Hum Hered* 80: 178-186.
26. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, et al. (2015) Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* 4: 7.
27. Shabalin AA (2012) Matrix eQTL: Ultra-fast eQTL analysis via large matrix operations. *Bioinformatics* 28: 1353-1358.
28. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B (Methodological)* 289-300.
29. Hu ZL, Park CA, Reecy JM (2015) Developmental progress and current status of the Animal QTLdb. *Nucleic Acids Res* 44: 827-833.
30. Kogelman LJ, Cirera S, Zhernakova DV, Fredholm M, Franke L, et al. (2014) Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA Sequencing in a porcine model. *BMC Med Genomics* 7.
31. Geisler C, Hepler C, Higgins M, Renquist B (2016) Hepatic adaptations to maintain metabolic homeostasis in response to fasting and refeeding in mice. *Nutr Metab* 13: 62.
32. Khanal P, Axel AM, Kongsted AH, Husted SV, Johnsen L, et al. (2015) Late gestation under- and overnutrition have differential impacts when combined with a post-natal obesogenic diet on glucose-lactate-insulin adaptations during metabolic challenges in adolescent sheep. *Acta Physiol* 213: 519-536.
33. Grala TM, Kay JK, Phyn CV, Bionaz M, Walker CG, et al. (2013) Reducing milking frequency during nutrient restriction has no effect on the hepatic transcriptome of lactating dairy cattle. *Physiol Genomics* 2013.
34. Heinonen M (2015) Players of Th-cell differentiation and immune dysregulation—focus on GIMAP family genes.
35. Banos G, Wall E, Coffey MP, Bagnall A, Gillespie S, et al. (2013) Identification of immune traits correlated with dairy cow health, reproduction and productivity. *PLoS One* 8: e65766.
36. Patience JF, Rossoni-Serão MC, Gutiérrez NA (2015) A review of feed efficiency in swine: Biology and application. *J Anim Sci Biotechnol* 6: 33.
37. Gondret F, Vincent A, Houée-Bigot M, Siegel A, Lagarrigue S, et al. (2017) A transcriptome multi-tissue analysis identifies biological pathways and genes associated with variations in feed efficiency of growing pigs. *BMC Genomics* 18: 244.
38. Mpetile Z (2014) Effect of divergent selection for residual feed intake on immune system of Yorkshire pigs.
39. Johnson R (2012) Fueling the immune response: What's the cost? In: *Feed efficiency in swine*. Springer 211-223.
40. Vatamaniuk M, Horyn O, Vatamaniuk O, Doliba N (2003) Acetylcholine affects rat liver metabolism via type 3 muscarinic receptors in hepatocytes. *Life Sci* 72: 1871-1882.
41. Picciotto MR, Higley MJ, Mineur YS (2012) Acetylcholine as a neuromodulator: Cholinergic signaling shapes nervous system function and behavior. *Neuron* 76: 116-129.
42. Johansen M, Lund P, Weisbjerg MR (2018) Feed intake and milk production in dairy cows fed different grass and legume species: A meta-analysis. *Animal* 12: 66-75.
43. Lemos MV, Chiaia HLJ, Berton MP, Feitosa FL, Aboujaoud C, et al. (2016) Genome-wide association between single nucleotide polymorphisms with beef fatty acid profile in Nelore cattle using the single step procedure. *BMC Genomics* 17: 213.
