

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
DEPARTAMENTO DE GENÉTICA E EVOLUÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA
MOLECULAR

ANDRESSA OLIVEIRA DE LIMA

GENES E VARIANTES GENÉTICAS NA REGULAÇÃO
DA EFICIÊNCIA DE GADO NELORE

SÃO CARLOS -SP
2019

ANDRESSA OLIVEIRA DE LIMA

GENES E VARIANTES GENÉTICAS NA REGULAÇÃO DA EFICIÊNCIA ALIMENTAR
EM GADO NELORE

Tese apresentada ao programa de Pós-graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos – UFSCar, como parte dos requisitos para obtenção do título de Doutor em Ciências (Ciências Biológicas), área de concentração: Genética e Evolução.

Orientadora: Prof. Dra. Luciana Correia de Almeida Regitano

Co-orientadora: Dra. Polyana Cristine Tizioto

São Carlos-SP
2019



UNIVERSIDADE FEDERAL DE SÃO CARLOS


Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

Folha de Aprovação


Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado da candidata Andressa Oliveira de Lima, realizada em 08/05/2019:



Profa. Dra. Luciana Correia de Almeida Regitano
EMBRAPA



Prof. Dr. Anderson Ferreira da Cunha
UFSCar



Prof. Dr. Marcos Roberto Chiaratti
UFSCar



Prof. Dr. Luis Artur Loyola Chardulo
UNESP



Profa. Dra. Aline Silva Mello Cesar
ESALQ/USP

I dedicate this work to my parents, José e Ana Luiza, my brothers Guilherme and Leandro, and to my boyfriend, Fernando, for the love, patience and supporting me in the life journey.

ACKNOWLEDGMENTS

To Dr. Luciana Correia de Almeida Regitano for advise me during this journey, patience, friendship, and the contribution in my academic formation.

To International collaborator and supervisor at Iowa State University (ISU), prof. Dr. James E. Koltés, for the support and help during my sandwich doctorate scholarship and collaboration in this study.

To Dr. Polyana Cristine Tizioto for the friendship, support and help during my Ph.D and the contribution in this study.

To CAPES for the Ph.D. scholarship, CNPq for the sandwich Doctorate scholarship, and CNPq and FAPESP (#grant 2012-23638-8) for the financial support in the project.

To UFSCar and the Graduation Program of Evolutionary Genetics and Molecular Biology for the contribution in my academic formation.

To my parents José Garcia and Ana Luiza for the love and support, and my brothers, Leandro e Guilherme, for friendship and support.

To my boyfriend, Fernando Minussi, for the love, friendship, and supporting me during my Ph.D.

To the members and ex-members from animal Biotechnology group in the Embrapa Southeast Livestock for the support, help and the friendship constructed during my Ph.D., especially Juliana Afonso, Marina Ibelli, Jessica Malheiros, and Wellison Diniz.

To Dr. Flávia Bressani for support in the biotechnology lab and friendship.

To the members of Animal Breeding and Genetics at ISU for the support and help during my sandwich doctorate scholarship, especially for Mary Shue.

“There should be no boundaries to human endeavor. We are all different. However bad life may seem, there is always something you can do and succeed at. While there’s life, there is hope.”

