Potential Functional Variants in Innate Immune Response Genes Associated with Feed Efficiency in Beef Cattle

**CURRENT STATUS:** POSTED

Gabriela Ribeiro  
Universidade de Sao Paulo

Aline Silva Mello Cesar  
Universidade de Sao Paulo

Pâmela Almeida Alexandre  
Universidade de Sao Paulo

José Bento Sterman Ferraz  
Universidade de Sao Paulo

Heidge Fukumasu  
Universidade de Sao Paulo

✉ fukumasu@usp.br **Corresponding Author**  
ORCiD: https://orcid.org/0000-0002-3265-5090

**DOI:**  
10.21203/rs.3.rs-18657/v1

**SUBJECT AREAS**  
Animal Science

**KEYWORDS**  
SNP, INDEL, transcriptome, variant calling, liver
Abstract

**Background:** Identifying and selecting animals for feed efficiency (FE) is extremely important for the beef production chain. Currently, the most common parameter to access the FE animals is residual feed intake (RFI), which is the residual of the linear regression that estimates DMI based on average daily gain and mid-test metabolic body weight. However, it relies on costly and time-consuming data collection, creating a growing demand for alternative approaches to identify genetically superior animals for FE. This study aimed to detect potential liver-specific functional variants from RNA-seq data of 16 Nellore bulls divergently selected for FE.

**Results:** The variant call analysis detected 247 missense SNPs and nine insertion-deletions (INDELs) that alter the protein functions. These variants were found within 190 genes differentially found ($P < 0.05$) in liver tissue between high FE (HFE) and low FE (LFE) animals. To better understand the role of these variants in biological pathways, we performed a functional enrichment analysis, which highlighted six genes involved in complement cascade and cascade complement regulation pathways, and 20 genes involved in the regulation of the innate immune system. They had four different significant variants in the complement factor H ($\text{CFH}$) family genes, and all were homozygous in HFE animals rather than some degree of heterozygous in LFE animals.

**Conclusion:** We developed a pipeline to detect potential liver-specific functional variants from RNA-seq data from animals divergently selected for feed efficiency. With this approach, we found potential functional variants in innate immune response genes associated with feed efficiency in beef cattle.

**Background**

Studies on feed efficiency (FE) in beef cattle are very important. The identification and selection of efficient animals can improve the productivity by reducing feed costs, which can reach 75% in feedlot systems (1). FE selection is also justified by the demand for less environmental impact of livestock. Feed efficient animals are recognized as more sustainable due to decreased production of waste and greenhouse gases.

One of the most common ways to evaluate FE is by residual feed intake (RFI), an independent measure of the level size and growth rate in beef cattle (2). At least five major physiological
processes contribute to RFI as intake of feed, digestion of feed, metabolism, physical activity and thermoregulation (3). Thus, the liver is considered a central organ for FE since it is responsible for the metabolism of nutrients as proteins, lipids and carbohydrates, along with other functions as metabolism of bilirubin, bile acids, xenobiotics, protein synthesis and immunity (4). Our group was the first to described a pathophysiological mechanism associated with FE in beef cattle: liver inflammation due to altered metabolism and/or bacterial translocation/infection (5), which was in part corroborated by others (6–8). We also unravel the metabolic pathways related to FE in Nellore cattle showing an increased bacterial load in low feed efficient animals (LFE) which is in part responsible for the hepatic lesions and inflammation in these animals (9).

The RFI estimation in beef cattle is costly and time-consuming, therefore other approaches to identify genetically superior animals are needed. A usual strategy is the use of molecular markers for genomic selection based on SNPs (single nucleotide polymorphisms) genotyping, which allows the detection of DNA variants associated with FE. However, the majority of the SNPs previously identified in Genome-Wide Association studies (GWAs) and expression quantitative loci (eQTL) association study (10) are non-causal variants (11–13), which are located in intron or intergenic genomic regions. There is an urgent need to establish the functional (causal) variants of this important phenotype. One possibility is the analysis of whole genome sequencing, an approach which is still very costly. Another possibility is the whole exome sequencing that has a reduced cost, but still lack information regarding the importance of the polymorphism within cell and/or tissue architecture. Therefore, we propose an approach to overcome these limitations based on the identification of genetic variants from RNA-sequencing (RNA-seq) data from a physiologically phenotype-related organ, followed by a classification of the potential functional variants according to their effects on protein expression and function.

In this work, we developed a pipeline to detect potential liver-specific functional variants from RNA-seq data from animals divergently selected for feed efficiency. With this approach, we found potential functional variants in innate immune response genes associated with feed efficiency in beef cattle.

Methods
**Phenotypic data and biological sample collection**

All animal protocols were approved by the Institutional Animal Care and Use Committee of Faculty of Food Engineering and Animal Sciences, University of São Paulo (FZEA-USP – protocol number 14.1.636.74.1).