44. Reyer H, Oster M, Magowan E, Dannenberger D, Ponsuksili S, et al. (2017) Strategies towards improved feed efficiency in pigs comprise molecular shifts in hepatic lipid and carbohydrate metabolism. *Int J Mol Sci* 18: 1674.
45. Moon YA, Ochoa CR, Mitsche MA, Hammer RE, Horton JD (2014) Deletion of ELOVL6 blocks the synthesis of oleic acid but does not prevent the development of fatty liver or insulin resistance. *J Lipid Res* 55: 2597-2605.
46. Corominas J, Marchesi JA, Puig-Oliveras A, Revilla M, Estellé J, et al. (2015) Epigenetic regulation of the ELOVL6 gene is associated with a major QTL effect on fatty acid composition in pigs. *Genet Sel Evol* 47: 20.
47. Breckenridge W, Kuksis A (1967) Molecular weight distributions of milk fat triglycerides from seven species. *J Lipid Res* 8: 473-478.
48. Chen S, He H, Liu X (2017) Tissue expression profiles and transcriptional regulation of elongase of very long chain fatty acid 6 in bovine mammary epithelial cells. *PLoS One* 12: e0175777.
49. Cohen-Zinder M, Asher A, Lipkin E, Feingersch R, Agmon R, et al. (2016) FABP4 is a leading candidate gene associated with residual feed intake in growing Holstein calves. *Physiol Genomics* 48: 367-376.
50. Resnyk CW, Carré W, Wang X, Porter TE, Simon J, et al. (2017) Transcriptional analysis of abdominal fat in chickens divergently selected on bodyweight at two ages reveals novel mechanisms controlling adiposity: Validating visceral adipose tissue as a dynamic endocrine and metabolic organ. *BMC Genomics* 18: 626.
51. Karisa B, Moore S, Plastow G (2014) Analysis of biological networks and biological pathways associated with residual feed intake in beef cattle. *Anim Sci J* 85: 374-387.
52. de Oliveira PS, Cesar AS, do Nascimento ML, Chaves AS, Tizioto PC, et al. (2014) Identification of genomic regions associated with feed efficiency in Nelore cattle. *BMC Genet* 15: 100.
53. Lu D, Miller S, Sargolzaei M, Kelly M, Vander Voort G, et al. (2013) Genome-wide association analyses for growth and feed efficiency traits in beef cattle. *J Anim Sci* 91: 3612-3633.
54. Snelling W, Allan M, Keele J, Kuehn L, Mcdanel T, et al. (2010) Genome-wide association study of growth in cross-bred beef cattle 1 2. *J Anim Sci* 88: 837-848.
55. Cole JB, Wiggans GR, Ma L, Sonstegard TS, Lawlor TJ, et al. (2011) Genome-wide association analysis of thirty-one production, health, reproduction and body conformation traits in contemporary US Holstein cows. *BMC Genomics* 12: 408.
56. Lindholm-Perry AK, Kuehn LA, Rempel LA, Smith TP, Cushman RA, et al. (2012) Evaluation of bovine chemerin (rarres2) gene variation on beef cattle production traits1. *Front Genet* 3: 39.
57. Dobson H, Smith R, Royal M, Knight C, Sheldon I (2007) The high-producing dairy cow and its reproductive performance. *Reprod Domest Anim* 42: 17-23.