Stephen Hawking

RESUMO

Genes e variantes genéticas na regulação da eficiência alimentar de gado Nelore: Os custos com alimentação podem representar mais da metade do custo total de produção. A eficiência alimentar é uma característica complexa, que pode contribuir com a redução de gastos com alimentação dos animais, podendo ainda estar associada à redução dos impactos ambientais relacionados com a ocupação territorial e com a emissão de gases poluentes, como o metano. Apesar do perfil de expressão em grupos contrastantes para eficiência alimentar ter sido investigado em diferentes tecidos, resultando na identificação de vias e processos biológicos relacionados a esse fenótipo, a relação da expressão gênica com a variação contínua de características de eficiência alimentar é pouco explorada na literatura. Assim, dentre os genes previamente identificados como sendo diferencialmente expressos (DE) em tecido hepático de grupos extremos de consumo alimentar residual (CAR) de bovinos da raça Nelore, selecionou-se genes em vias biológicas importantes para eficiência alimentar. Com base nessa seleção, verificou-se a influência da expressão dos genes *COL1A1*, *CTGF*, *CYP2B6*, *EGR1*, *PRUNE2* e de uma isoforma deste último (*PRUNE2_isoform*) na variação de características relacionadas à eficiência alimentar, a saber: consumo de matéria seca (CMS), consumo alimentar residual (CAR), conversão alimentar (CA), eficiência alimentar (EA), ganho de peso diário (GPD), índice de Kleiber (IK), peso médio (PM), peso médio metabólico (PMM) e taxa relativa de crescimento (TRC). Para isso, realizou-se ensaios de PCR quantitativa em tempo real (RT-qPCR) dos genes e isoforma selecionados em 52 amostras do tecido hepático de bovinos Nelore, seguido pela análise de associação por meio de um modelo linear misto. Esse estudo revelou que o padrão de expressão do gene *PRUNE2* desfavorece a eficiência alimentar, e esse fato pode estar relacionado indiretamente à disfunção mitocondrial. No entanto, o padrão de expressão da sua isoforma (*PRUNE2_isoform*) aparentemente atua de forma contrária à expressão total do gene, favorecendo a eficiência alimentar. Além disso, para melhor entendimento dos mecanismos genéticos envolvidos com a eficiência alimentar, nós realizamos análise de rede de coexpressão gênica, construída pela metodologia *weighted gene co-expression network analysis* (WGCNA) a partir de 180 amostras de tecido muscular (*Longissimus thoracis*) de bovinos Nelore, identificou-se 391 potenciais biomarcadores (*hub genes*) relacionados com a variação da eficiência alimentar. Tais *hub genes* participam de vias relacionadas com síntese de proteínas, crescimento muscular e resposta imune. Entre esses *hub genes* nós destacamos os *CCDC80*, *FBLN5*, *SERPINF1* e *OGN* associados com as características CA, EA, GPD, IK e TRC. Esses genes estão relacionados com a homeostase da glicose, estresse oxidativo e formação óssea. Entre os *hubs genes*, identificou-se que 13 fatores de transcrição descritos para bovinos, e seis desses revelaram potenciais reguladores de outros *hub genes* identificados nesse estudo. Entre eles, o fator de transcrição *TCF4* pode desempenhar uma função importante no crescimento muscular e é um potencial regulador de genes previamente identificados por nosso grupo de pesquisa como DE no tecido muscular em grupos extremos de consumo alimentar residual (CAR). Finalmente, identificou-se potenciais regiões regulatórias e variantes funcionais relacionadas aos potenciais biomarcadores identificados nesse estudo.

Palavras-chave: *Bos indicus*, eficiência alimentar, elementos regulatórios, redes gênicas de co-expressão, RT-qPCR, SNPs, WGCNA.

ABSTRACT

Genes and genetic variants in the regulation of feed efficiency in Nelore cattle: Feeding accounts for most of the costs in beef cattle production. To reduce it, as well as the environmental impact, greenhouse gas emission, and land occupation have been mandatory to improve the animal feed efficiency. Multifactorial, feed efficiency (FE) has been evaluated by different indexes and approaches. Among them, genomic studies from animals genetically divergent for FE pointed out candidate genes and pathways such as energy metabolism, inflammatory and oxidative stress. However, these studies did not take into account the continuous variation of the gene expression within the population. Thereby, from a previous differential approach carried out in divergent Nelore steers for FE, we selected six hepatic candidate genes (*COL1A1*, *CTGF*, *CYP2B6*, *EGR1*, *PRUNE2*, and the *PRUNE2_isoform*) based on their biological role on pathways related to feed efficiency. We carried out a real-time quantitative PCR (RT-qPCR) assay in 52 Nelore steers to evaluate the hepatic expression profile of the overmentioned genes and their association with FE related-traits such as average daily gain (ADG), body weight (BW), dry matter intake (DMI), feed conversion ratio (FCR), feed efficiency ratio (FE), Kleiber index (KI), metabolic body weight (MBW), residual feed intake (RFI), and relative growth rate (RGR). Based on a linear mixed model, we identified that the total expression of *PRUNE2* has an unfavorable effect on feed efficiency related-traits, likely related to mitochondrial dysfunction. However, when taking only the *PRUNE2_isoform*, we observed a favorable effect on the evaluated trait. Still, to shed light on the genetic mechanisms affecting feed efficiency in Nelore, we applied a co-expression approach using muscle RNAseq data from 180 animals. Based on the weighted gene co-expression network analysis (WGCNA) software, we identified 391 potential biomarkers (hub genes) related to feed efficiency variation. These hub genes partook in protein synthesis, muscle growth, and immune response pathways. Among the hub genes, we highlighted *CCDC80*, *FBLN5*, *SERPINF1*, and *OGN* genes, which were associated to ADG, FCR, FE, KI, and RGR traits, and were related to glucose homeostasis, oxidative stress, and osteogenesis. Furthermore, we found 13 transcription factors among the hub genes described for bovine and six of them are putative regulators for the others hub genes identified in this study. Among them, the *TCF4* may have a role in muscle growth metabolism and regulator of DE genes for divergent RFI in muscle previously identified by our research group. Finally, we identified potential regulatory regions and functional variants related to the potential biomarkers identified in this study.