All procedures to collect the phenotypes and biological samples were carried out at FZEA-USP, Pirassununga, State of São Paulo, Brazil. Ninety-eight Nellore bulls (16 to 20 months old and 376 ± 29 kg BW) were evaluated in a feeding trial comprised of 21 days of adaptation to feedlot diet followed by a 70-day period of data collection. Total mixed ration was offered *ad libitum* and daily dry matter intake (DMI) was individually measured. Animals were weighted at the beginning, at the end and every two weeks during the experimental period. Feed efficiency (FE) was estimated by residual feed intake (RFI) (2). At the end of the experimental period, liver biopsy samples from the left hepatic lobes were collected from each animal and quickly frozen in liquid nitrogen and stored at -80 °C. Further information about management and phenotypic measures of the animals used in this study can be found elsewhere (5).

**RNA-seq data**

Samples of 8 animals from each FE group (high and low) were selected for RNA-seq using RFI measure extremes. The total RNA from liver samples was extracted by using the RNeasy mini kit (QIAGEN, Crawley, West Sussex, UK) according to the instructions provided by the manufacturer. The total RNA quality and quantity were assessed using automated capillary gel electrophoresis on a Bioanalyzer 2100 with RNA 6000 Nano Labchips according to the manufacturer’s instructions (Agilent Technologies Ireland, Dublin, Ireland). Samples that presented RNA integrity number (RIN) less than 8.0 were discarded. The mRNA libraries were constructed using the TruSeq™ Stranded mRNA LT Sample Prep Protocol and sequenced on Illumina HiSeq 2500 equipment in a HiSeq Flow Cell v4 using HiSeq SBS Kit v4 (2x100pb). FastQC software (Babraham Institute, Cambridge, UK, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to obtain the sequencing quality. Removal of PolyA / T tails and adapters was performed using Seqyclean software (University of Idaho: Institute of Bioinformatics and Evolutionary Studies, Moscow, USA,
https://bitbucket.org/izhbannikov/seqyclean), so only bases with quality $\geq$ 20 and complete readings
with at least 50 bp, were further studied for further analysis. The alignment of the reads was done
using the STAR software version 2.7 (14) with the reference genome *Bos taurus* UMD3.1 (Ensembl,
ftp://ftp.ensembl.org/pub/release-98/fasta/bos_taurus/dna/), allowing two mismatches per read.

**Call of functional genomic variations associated with FE**

GATK software (15) was used to perform the call the variants analysis. HaplotypeCaller command was
applied to identify the variants (SNPs and insertions and deletions (INDELs). The variants underwent
quality control on the GATK software as follows: (Variant Quality Score - QUAL) > 30, depth of
sequencing (Deph Plot - DP) > 4, amount of available coverage (QualByDepth - QD) > 3, polarization
trend (FisherStrand - FS) > 30 and the general mapping quality of readings that support a variant call
(RMSMappingQuality - MQ) < 35. Statistical analysis was performed by Plink software (16) considering
MAF < 40% and call rate equal to 50%. For the allele frequency test between HFE and LFE groups, the
Cochran-Armitage trend analysis was used considering significant differences when $P<0.05$.

**Characterization of the effects of variants on protein sequence and function**

The potential functional variants were analyzed in the Variant Effect Predictor (VEP) online tool (17)
which predicts the functional effects of the variants. Potential functional variants were analyzed by
the Scale-Invariant Feature Transform (SIFT) score, a statistical tool of the VEP online software. SIFT is
an algorithm that predicts whether an amino acid substitution affects the function of the protein. This
analysis was performed on the basis of the gene sequence homology and the physical properties of
the amino acids, where the like sequences are first sought, then strictly related sequences (which
may share functions similar to the query sequence), with the alignment of the sequences chosen is
possible to calculate the normalized probability for all possible substitutions of the alignment. Thus,
generating a SIFT score ranging from 0 to 1, being classified as deleterious values below 0.05 and
above 0.05 is considered as tolerated. The deleterious classification indicates that the amino acid
change will have a great impact on the proteins, since the tolerated classification indicates that the
impact will not be so great. With this information it was possible to compare the position of the
variants in the protein with the protein database of the National Center for Biotechnology (NCBI), in
order to find information about the name of the region and what function that region performs.

**Functional enrichment analysis**

The functional enrichment analysis was performed with Panther version 14.1 (18) to identify biological pathways over-represented in the set of genes with potential functional variants. A multiple test correction was used and significant pathways were considered when p<0.05.