Supplementary Table 1: List of differentially expressed genes for Holstein.

	Ensembl.Gene.ID	Associated.Gene.Name	baseMean	log2FoldChange	padj
1	ENSBTAG00000000170	Uncharacterized protein	226.9268	-0.42496	0.013405
2	ENSBTAG00000000654	<i>ARMC4</i>	58.66466	-0.59056	0.000358
3	ENSBTAG00000001009	<i>HCLS1</i>	440.5817	0.322596	0.028031
4	ENSBTAG00000001154	<i>DGAT2</i>	511.454	-0.36983	0.0292
5	ENSBTAG00000001204	<i>KIAA1462</i>	225.2099	-0.41583	0.017639
6	ENSBTAG00000002224	<i>UHRF1</i>	77.58827	-0.49601	0.000209
7	ENSBTAG00000002526	<i>BDH2</i>	1382.085	-0.57884	3.67E-16
8	ENSBTAG00000002705	<i>REC8</i>	304.6793	-0.36531	0.00171
9	ENSBTAG00000003696	<i>CCDC64</i>	45.09122	0.445439	0.026992
10	ENSBTAG00000003718	<i>HACL1</i>	6329.552	0.315377	0.038742
11	ENSBTAG00000004076	<i>OXER1</i>	223.6781	-0.43485	0.010458
12	ENSBTAG00000004558	<i>C15orf48</i>	89.05707	0.504892	0.003366
13	ENSBTAG00000004908	<i>CHRNE</i>	246.4988	-0.74877	6.38E-08
14	ENSBTAG00000005287	<i>CYP7A1</i>	4126.209	0.456229	0.013057
15	ENSBTAG00000005629	<i>AIM1L</i>	913.5876	-0.29846	0.003366
16	ENSBTAG00000006452	<i>CD3D</i>	77.7469	0.40702	0.036637
17	ENSBTAG00000006599	<i>INHBE</i>	605.9473	-0.4237	0.044088
18	ENSBTAG00000006675	<i>PCSK6</i>	3039.353	-0.1866	0.028031
19	ENSBTAG00000006934	<i>CYP11A1</i>	649.8806	0.48638	0.004844
20	ENSBTAG00000006978	<i>HSD17B4</i>	13797.37	0.301172	0.026992
21	ENSBTAG00000006999	<i>RYR1</i>	148.3439	0.516872	0.001988
22	ENSBTAG00000007554	<i>IFI6</i>	136.8134	0.375726	0.049041
23	ENSBTAG00000007828	<i>SLA</i>	118.2561	0.322695	0.038742
24	ENSBTAG00000007895	<i>SLC20A1</i>	880.0804	-0.56325	1.36E-05
25	ENSBTAG00000008160	<i>MBOAT2</i>	440.5952	0.344004	0.043864
26	ENSBTAG00000008424	<i>ABR</i>	459.6035	0.329325	0.031825
27	ENSBTAG00000008913	<i>TMEM98</i>	333.1394	-0.51804	0.001004
28	ENSBTAG00000009085	<i>SLC35A5</i>	1691.358	0.277516	0.006652
29	ENSBTAG00000009137	<i>NKG7</i>	215.4444	0.380987	0.028031
30	ENSBTAG00000009263	<i>MFSD1</i>	2661.739	0.240088	0.013057
31	ENSBTAG00000010463	Uncharacterized protein	394.1386	0.383919	0.002458
32	ENSBTAG00000010564	<i>ELOVL6</i>	994.7373	0.43905	0.026992
33	ENSBTAG00000011771	<i>FICD</i>	107.9076	-0.35997	0.037153
34	ENSBTAG00000011832	<i>ALDH18A1</i>	404.9092	0.313185	0.025962
35	ENSBTAG00000012007	<i>SOCS2</i>	835.3211	0.422255	0.043864
36	ENSBTAG00000012995	<i>CCDC109B</i>	52.93337	0.424138	0.037153
37	ENSBTAG00000013596	<i>NR1H4</i>	1215.489	0.241071	0.011983
38	ENSBTAG00000014064	<i>FGFR2</i>	1554.53	-0.40189	0.001141
39	ENSBTAG00000014791	<i>CTH</i>	224.7532	-0.5345	1.74E-06
40	ENSBTAG00000015313	<i>CEACAM19</i>	51.60282	-0.94436	1.81E-14
41	ENSBTAG00000015419	<i>ARHGEF37</i>	204.9786	0.471594	0.001988
42	ENSBTAG00000016542	<i>LAMB3</i>	1783.587	0.424674	0.024742
43	ENSBTAG00000017567	<i>ACACA</i>	844.7716	0.404374	0.005072
44	ENSBTAG00000018116	<i>MTFP1</i>	88.16198	-0.38701	0.025962
45	ENSBTAG00000018548	<i>INTS7</i>	6522.273	0.238158	0.001988
46	ENSBTAG00000018604	<i>SEMA4G</i>	4847.004	-0.16063	0.049041
47	ENSBTAG00000018723	<i>SLC25A34</i>	96.51437	-0.44163	0.0292
48	ENSBTAG00000019585	<i>MYOM1</i>	962.386	0.453079	0.025089