Keywords: *Bos indicus*, feed efficiency, regulatory elements, co-expression network, RT-PCR, SNPs, WGCNA.

LIST OF TABLES

	Page
Table 2.1. Gene symbol, primers sequences, amplicon sizes for the RT-qPCR gene expression assays for five feed efficiency candidate genes and reference genes.....	27
Table 2.2. Means, standard deviations (SD) and ranges of the measured feed efficiency-related traits of Nelore steers for residual feed efficiency.	29
Table 2.3. Summary of the significant association of target genes' expression levels to feed efficiency-related traits.....	30
Table 3.1. Mean, standard deviation (SD), minimum (Min) and maximum (Max) observed for feed efficiency related-traits in Nelore steers.	45
Table 3.2. Description of significantly feed efficiency related-traits associated modules in Nelore cattle.	47
Table 3.3. Description of the putative transcription factors binding sites in the cis-eQTLs for the hub genes <i>PCDH18</i> and <i>SPARCL1</i>	51

LIST OF FIGURES

	Page
Figure 1.1. A representative example of Nelore steer.....	12
Figure 1.2. Cis-regulatory elements involved in the eukaryotic gene expression regulation.	16
Figure 2.1. Pearson’s correlation matrix among the adjusted expression values (CtA) for all the target genes.	31
Figure 2.2. mRNA sequence alignment and functional protein domain annotation for PRUNE2 gene isoforms.	32
Figure 3.1. Clustering dendrogram of genes in modules. Each color represents one module eigengene (ME).	46
Figure 3.2. Linear module trait association for all FE related-traits. Beta coefficients are displayed on top and p-values in parenthesis.	47
Figure 3.3. Venn diagram showing the overlapped hub genes with linear expression effect on more than one trait.	48
Figure 3.4. Number of hub genes per pathway identified by functional annotation and enrichment analysis performed by ClueGO software.	49
Figure 3.5. Hub genes and the enriched pathways interaction.....	49
Figure 3.6. Integration of the TFs (hub-TFs) with their targets.	50
Figure 3.7. Integrative network of the hub genes identified in the present co-expression network, with regulatory elements (ASE, TFs, eQTLs) and genes previously associated with FE (DEGs and QTLs).	52

TABLE OF CONTENTS

	Page
Chapter 1 INTRODUCTION AND BACKGROUND	12
1.1. INTRODUCTION.....	12
1.1.2. Beef cattle in Brazil	12
1.1.3. Nelore breed.....	12
1.1.4. Feed efficiency related-traits.....	13
1.1.5. Genomic regions and candidate genes associated with feed efficiency	14
1.1.6. Systems biology and feed efficiency	15
1.1.7. Cis- factors affecting the gene expression patterns.....	16
1.2. OBJECTIVES	17
1.2.1. Specific objectives:	17
1.3. References.....	19
Chapter 2 The expression pattern of <i>PRUNE2</i> gene is related to feed efficiency in Nelore cattle	23
2.1. INTRODUCTION	24
2.2. MATERIAL AND METHODS	25
2.2.1. Production of experimental animals and sampling.....	25
2.2.2. Real-time PCR analysis (RT-qPCR).....	26
2.2.3. Expression data normalization and association analysis	27
2.2.4. Protein domain functional annotation for <i>PRUNE2_isoform</i>	28
2.3. RESULTS	29
2.4. DISCUSSION	32
2.5. CONCLUSION	35
2.6. References.....	35
Chapter 3 Potential biomarkers for feed efficiency related-traits in Nelore cattle identified by co-expression network.....	39
3.1. INTRODUCTION.....	40
3.2. MATERIALS AND METHODS	41
3.2.1. Animals and phenotypic traits	41
3.2.2. Library preparation, RNA sequencing, and data processing	42
3.2.3. Co-expression Network and module trait association (MTA) analyses	42
3.2.4. Hub genes identification, functional annotation and pathway over-representation analyses	43
3.2.5. Enrichment analysis to identify TFBS in hub genes.....	44
3.2.6. Data integration for identification of regulatory regions and functional variants ...	44