**Results**

**Characterization of feed efficiency groups**

Two groups of eight animals with extreme values of feed efficiency were selected and named High Feed Efficient (HFE, lowest RFI) and Low Feed Efficient (LFE, highest RFI). These groups were significantly different for feed efficiency traits RFI, feed conversion ratio (FCR), residual weight gain (RWG), residual intake and weight gain (RIG), for dry matter intake (DMI) and average daily gain (ADG) (Table 1). The LFE animals presented higher backfat thickness at the end of the experiment (P<0.05), supporting that HFE animals are more feed efficient because they eat less, have similar ADG and are leaner than LFE animals (5,19). A total of 11,361 genes were expressed in liver samples from HFE and LFE animals and eight genes were differentially expressed (DE) (P ≤ 0.1): NR0B2, SOD3, RHOB, Bta-mir-2904-2, FTL, CYP2E1, GADD45G and FASN (5).

**Characterization of expressed SNPs and INDELs associated with feed efficiency**

The variant calling analysis detected 268,393 and 37,587 SNPs and INDELs, respectively, from the 16 samples. These variants were tested for Minor Allele Frequency (MAF)<0.4 and call rate=50%, which left 68,852 SNPs and 5,148 INDELs, respectively, for statistical analysis. From these, 2,149 SNPs and 139 INDELs were significantly different between HFE and LFE animals (P<0.05). Part of these variants were found in exonic regions (51% for SNPs and 29% for INDELs) but a considerable fraction was found outside the gene (43% for SNPs and 66% for INDELs) or in introns (6% for SNPs and 5% for INDELs, Fig. 1).

We considered the variants with moderate and high impact as potential functional variants totalizing 256 variants (247 SNPs and 9 INDELs) selected for further analysis. They generated effects on proteins as follows: shorter proteins due to frameshift INDELs, bigger proteins due to insertion of one
or two AA by inframe insertion INDELs, changes in single amino acid (AA) on proteins due to missense SNPs and a stop retained INDEL (able to change at least one base of the terminator codon but maintaining the function of the terminator). These 256 potential functional variants were distributed across 190 genes, as some variants were found in spliced isoforms and there were genes with more than one variant (Supplementary table 1).

Depicting the biology of potential functional variants

In order to understand the importance of these 256 potential functional variants we performed the functional enrichment analysis for the selected genes, which detected three different significant pathways (False Discovery Rate - FDR<0.10): “Regulation of Complement cascade”, "Complement cascade" and "Innate Immune System" with fold enrichments of 14.79x, 12.19x and 2.57x respectively (Table 2).

The three pathways associated with FE by our approach were all related to the innate immune system, in particular with the regulation of the complement cascade. The first two pathways enriched from the same set of six genes (Table 3) and the Innate Immune System enriched from twenty genes (Table 3), including the six genes from the complement cascade/regulation of complement cascade pathways. The list of these 20 genes, their potential functional variants, the impact on protein sequence and the frequency in each group can be found in the Supplementary table 2. Considering the “regulation of complement cascade” and “complement cascade” as one pathway since they were enriched for the same set of genes, it calls the attention the overrepresentation of Complement H factor genes (complement factor H-related 5, complement factor H and complement factor H precursor). There were four different significant variants in these genes and all were homozygous in HFE animals instead of some degree of heterozygosity in LFE animals (Table 4).

Effects of other important potential causal variants

At last, we analyzed the effects of high impact INDELs on protein function since they altered the protein sequence considerably but were not enriched in the processes already described. The frameshift INDEL on GAS2L1 caused a deletion of 7 AA in the calponin homology domain and had an allele frequency of 35% (5/14) in the HFE animals whereas it was not found in LFE group (0/8,
Another frameshift INDEL generated a deletion of 7 AAs in the ATP binding cassette (ABC) domain of the TAP2 and had an allele frequency of 50% (7/14) in the HFE group whereas only 25% in the LFE group (3/12, p<0.05). Interestingly, this INDEL was found significantly associated not only in one but also in two spliced isoforms of TAP2. An INDEL in the transcription factor RREB1 caused an insertion of 2 AAs in one of the C2H2 Zinc finger domains of Bos indicus RREB1 and was associated with LFE since the allele frequency in this group was 30% (3/10) but in the HFE no allele was found (0/14, P<0.05). An INDEL was found in SEMA4F, which generated a premature stop codon reducing the COOH-terminal of the SEMA4F in 5 AAs. This region contains the PDZ domain of the protein responsible for anchoring the SEMA4F in the membrane to cytoskeletal components. The allele with the deletion of 40 bp was found in 50% in LFE group (5/10) and 16% in HFE group (2/12, P<0.05). An inframe insertion INDEL in the ECSIT insert a glutamic acid in the position 423 at the COOH terminal of ECSIT and this variant is associated with feed efficiency as the allele frequency was 29% (4/14) in the HFE and 0% (0/12) in the LFE group (P<0.05).