49	ENSBTAG00000020116	<i>JSP.1</i>	2041.873	0.323612	0.016775
50	ENSBTAG00000020371	<i>ACOT8</i>	312.3978	0.417754	0.013405
51	ENSBTAG00000020375	Uncharacterized protein	5129.183	0.421767	0.031078
52	ENSBTAG00000020499	Uncharacterized protein	68.21442	0.571805	0.000358
53	ENSBTAG00000020755	<i>SELP</i>	478.8212	-0.39216	0.014977
54	ENSBTAG00000021746	<i>ANXA5</i>	333.8561	-0.38421	0.038742
55	ENSBTAG00000023851	<i>FAM102A</i>	229.0066	-0.50952	0.001422
56	ENSBTAG00000023929	<i>FOSL2</i>	189.6217	0.424473	0.028031
57	ENSBTAG00000024044	<i>CDKL4</i>	82.11652	0.51475	0.001988
58	ENSBTAG00000025258	Uncharacterized protein	102.6902	0.544344	0.00171
59	ENSBTAG00000025898	<i>TBC1D8</i>	442.302	0.271973	0.047591
60	ENSBTAG00000026779	<i>LYZ</i>	516.8438	0.643794	6.78E-06
61	ENSBTAG00000030966	<i>TAF6</i>	419.5936	-0.25649	0.010458
62	ENSBTAG00000035998	<i>CKB</i>	332.0427	0.385047	0.049041
63	ENSBTAG00000037913	Uncharacterized protein	436.5267	0.212286	0.042867
64	ENSBTAG00000037917	<i>SLC17A1</i>	2786.405	0.438054	0.017094
65	ENSBTAG00000038496	<i>CR2</i>	1355.757	-0.54159	3.72E-06
66	ENSBTAG00000038962	<i>SLC6A11</i>	2637.353	-0.37017	0.010028
67	ENSBTAG00000039731	<i>RND3</i>	1761.21	-0.25436	0.028031
68	ENSBTAG00000046076	Uncharacterized protein	124.7543	-0.41978	0.04847
69	ENSBTAG00000046730	Uncharacterized protein	139.8274	0.365363	0.049041
70	ENSBTAG00000047529	Uncharacterized protein	110.8376	-0.53501	0.001896

+ve log2 fold change = upregulated in low feed efficiency group; -ve log2 fold change = downregulated in low feed efficiency group.

Supplementary Table 2: List of hub genes for Holstein.

	Ensembl gene ID	Gene name	Module membership	Gene significance
1	ENSBTAG00000000197	<i>TRMT10A</i>	0.801	0.576
2	ENSBTAG00000001774	<i>SPRY2</i>	-0.814	-0.52
3	ENSBTAG00000001950	<i>RDH11</i>	0.852	0.441
4	ENSBTAG00000002412	<i>CYB5B</i>	0.907	0.633
5	ENSBTAG00000002435	<i>PTPRE</i>	0.852	0.767
6	ENSBTAG00000002714	<i>GNAI1</i>	0.901	0.557
7	ENSBTAG00000002827	<i>ACAT2</i>	0.946	0.691
8	ENSBTAG00000002966	<i>DNAJC13</i>	0.813	0.71
9	ENSBTAG00000003068	<i>MSMO1</i>	0.852	0.579
10	ENSBTAG00000003305	<i>NCF1</i>	0.802	0.642
11	ENSBTAG00000003696	<i>CCDC64</i>	0.837	0.679
12	ENSBTAG00000003718	<i>HACL1</i>	0.854	0.705
13	ENSBTAG00000003948		0.919	0.559
14	ENSBTAG00000004075	<i>IDI1</i>	0.87	0.607
15	ENSBTAG00000004688	<i>DHCR24</i>	0.859	0.555
16	ENSBTAG00000005183	<i>MVK</i>	0.906	0.497
17	ENSBTAG00000005498	<i>SQLE</i>	0.816	0.442
18	ENSBTAG00000005650	<i>SKAP2</i>	0.826	0.589
19	ENSBTAG00000005976	<i>HSD17B7</i>	0.809	0.55
20	ENSBTAG00000006999	<i>RYR1</i>	0.929	0.763
21	ENSBTAG00000007014	<i>CEP63</i>	0.823	0.623
22	ENSBTAG00000007079	<i>LCP1</i>	0.806	0.583
23	ENSBTAG00000007840	<i>HMGCR</i>	0.888	0.522
24	ENSBTAG00000007844	<i>CETN2</i>	0.836	0.335