3.3. RESULTS	44
3.3.1. Background of animals and expression data.....	45
3.3.2. Weighted gene co-expression network analysis (WGCNA) and module trait association (MTA)	45
3.3.3. Hub genes and their association with feed efficiency.....	46
3.3.4. Hub genes' pathway over-representation analyses	48
3.3.5. Enrichment analysis to identify TFBS in hub genes.....	50
3.3.6. Data integration for identification of regulatory regions and functional variants ...	51
3.4. DISCUSSION	52
3.4.1. Module-trait association analysis.....	52
3.4.2. Hub genes identification and their linear expression association with feed efficiency.....	53
3.4.3. Pathway over-representation analysis for the hub genes	54
3.4.3.1. Protein synthesis metabolism.....	54
3.4.3.2. Immune system	55
3.5. Enrichment analysis to identify TFBS in hub genes.....	58
3.6. Data integration for identification of regulatory regions and functional variants ...	58
3.7. CONCLUSION	59
3.8. References.....	60
4.0 GENERAL CONCLUSION	68
5.0 SUPPLEMENTARY INFORMATION	68

Chapter 1 INTRODUCTION AND BACKGROUND

1.1. INTRODUCTION

1.1.2. Beef cattle in Brazil

The Brazilian agribusiness accounts for 22% of national gross domestic product (GDP), being livestock responsible for 31% of this amount (ABIEC, 2018). Besides, the commercial bovine herd is composed of around 222 million head (ABIEC, 2018). In 2018, Brazil produced around 9.71 tons carcass weight equivalent (CWE), being the first one exporter and the second producer of the world, and out of that, a total of 2 million were exported mainly to Hong Kong, China, Egypt, and Russia (ABIEC, 2018).

1.1.3. Nelore breed

Nelore breed represents one of the first *Bos indicus* animals that came from India to Brazil, where they were multiplied in all country because of their adaptability for tropical climate (OLIVEIRA et al., 2002). In this context, about 80% of the national herd is composed of pure zebu or crossbred breed cattle (MARIANTE et al., 1984; OLIVEIRA, 2002). Nowadays, the most popular *Bos indicus* beef breed in Brazil is the Nelore (Figure 1.1), which composes the majority part of the herd (USDA, 2018).

Figure 1.1. A representative example of Nelore steer.



Source: Karina Santos.

Nelore breed has rusticity and adaptation to the extensive breeding system, high digestion capacity of low-quality fibers, natural resistance to parasites, low metabolism and greater productive and reproductive performance when compared to *Bos taurus* (OLIVEIRA et al., 2002).

1.1.4. Feed efficiency related-traits

Animal feeding is related to the energy of maintenance (ARTHUR; HERD, 2008) since feed intake and digestibility depend on the interaction among food, animal and environment (MERTENS et al., 1987). Feed efficiency can contribute to improve meat production and carcass quality allied to the reduction of the costs of animal feeding (NASCIMENTO et al., 2014), which depends strongly on the production system, but may represent up to 60% of the total cost of production (ARTHUR; HERD, 2008; CONNOR, 2015; MONTAÑO-BERMUDEZ et al., 1990). In addition, efficient animals have been associated with reducing environmental impacts, as greenhouse gases emissions and the use of natural resources for meat production (BASARAB et al., 2003 KHIAOSA-ARD; ZEBELI, 2014; VELAZCO, 2017).

Feed efficiency (FE) is traditionally evaluated by feed conversion ratio (feed intake/gain; kg/kg) (BRODY, 1945) and its inverse, feed efficiency (gain/feed intake; kg/kg), but both are associated with growth rates (HERD; BISHOP, 2000). However, proposed by Koch et al. (1963), residual feed intake (RFI), measures the difference between feed intake observed and feed intake predicted, considering the individual maintenance requirements and is independent of growth rates (HERD; ARTHUR, 2009), which can reduce maintenance demand for adult animals, as proposed by Koch et al. (1963) and confirmed by Lima et al. (2017) and Khiaosa-Ard and Zebeli (2014). This trait is estimated from the residuals, resulting from a regression of feed intake (DMI, Dry matter intake), $BW^{0.75}$ (MBW, metabolic body weight), and average daily gain (ADG, regression coefficient of weight in days) on intake (DMI) by the equation: $DMI = \beta_0 + \beta_1 * (MBW) + \beta_2 * (ADG) + \epsilon$. Besides these FE measures, there are the Kleiber index (KI, ADG/MBW) (Kleiber, 1936) and the relative growth rate (RGR, $100 * (\log BW_{\text{final}} - \log BW_{\text{Initial}}) / \text{total days of the experiment}$) (FITZHUGH; TAYLOR, 1971).