Another inframe insertion INDEL was found in the ENSBTAG0000011926 gene, which is probably the transcription factor ZFN665-like that inserted a leucine in one C2H2 Zn finger region of the protein. The variant with the inserted leucine is associated with lower feed efficiency with an allele frequency of 38% (3/8) instead of 6% (1/15) in the HFE group (P>0.05). Another potential functional variant is the insertion of one Alanine (252) just one position from the phosphorylation serine site of RPP30 (S251) with a higher frequency in the LFE group (LFE=50%, 5/10; HFE=14%, 2/14; p<0.05). The INDEL in the MGAT2 although considered a frameshift INDEL, it did not change the function since it generated a stop retained codon.

Discussion
There is increasing evidence for the importance of the immune system on feed efficiency in a variety of domestic species. Our group previously showed that feed efficiency is associated with altered inflammatory response in beef cattle partially because of bacterial translocation from the digestive tract to the liver (5,9). Here we reported for the first time the existence of potential functional variants in immune system-related genes associated with feed efficiency, with the potential to alter
the function of the innate immune system by regulation of the complement cascade. To achieve these results, we performed a screening for potential functional variants from transcriptomic data of animals evaluated for feed efficiency using RNA-seq data and bioinformatic tools such as GATK, VEP, functional enrichment using Panther and evaluation of effects from the potential variants.

Our pipeline found potential functional variants in genes that significantly enriched for three biological pathways related to the Immune System: “Regulation of Complement Cascade”, “Complement Cascade” and “Innate Immune System”. Similar genes enriched for Regulation of Complement Cascade and Complement Cascade (C8G, CFHR5, CFH, IGHG, CPN2) and it calls the attention the overrepresentation of potential functional variants in Complement H factor genes (CFHR5, CFH and CFH precursor). The complement system is one of the major biological processes of the innate immune system related to the ability of antibodies and phagocytic cells to clear microbes and damage cells inducing inflammation. The system consists of a number of small proteins mainly synthesized by hepatocytes, but can also be secreted by endothelial cells, white blood cells and epithelial cells (20,21). This enzymatic cascade helps in the defense of infections and is one of the main effectors of humoral immunity, regulating various biological processes such as phagocytosis, opsonization, leukocyte chemotaxis, release of mast cells, basophils and active oxygen species by leukocytes, vasoconstriction, smooth muscle contraction, increased vessel permeability, platelet aggregation and cytolysis (22,23). Activation of this cascade can be performed by 3 pathways (classical, lectin and alternative) and the genes found in our study regulate the alternative pathway, which is triggered in the presence of an exogenous activator, such as the presence of fungi, bacteria, some types of viruses and parasites, which activate C3 molecules and trigger the cascade (24–26). As already mentioned, we showed in previous works that less feed efficient animals have increased inflammatory response in the liver (5) and there is a higher level of endotoxins in the blood of LFE animals (9) which corroborates with the activation of the complement cascade by the alternative pathway. In fact, another study comparing the hepatic transcriptome of high and low FE animals found the complement system as the most significant canonical pathway enriched from the differential expressed genes (7).
The proteins responsible for regulation of the complement activation (RCA) can be divided into two main groups: membrane-bound regulators and soluble regulators (27). Factor H Complement belongs to the RCA family and we found different alleles predominantly in LFE animals. The CFH also has five additional members, represented by five separate CFH-related genes (CFHR), ranging from 1 to 5. The proteins produced by CFHR contain a set of domains that are homologous to those of CFH (28). These proteins act at the C3 level (29) negatively regulating complement activation, acting as a cofactor for factor I-mediated C3b cleavage and facilitating C3 convertase acceleration of deterioration (28,30). In this work, we found heterozygosity in the LFE animals for the CFH genes whereas the HFE animals were homozygous. Another gene that influences the complement system in all pathways (classical, lectin and alternative) is the C8 gene found in heterozygosity only in LFE animals. The C8 protein is part of the complex membrane attack complex (MAC) that assembles on bacterial membranes to form a pore, allowing the disruption of bacterial membrane organization, cell death and consequently inflammatory response (31,32).

The genetic modulation of the innate immune system supports our previous studies that demonstrate the relationship between feed efficiency and hepatic inflammation (5,9). Along with the genes related to complement cascade regulation, other genes with potential functional variations enriched for the innate immune system as BOLA, ECSIT, GLB1, ADA2, RIPK3, LILRA4, SIRPB1 and YEAR6. Olivieri and colleagues found candidate genes associated with feed efficiency traits in Nellore cattle involved in immune system as well as other processes (33). Specifically, they proposed the NLRP14, which was shown to regulate the innate immune signaling. In humans, a germline mutation in NLRP14 impairs its function and affects the innate immune signaling (34) as well as other polymorphisms in complement genes are linked to human diseases as hemolytic uremia and age-dependent macular degeneration (35), supporting the idea that potential functional variations affect the innate immune system.