25	ENSBTAG00000008160	<i>MBOAT2</i>	0.865	0.534
26	ENSBTAG00000008329	<i>CYTIP</i>	0.823	0.477
27	ENSBTAG00000010347	<i>EZR</i>	0.85	0.506
28	ENSBTAG00000011146	<i>RAB8B</i>	0.884	0.473
29	ENSBTAG00000011839	<i>HMGCS1</i>	0.871	0.507
30	ENSBTAG00000012059	<i>MVD</i>	0.831	0.364
31	ENSBTAG00000012170	<i>UBL3</i>	0.813	0.729
32	ENSBTAG00000012432	<i>FDFT1</i>	0.821	0.529
33	ENSBTAG00000012695	<i>LCK</i>	0.837	0.534
34	ENSBTAG00000013284		0.886	0.736
35	ENSBTAG00000013303	<i>ACSS2</i>	0.866	0.571
36	ENSBTAG00000013749	<i>RHOQ</i>	0.868	0.525
37	ENSBTAG00000014517	<i>KLB</i>	0.857	0.64
38	ENSBTAG00000015327	<i>SPTAN1</i>	0.899	0.637
39	ENSBTAG00000015980	<i>FASN</i>	0.859	0.49
40	ENSBTAG00000016445	<i>YME1L1</i>	0.807	0.717
41	ENSBTAG00000016465	<i>DHCR7</i>	0.903	0.521
42	ENSBTAG00000016709	<i>NT5C3A</i>	0.824	0.615
43	ENSBTAG00000016721	<i>ZNF791</i>	0.824	0.559
44	ENSBTAG00000016740	<i>ACLY</i>	0.918	0.52
45	ENSBTAG00000018936	<i>LSS</i>	0.839	0.58
46	ENSBTAG00000018959	<i>RAB11A</i>	0.828	0.67
47	ENSBTAG00000020984	<i>RAPGEF4</i>	0.856	0.775
48	ENSBTAG00000021842		0.804	0.492
49	ENSBTAG00000030951		0.844	0.508
50	ENSBTAG00000036260	<i>LPXN</i>	0.801	0.391
51	ENSBTAG00000037413	<i>TMEM164</i>	0.81	0.468
52	ENSBTAG00000047970		0.835	0.558

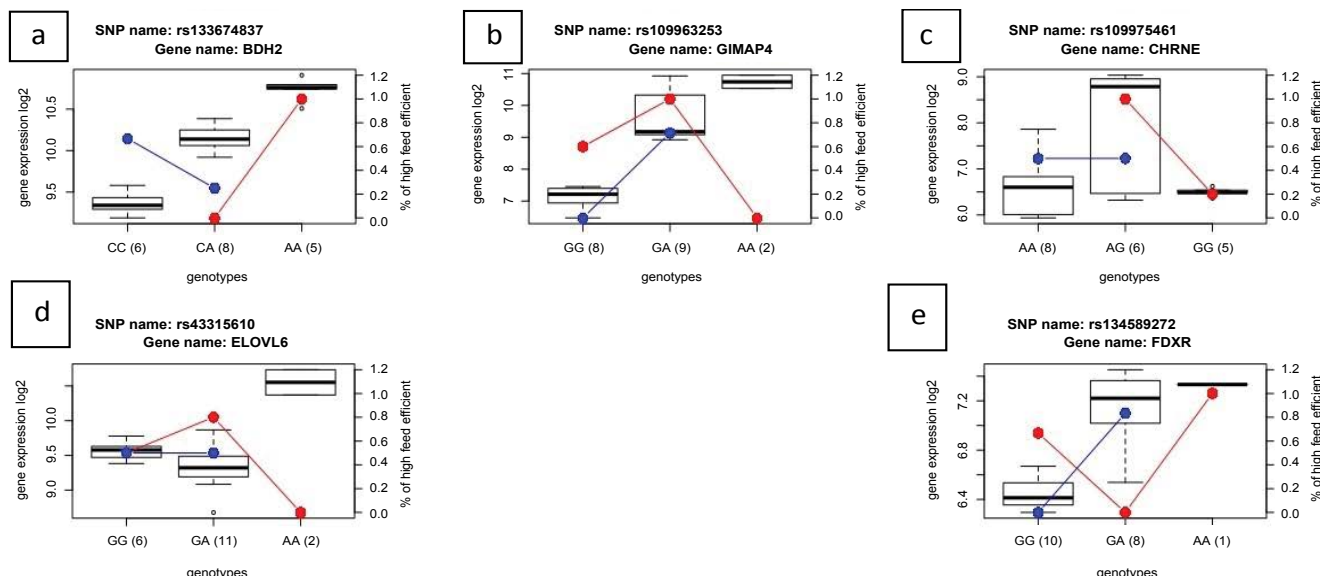
Supplementary Table 3: List of differentially expressed genes for Jersey.