Thus, feed efficiency measures can improve profitability in beef production since feed-efficient animals consume less food and improve the ratio of feed intake to gain in comparison to inefficient animals (BASARAB et al., 2007). However, it is important to

consider the biological mechanisms related to the different measures, as selection for growth rates associated traits can increase nutrient demand in adult animals.

In this context, studies that investigate biological processes and pathways can contribute to the understanding of the biological mechanisms involved in the differences of these phenotypes and to the identification of essential pathways related to feed efficiency variation.

1.1.5. Genomic regions and candidate genes associated with feed efficiency

Genomics tools as genotyping in high-density SNP panels, DNA and RNA sequencing at low-cost can make possible to investigate and help to understand the complexity of the genomes (LEVY; MYERS, 2016). Thus genome-wide studies and gene expression profiling can add more layers of information to feed efficiency traits, since they have polygenic nature resulting in complex biological mechanisms (MOORE et al., 2009).

Genome-wide association studies (GWAS) are useful to detect genomic regions associated with phenotypes (VISSCHER et al., 2017). Based on that, some candidate genes for feed efficiency, that are related to energy, protein and lipids metabolisms, and ion transport immune response in Nelore cattle were reported (De OLIVEIRA et al., 2014; OLIVEIRA et al., 2016)

Gene expression profiling is also an important tool to investigate changes in global gene expression by estimating the abundance of mRNA on a biological material (RØSOK; SIOUD, 2007). Zarek et al. (2017) reported differentially expressed (DE) genes in liver for gain: intake from several pure and crossbred beef cattle. One of these genes is *CAMK2* that is involved in the glucose regulation. The same authors highlighted the role of immune response genes on feed efficiency.

Immune response genes were also related to FE in pigs. Gondret et al. (2017), in a multi-tissue study, reported immune response enriched DE genes that were activated in higher efficiency animals. Horodyska et al. (2018) also highlighted the importance of the immune response to FE. They investigated DE genes for FE in muscle and identified genes like *PIK3C2B*, which is related to T-cell activation, and *IL-8*, that is involved in leukocytes trafficking to the site of inflammation. In addition, Horodyska et al. (2019) studying hepatic tissue reported that in efficient animals pathways related to immune cell activation and differentiation were activated.

The oxidative stress processes were found activated in inefficient Nelore animals (TIZIOTO et al., 2015; TIZIOTO et al., 2016). In addition, these authors pointed out the *EGR1* as a potential regulator of genes related to oxidative stress.

A common feature of all gene expression studies reported so far is the use of disruptive sampling aiming to compare contrasting individuals with regard to feed efficiency related-phenotypes. Thus, the rationale for these experiments is that genes whose expression differed between extremes are a candidate to affect the quantitative variation of the trait, but it is not possible from this design to quantify the effect of each gene in the trait variation.

1.1.6. Systems biology and feed efficiency

The biological network's approach is effective to investigate the interaction between, for example, gene-gene or protein-metabolite using mathematical probabilistic or statistical network modeling (LIANG; KELEMEN, 2017).

The biological system is classified by the nature of compounds and interactions involved (SERIN et al., 2016). Co-expression network analysis can indicate genes with similar expression patterns and simultaneously active in the same biological processes and different conditions (DAM et al., 2018; SERIN et al., 2016). The similarity in this approach is inferred by the correlation measure between each pair of genes or by mutual information (DAM et al., 2018). In addition, with the co-expression network, it is possible to identify clusters of genes (highly connected genes) and infer the relationship of biological processes and the phenotype (DAM et al., 2018).

One of the used co-expression network methods is Weighted Gene Correlation Network Analysis (WGCNA). This approach contains a comprehensive set of functions and the main steps are: i) network construction (based on correlation measures) ii) identification of the modules (clusters of densely interconnected genes using hierarchical clustering method) iii) finding the relation between modules with external information (correlation between the module information and the trait) (LANGFELDER; HORVATH, 2008).