Although we haven’t found the same candidate genes from the GWAS study of Olivieri and colleagues (2016), our approach confirmed the innate immune system as an important pathway for feed efficiency in beef cattle.

Our initial idea was to identify causal genetic variants for feed efficiency and found these to be
present exclusively in one group (LFE or HFE). However, in our study none of the 256 potential functional variants affecting protein sequences were found exclusively in one group and that is the reason why we choose not to call these as causal variants. It is shown for complex traits that even the most important loci in the genome have small effects sizes, and that, together, the significant hits only explain a modest fraction of the predicted genetic variance (36). Indeed, in one of the largest performed GWAS study for feed efficiency in beef cattle, the authors found ten significant quantitative trait locus (QTL) for RFI but none explained more than 2.5% of the additive genetic variance in any population (12).

We are aware that one possible limitation of the present work is the sample size but even in the present condition the results are well supported by the literature in transcriptomic (5–7) and GWAs studies (33), including other species as pigs (37,38) and poultry (39). Therefore, the existence of potential functional variations in genes of the innate immune response affecting feed efficiency in beef cattle seems plausible and worth future studies.

Conclusion
In this research, we found possible hepatic causal variants that modulate the expression/function of genes involved in the pathways innate immune response. Having as one of the main findings, four applicable homozygous variants for HFE animals in the CFH and CFH5 genes, which are responsible for phagocytosis of microbes and damaged cells that induce infection. Occasional changes in these variants regulate the alteration of immunity-related biological pathways, may alter or modify the model, which in our case is economical.

Abbreviations
ADG Average Daily Gain
DMI Dry Matter Intake
FCR Feed Conversion Ratio
FDR False Discovery Rate
FE Feed Efficiency
GATK Genome Analysis Toolkit
GWAS Genome-Wide Association Studies
HFE High Feed Efficiency
INDELs Insertion and Deletions
LFE Low Feed Efficiency
MAC Complex Attack Complex
MAF Minor Allele Frequency
NCBI Natural Center for Biotechnology Information
QTL Quantitative Trait Locus
RCA Regulation of the Complement Activation
RFI Residual Feed Intake
RIG Residual Intake and Weigh Gain
RIN RNA Integrity Number
RWG Residual Weigh Gain
SIFT Scale-Invariant Feature Transform
SNP Single Nucleotide Polymorphisms
VEP Variant Effect Predictor

Declarations

Funding
This work was financially supported by “Fundação de Amparo à Pesquisa do Estado de São Paulo” (FAPESP 2014/02493-7; 2014/07566-2;). FAPESP also granted the fellowships to G.R. (2016/11585-8; 2019/01234-12019/18647-7) and PAA (2015/22276-3; 2017/14707-0).

Availability of data and materials
The dataset supporting the conclusions of this article is available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress/) under accession number E-MTAB-3376,
http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-3376/

Authors’ contribution
GR, PAA, JBSF and HF designed the experimental study; GR, PAA and ASMC conducted the
bioinformatic analysis; GR and HF wrote the initial manuscript. All authors read, made critical revisions, and approved the final manuscript for submission.

**Ethics approval and consent to participate**

All animal protocols were approved by the Institutional Animal Care and Use Committee of Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo (FZEA-USP – protocol number 14.1.636.74.1).

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Acknowledgements**

The authors thank Mrs. Elisângela C. M. Oliveira and Ms. Arina L. Rochetti for all technical support.

**References**

1. Paper R. Simulation Modelling of the Cost of Producing and Utilising Feeds for Ruminants. 2010;14(2).

2. Koch RM, Swiger LA, Chambers D, Gregory KE. Efficiency of feed use in beef cattle. J Anim Sci. 1963;22:486–94.

3. Arthur JPF, Herd RM. Residual feed intake in beef cattle. Rev Bras Zootec. 2008 Jul;37(spe):269–79.

4. Stalker MJ, Hayes MA. Liver and biliary system. In: Jubb, Kennedy and Palmer’s Pathology of domestic animals. Elsevier Philadelphia, PA; 2007. p. 297–388.

5. Alexandre PA, Kogelman LJA, Santana MHA, Passarelli D, Pulz LH, Fantinato-Neto P, et al. Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. BMC Genomics. 2015;16(1).

6. Paradis F, Yue S, Grant JR, Stothard P, Basarab JA, Fitzsimmons C. Transcriptomic analysis by RNA sequencing reveals that hepatic interferon-induced genes may be
associated with feed efficiency in beef heifers. J Anim Sci. 2015 Jul;93(7):3331-41.

7. Tizioto PC, Coutinho LL, Decker JE, Schnabel RD, Rosa KO, Oliveira PSN, et al. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. BMC Genomics. 2015;16(1):1-14.