	Ensembl.Gene.ID	Associated.Gene.Name	baseMean	log2FoldChange	padj
1	ENSBTAG00000006525	<i>FDXR</i>	125.97	-0.64501	6.21E-13
2	ENSBTAG00000008066	<i>PKDREJ</i>	76.74951	0.561566	1.15E-05
3	ENSBTAG00000013689	<i>MCTP2</i>	148.2616	0.528627	1.07E-05
4	ENSBTAG00000027727	Uncharacterized protein	284.1996	0.479634	0.000373
5	ENSBTAG00000038487	<i>ZNF613</i>	155.3981	-0.39026	0.026308
6	ENSBTAG00000046257	<i>GIMAP4</i>	650.1794	-0.38726	0.0024
7	ENSBTAG00000005182	<i>BOLA-A</i>	434.9538	-0.38721	0.001332
8	ENSBTAG00000014402	<i>GIMAP8</i>	713.0488	-0.38164	0.008566
9	ENSBTAG00000045727	Uncharacterized protein	921.1041	0.380952	0.032465
10	ENSBTAG00000019026	<i>EXTL2</i>	34.58975	0.377761	0.03724
11	ENSBTAG00000037440	<i>ZNF197</i>	281.1009	0.357972	0.016391
12	ENSBTAG00000021751	<i>RASEF</i>	36.02273	-0.3504	0.010571
13	ENSBTAG00000027205	<i>PGBD5</i>	30.06234	-0.34041	0.026308
14	ENSBTAG00000031737	<i>TMEM102</i>	26.52997	0.338561	0.03724
15	ENSBTAG00000009087	<i>GNG10</i>	1516.44	-0.32465	0.026308
16	ENSBTAG00000040323	Uncharacterized protein	1003.599	-0.32071	0.026308
17	ENSBTAG00000014161	<i>ARMC10</i>	258.84	-0.2983	0.026308
18	ENSBTAG00000013106	<i>C19orf81</i>	26.34468	0.295346	0.026308
19	ENSBTAG00000047379	<i>CYP3A4</i>	2422.366	0.286561	0.043386