Besides that, it is possible to identify potential biomarkers for the phenotype of interest through the identification of hub genes in co-expression networks (ALEXANDRE et al., 2015). In this analysis, the goal is to summarize the most representative nodes which are the highly connected nodes in the modules (LANGFELDER; HORVATH, 2008). The genes which have a crucial biological role will presumably be the highly connected nodes (hub genes) (RHEE; MUTWIL, 2014) when compared to the other genes in the networks.

Studies using co-expression network approach identified pathways and potential biomarkers associated with feed efficiency. Alexandre et al. (2015) identified biomarkers as the *DOLK* gene, related to inflammation response, and the *PPP3CB* gene, associated with FE and energy expenditure in cattle. Salleh et al. (2018) reported candidate genes related to immune response and highlighted the role of the immune system in feed efficiency. Weber et al. (2016) identified cattle transcription factors as *HHEX* and *CDKNA2A/B*, related to the insulin pathway, and the *GATA3* and *STAT3* genes, related to feeding behavior, in co-expression networks. Furthermore, De Oliveira et al. (2018) investigated the interaction between mRNA-miRNA co-expression networks associated with FE in Nelore and reported pathways related to the immune system, oxidative stress, and lipid metabolism.

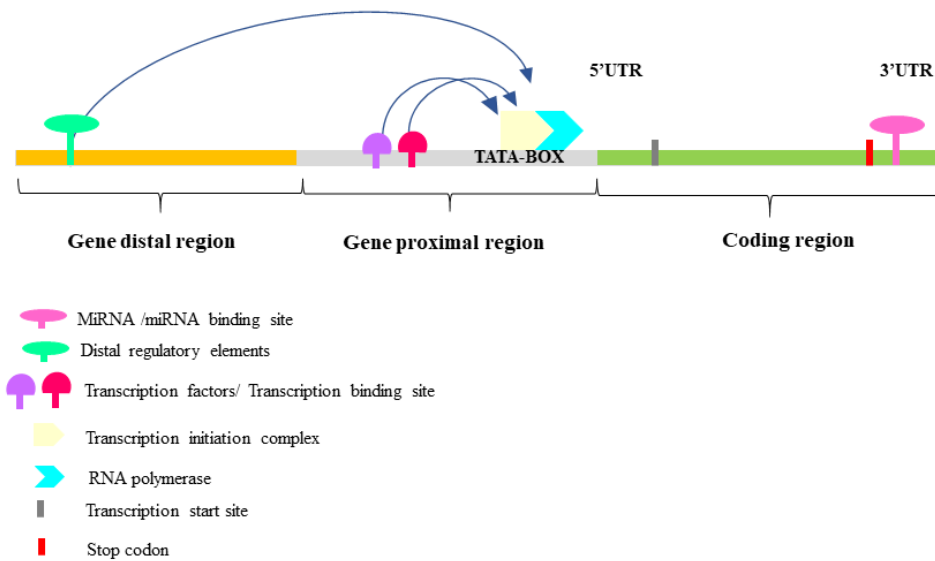
Hence, the biological networks are being used in several studies as a useful approach to investigate many genes simultaneously expressed in several pathways and under different conditions. This approach is not only important to understand the mechanisms related to feed efficiency in beef production, but also to identify biomarkers for this trait.

1.1.7. Cis- factors affecting the gene expression patterns

The majority part of genome is composed of noncoding regions (MATTICK, 2001; FEHLMANN et al., 2017). Among these regions, those that are conserved within species are enriched with cis-regulatory elements (DOUGLAS; HILL, 2014), which influence the pattern of gene expression.

Elements in the upstream region as core promoter, transcription factor binding site (TFBS) and distal regulatory elements can contribute to the regulation of the gene expression (HERNANDEZ-GARCIA; FINER, 2014) (Figure 1.2). The transcription is initiated when the transcriptional initiation complex attaches to the core promoter region (in Figure 1.2 represented by TATA-BOX). Simultaneously the transcription factors bind in specific DNA sequence motifs (transcription factor binding sites) and interact with the transcriptional initiation complex leading to enhancement or suppression of the gene expression. Besides that, distal cis-regulatory elements can get closer to the transcriptional initiation complex by conformational changes in the structure of DNA and chromatin and participate in the regulation of gene expression (HERNANDEZ-GARCIA; FINER, 2014).