8. Weber KL, Welly BT, Van Eenennaam AL, Young AE, Port-Neto LR, Reverter A, et al. Identification of Gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-seq. PLoS One. 2016;11(3):1-19.

9. Fonseca LD, Eler JP, Pereira MA, Rosa AF, Alexandre PA, Moncau CT, et al. Liver proteomics unravel the metabolic pathways related to Feed Efficiency in beef cattle. Sci Rep [Internet]. 2019 Dec 29 [cited 2019 Apr 29];9(1):5364. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30926873

10. Cesar ASM, Regitano LCA, Reecy JM, Poleti MD, Oliveira PSN, de Oliveira GB, et al. Identification of putative regulatory regions and transcription factors associated with intramuscular fat content traits. BMC Genomics [Internet]. 2018 Jun 27 [cited 2020 Feb 24];19(1):499. Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-018-4871-y

11. Santana MHAMHA, Utsunomiya YTYT, Neves HHRHHR, Gomes RCRC, Garcia JFJF, Fukumasu H, et al. Genome-wide association analysis of feed intake and residual feed intake in Nellore cattle. BMC Genet. 2014;15:1-8.

12. Saatchi M, Beever JE, Decker JE, Faulkner DB, Freetly HC, Hansen SL, et al. QTLs associated with dry matter intake, metabolic mid-test weight, growth and feed efficiency have little overlap across 4 beef cattle studies. BMC Genomics. 2014;15(1).

13. Oliveira PSN, Cesar ASM, Nascimento ML, Chaves AS, Tizioto PC, Tullio RR, et al. Identification of genomic regions associated with feed efficiency in Nelore cattle. BMC Genet. 2014;15:100.
14. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013 Jan;29(1):15–21.

15. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010 Sep;20(9):1297–303.

16. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.

17. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biol [Internet]. 2016 Dec 6 [cited 2019 Nov 18];17(1):122. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-0974-4

18. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, et al. PANTHER: A library of protein families and subfamilies indexed by function. Genome Res. 2003;13(9):2129–41.

19. Novais FJ, Pires PRL, Alexandre PA, Dromms RA, Iglesias AH, Ferraz JBS, et al. Identification of a metabolomic signature associated with feed efficiency in beef cattle. BMC Genomics [Internet]. 2019 Dec 7 [cited 2019 Apr 29];20(1):8. Available from: https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-018-5406-2

20. Peng Q, Li K, Anderson K, Farrar CA, Lu B, Smith RAG, et al. Local production and activation of complement up-regulates the allostimulatory function of dendritic cells through C3a–C3aR interaction. Blood. 2008 Feb;111(4):2452 LP – 2461.

21. Strainic MG, Liu J, Huang D, An F, Lalli PN, Muqim N, et al. Locally produced complement fragments C5a and C3a provide both costimulatory and survival signals to naive CD4+ T cells. Immunity. 2008/03/06. 2008 Mar;28(3):425–35.
22. Frank MM, Fries LF. The role of complement in inflammation and phagocytosis. Immunol Today. 1991;12(9):322–6.

23. Haeney MR. The role of the complement cascade in sepsis. J Antimicrob Chemother. 1998 Jan;41(suppl_1):41–6.

24. Iturry-Yamamoto GR, Portinho CP. Sistema complemento: ativação, regulação e deficiências congênitas e adquiridas. Rev Assoc Med Bras. 2001;47(1):41–51.

25. Frank MM. Complement: A brief review. J Allergy Clin Immunol. 1989;84(4, Part 1):411–20.

26. Ochs HD, Wedgwood RJ, Frank MM, Heller SR, Hosea SW. The role of complement in the induction of antibody responses. Clin Exp Immunol. 1983 Jul;53(1):208–16.

27. Jiang C, Zhang J, Yao J, Liu S, Li Y, Song L, et al. Complement regulatory protein genes in channel catfish and their involvement in disease defense response. Dev Comp Immunol. 2015;53(1):33–41.

28. Zipfel PF, Skerka C, Hellwage J, Jokiranta ST, Meri S, Brade V, et al. Factor H family proteins: on complement, microbes and human diseases. Biochem Soc Trans. 2002 Nov;30(6):971 LP – 978.

29. Józsi M, Zipfel PF. Factor H family proteins and human diseases. Trends Immunol. 2008;29(8):380–7.

30. Zipfel PF, Heinen S, Józsi M, Skerka C. Complement and diseases: Defective alternative pathway control results in kidney and eye diseases. Mol Immunol. 2006;43(1):97–106.

31. Morgan BP. The membrane attack complex as an inflammatory trigger. Immunobiology. 2016;221(6):747–51.

32. Oikonomopoulou K, Ricklin D, Ward PA, Lambris JD. Interactions between coagulation and complement--their role in inflammation. Semin Immunopathol. 2011/08/03. 2012
33. Olivieri BF, Mercadante MEZ, Cyrillo JNDSG, Branco RH, Bonilha SFM, De Albuquerque LG, et al. Genomic regions associated with feed efficiency indicator traits in an experimental nellore cattle population. PLoS One. 2016 Oct 1;11(10).