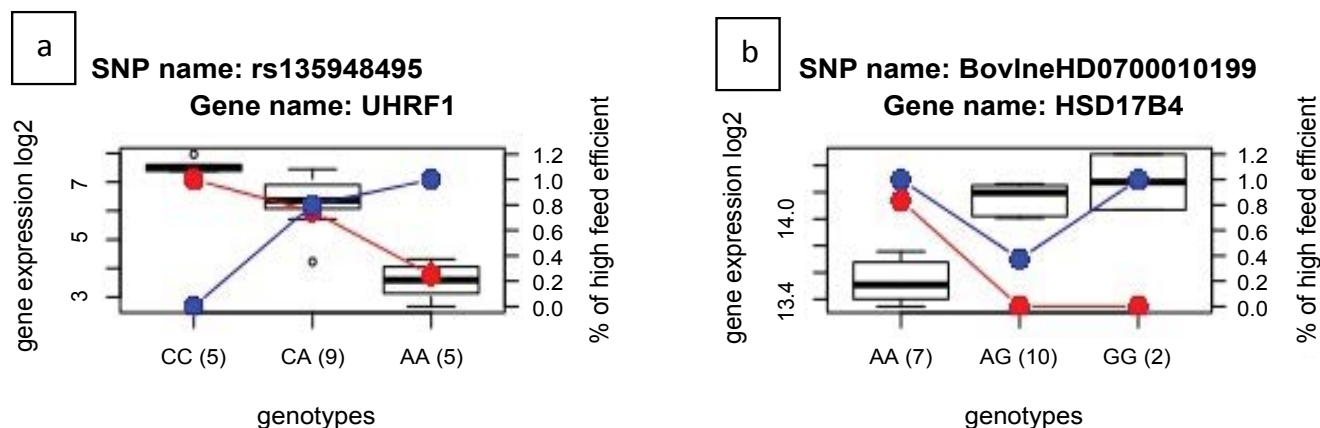
+ ve log2 fold change = upregulated in low feed efficiency group; - ve log2 fold change = downregulated in low feed efficiency group.

Supplementary Table 4: List of hub genes for Holstein.

	Ensembl gene ID	Gene name	Module membership	Gene significance
1	ENSBTAG00000000431	<i>TRDC</i>	0.858	-0.411
2	ENSBTAG00000000432	<i>TRAC</i>	0.86	-0.526
3	ENSBTAG00000000715		0.889	-0.487
4	ENSBTAG00000001198		0.81	-0.555
5	ENSBTAG00000002669	<i>RASSF4</i>	0.802	-0.722
6	ENSBTAG00000003037		0.829	-0.485
7	ENSBTAG00000004894		0.907	-0.497
8	ENSBTAG00000004917	<i>KLRK1</i>	0.826	-0.437
9	ENSBTAG00000005628		0.818	-0.49
10	ENSBTAG00000005892	<i>ZAP70</i>	0.864	-0.609
11	ENSBTAG00000006452	<i>CD3D</i>	0.9	-0.494
12	ENSBTAG00000006552	<i>LAMP3</i>	0.827	-0.501
13	ENSBTAG00000007191	<i>CCL5</i>	0.909	-0.48
14	ENSBTAG00000008401	<i>PFKFB3</i>	0.808	-0.547
15	ENSBTAG00000009381	<i>LCP2</i>	0.857	-0.654
16	ENSBTAG00000012695	<i>LCK</i>	0.852	-0.51
17	ENSBTAG00000013730	<i>CD5</i>	0.857	-0.403
18	ENSBTAG00000014725	<i>CD27</i>	0.822	-0.474
19	ENSBTAG00000015708	<i>CXCR6</i>	0.879	-0.469
20	ENSBTAG00000015710	<i>CD3E</i>	0.875	-0.537
21	ENSBTAG00000017256	<i>CD2</i>	0.914	-0.474
22	ENSBTAG00000019403	<i>MALSU1</i>	0.8	-0.536
23	ENSBTAG00000020904	<i>JAK3</i>	0.857	-0.439
24	ENSBTAG00000027246	<i>UBD</i>	0.888	-0.621
25	ENSBTAG00000030426		0.889	-0.379
26	ENSBTAG00000037510		0.853	-0.433
27	ENSBTAG00000038639	<i>CXCL9</i>	0.906	-0.425
28	ENSBTAG00000039588		0.815	-0.535
29	ENSBTAG00000047988		0.842	-0.365



Supplementary Figure 1: a-e) The boxplots show the five significant eQTLs with the associated genes for control diet group analysis. X-axis: genotypes; y-axis: gene expression (log2); red line: Holstein; blue line: Jersey.



Supplementary Figure 2: a,b) The boxplots show the two significant eQTLs with the associated genes for high concentrate diet group analysis. x-axis: genotypes; y-axis: gene expression (log2); red line: Holstein; blue line: Jersey.