34. Abe T, Lee A, Sitharam R, Kesner J, Rabadan R, Shapira SD. Germ-Cell-Specific Inflammasome Component NLRP14 Negatively Regulates Cytosolic Nucleic Acid Sensing to Promote Fertilization. Immunity. 2017 Apr 18;46(4):621–34.

35. Degn SE, Jensenius JC, Thiel S. Disease-causing mutations in genes of the complement system. Vol. 88, American Journal of Human Genetics. 2011. p. 689–705.

36. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. Vol. 169, Cell. Cell Press; 2017. p. 1177–86.

37. Horodyska J, Hamill RM, Reyer H, Trakooljul N, Lawlor PG, McCormack UM, et al. RNA-Seq of Liver From Pigs Divergent in Feed Efficiency Highlights Shifts in Macronutrient Metabolism, Hepatic Growth and Immune Response. Front Genet [Internet]. 2019 [cited 2019 Nov 18];10:117. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30838035

38. Ramayo-Caldas Y, Ballester M, Sánchez JP, González-Rodríguez O, Revilla M, Reyer H, et al. Integrative approach using liver and duodenum RNA-Seq data identifies candidate genes and pathways associated with feed efficiency in pigs. Sci Rep. 2018 Dec 1;8(1).

39. Zhou N, Lee WR, Abasht B. Messenger RNA sequencing and pathway analysis provide novel insights into the biological basis of chickens’ feed efficiency. BMC Genomics [Internet]. 2015 Dec 17 [cited 2019 Nov 18];16(1):195. Available from: http://www.biomedcentral.com/1471-2164/16/195

Tables
Table 1 - Phenotypic traits. Mean of the groups of high feed efficiency (n = 8) and low feed efficiency (n = 8).

| Trait      | HFE           | LFE           | P-value     |
|------------|---------------|---------------|-------------|
| BWi (kg)   | 413.20 ± 47.20| 407.80 ± 27.05| 0.78        |
| BWf (kg)   | 562.20 ± 49.50| 530 ± 30.76   | 0.10        |
| ADG (kg/d) | 2.13 ± 0.52   | 1.75 ± 0.23   | 0.08        |
| DMI (kg/d) | 10.20 ± 1.26  | 12.32 ± 0.95  | 2.09x10-3*  |
| FCR        | 4.96 ± 0.80   | 7.18 ± 0.62   | 2.82x10-5*  |
| RFI (kg/d) | −1.43 ± 0.33  | 1.66 ± 0.41   | 9.39x10-4*  |
| RWG (kg/d) | 0.38 ± 0.30   | −0.38 ± 0.14  | 7.98x10-5*  |
| RIG        | 1.80 ± 0.31   | −2.03 ± 0.37  | 9.39x10-4*  |
| REAi       | 67.45±5.38    | 65.84±4.08    | 0.5115      |
| REAf       | 83.38±7.44    | 82.79±4.54    | 0.8521      |
| REAg       | 15.93±10.75   | 16.95±5.02    | 0.812       |
| BFTi       | 1.087±1.19    | 1.97±1.32     | 0.1795      |
| BFTf       | 3.00±1.88     | 5.96±1.52     | 0.004*      |
| BFTg       | 1.91±2.06     | 3.99±1.15     | 0.03*       |
| RFTi       | 2.49±1.40     | 5.12±1.26     | 0.001*      |
| RFTf       | 5.69±2.58     | 9.50±2.05     | 0.006*      |
| RFTg       | 3.20±1.95     | 4.37±1.70     | 0.2206      |
HFE, high feed efficiency; LFE, low feed efficiency; BWi, initial body weight; BWf, final body weight; DMI, dry matter intake; ADG, average daily gain; FCR, feed conversion ratio; RFI, residual feed intake; RWG, residual body weight gain; RIG, residual intake and body weight gain; REAi, initial rib eye area; REAf, final rib eye area; REAg, gain of rib eye area; BFTi, initial back fat thickness; BFTf, final back fat thickness; BFTg, gain of back fat thickness; RFTi, initial rump fat thickness; RFTf, final rump fat thickness; RFTg, gain of rump fat thickness

*P ≤ 0.05

♦ Student’s t-test

☐ Mann–Whitney-Wilcoxon test

Adapted from Alexandre et al. 2015

Table 2. Significant pathways detected by functional enrichment analysis of genes with potential functional variants.

| Reactome pathways                  | Bos taurus REFLIST | Input genes | Fold Enrichment | P-value      | Q-value      |
|------------------------------------|--------------------|-------------|-----------------|--------------|--------------|
| Regulation of complement cascade   | 47                 | 6           | 14.79           | 5.97x10^{-06} | 9.60x10^{-03}|
| Complement cascade                 | 57                 | 6           | 12.19           | 1.65x10^{-05} | 1.32x10^{-02}|
| Innate immune system               | 901                | 20          | 2.57            | 1.31x10^{-04} | 6.99x10^{-02}|

Table 3. List of genes enriched for the significant pathways found from potential functional variants associated with Feed Efficiency.

| Pathways                          | Ensembl code          | Gene symbol | Gene name                      |
|-----------------------------------|-----------------------|-------------|--------------------------------|
| Regulation of Complement Cascade  | ENSBTAG00000010520    | C8G         | Complement C8 gamma chain      |
| Complement Cascade pathway        | ENSBTAG00000038171    | CFHR5       | Complement factor H-related 5  |
| Gene ID               | Description                                      |
|----------------------|--------------------------------------------------|
| ENSBTAG00000032656   | Complement factor H-related 5                  |
| ENSBTAG00000023177   | Complement factor H                              |
| ENSBTAG00000008959   | Complement C8 gamma chain                       |
| ENSBTAG000000018422  | Receptor interacting serine/threonine kinase 3  |
| ENSBTAG000000017313  | Adenosine deaminase 2                           |
| ENSBTAG000000038171  | Complement factor H-related                      |
| ENSBTAG00000007213   | Tyrosine-protein phosphatase nor substrate 1    |
| ENSBTAG00000006378   | Receptor (TNFRSF)-interacting serine/threonine  |
| ENSBTAG00000039520   | Signal-regulatory protein beta 1                |
| ENSBTAG00000026944   | Immunoglobulin receptor precursor               |
| ENSBTAG000000048135  | Member of immunoglobulin heavy chain            |
| ENSBTAG00000006859   | Carcinoembryonic antigen-related molecule       |
| ENSBTAG000000039520   | Signal-regulatory protein beta 1                |
| ENSBTAG00000039995   | Complement factor H                              |
| ENSBTAG0000002902   | Anoctamin                                       |
| ENSBTAG00000015254   | Beta-galactosidase                               |
| ENSBTAG00000006556   | Coatamer subunit beta                            |
| ENSBTAG000000032656   | Carboxypeptidase N subunit 2                    |
| ENSBTAG00000015049   | Evolutionarily conserved signaling pathway       |

**Innate Immune System**

| Gene ID               | Description                                      |
|----------------------|--------------------------------------------------|
| ENSBTAG00000032656   | Complement factor H-related 5                  |
| ENSBTAG00000008959   | Complement C8 gamma chain                       |
| ENSBTAG000000018422  | Receptor interacting serine/threonine kinase 3  |
| ENSBTAG00000039995   | Complement factor H                              |
| ENSBTAG00000006378   | Receptor (TNFRSF)-interacting serine/threonine  |
| ENSBTAG000000023177  | Complement factor H                              |
| ENSBTAG000000048135  | Member of immunoglobulin heavy chain            |
| ENSBTAG00000006859   | Carcinoembryonic antigen-related molecule       |
| ENSBTAG000000039520   | Signal-regulatory protein beta 1                |
| ENSBTAG000000026944   | Immunoglobulin receptor precursor               |
| ENSBTAG00000039995   | Complement factor H                              |
| ENSBTAG0000002902   | Anoctamin                                       |
| ENSBTAG00000015254   | Beta-galactosidase                               |
| ENSBTAG00000006556   | Coatamer subunit beta                            |
| ENSBTAG000000032656   | Carboxypeptidase N subunit 2                    |
| ENSBTAG00000015049   | Evolutionarily conserved signaling pathway       |
Table 4. Frequency and missense alteration of the potential functional variants enriched from the complement cascade pathway associated with feed efficiency.

| Gene ID                | Gene | SNP | Position protein | Protein alteration | LFE | HFE | P-value |
|------------------------|------|-----|------------------|--------------------|-----|-----|---------|
| ENSBTAG00000023177     | CFH  | T/A | 445              | I/K                | 2|6  | 0|16     | 0,021  |
| ENSBTAG00000038171     | CFHR5| T/C | 558              | P/S                | 3|11 | 0|16     | 0,031  |
| ENSBTAG00000039995     | CFH  | G/C | 15               | P/R                | 4|6  | 0|8      | 0,011  |
| ENSBTAG00000039995     | CFH  | C/T | 30               | Y/H                | 4|8  | 0|8      | 0,031  |

Figures
Figure 1

Relative frequency of significant SNPs and INDELs between HFE and LFE groups classified by consequence and impact on protein expression. SNPs and INDELs classified by the consequence on gene sequence (A and B) and by the impact on protein expression (C and D).

Supplementary Files

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

Supplementary tables.xlsx