Figure 1.2. Cis-regulatory elements involved in the eukaryotic gene expression regulation.



Source: Adapted from Hernandez-Garcia; Finer, 2014; TATA-BOX = Core promoter

Furthermore, miRNAs, short non-coding RNA, can directly bind to a sequence in the 3'UTR of the mRNA and can lead to mRNA degradation (TAMMEN et al., 2013; YUAN; WEIDHAAS, 2018). Based on this, alteration in the miRNA binding site can change the interaction miRNA-mRNA and lead to a complete disruption or the creation of new miRNA binding sites (YUAN; WEIDHAAS, 2018).

In this context, Cohen-Zinder et al. (2016), investigating variants in the promoter region of the *FABP4* gene, identified one SNP associated with RFI in Holstein's calves. Thus, variants in regulatory regions can change the sequences of the binding sites and modify the interaction between the regulatory elements and their targets, and this can contribute to the feed efficiency variation in beef cattle.

1.2. OBJECTIVES

The aim of this study was to identify potential regulators and genes, as well as functional variants, associated with the quantitative variation of feed efficiency-related traits.

1.2.1. Specific objectives

- i) To verify the relationship between gene expression and feed efficiency-related traits (Average daily gain-ADG, Body weight-BW, Feed conversion ratio-FCR, Feed efficiency ratio-FE, Kleiber index-KI, Metabolic body weight-MBW,

Relative growth ratio-RGR, Residual feed intake-RFI), in genes previously identified as DE in the liver of divergent residual feed intake Nelore steers;

- ii) To apply the Weighted Gene Co-expression Network Analysis (WGCNA) to identify potential gene expression regulators and essential genes (hub genes) in the muscle (*Longissimus thoracis*) associated with feed efficiency-related traits;
- iii) To verify the influence of the hub genes in feed efficiency-related traits;
- iv) To integrate data from previous studies in the same population with the hub genes found herein to identify regulatory regions and potential functional variants.

1.3. REFERENCES

ABIEC -Associação Brasileira das Indústrias Exportadoras de Carne: **Exportação Brasileira de carne bovina**. Brazilian livestock profile: Annual Report. Disponível em: <<http://www.brazilianbeef.org.br/download/sumarioingles2018.pdf>>. Acesso em: mar. 2019.

ALEXANDRE, P. A. et al. Liver transcriptomic networks reveal the main biological processes associated with feed efficiency in beef cattle. **BMC Genomics**, v. 16, n. 1073, p. 2-13, 2015.

ARTHUR, P.F.; HERD, R.M. Residual feed intake in beef cattle. **Revista Brasileira de Zootecnia**, Viçosa, v. 37, n. esp. 37, p. 269-279, 2008.

ARTHUR, P. F.; HERD, R. M. Efficiency of feed utilization by livestock — implications and benefits of genetic improvement. **Canadian Journal of Animal Science**, v. 85, n. 3, p. 281-290, 2005.

BASARAB, J. A. et al. Residual feed intake and body composition in young growing cattle. **Canadian Journal of Animal Science**, v. 83, n. 2, p. 180-204, 2003.

BASARAB, J. A. Relationships between progeny residual feed intake and dam productivity traits. **Canadian Journal of Animal Science**, v. 87, n. 4, p. 489-502, 2007.

BRODY, S. **Bioenergetics, and growth; with special reference to the efficiency complex in domestic animals**. Oxford: Reinhold, 1945. 1023 p.

CONNOR, E.E. Invited review: improving feed efficiency in dairy production: challenges and possibilities. **Animal**, v. 9, n. 3, p. 395-408, 2015.

COHEN-ZINDER, M. et al. Fabp4 is a leading candidate gene associated with residual feed intake in growing Holstein calves. **Physiological Genomics**, v. 48, n. 5, p. 367-376, 2016.

DAM, S. V. et al. Gene co-expression analysis for functional classification and gene-disease predictions. **Briefings in Bioinformatics**, v. 19, n. 4, p. 575–592, 2018.

DE OLIVEIRA, P.S.N. et al. Identification of genomic regions associated with feed efficiency in nelore cattle. **BMC Genetics**, v. 15, n. 1, p. 1-10, 2014.

DE OLIVEIRA, P. S. N. et al. An integrative transcriptome analysis indicates regulatory mRNA-miRNA networks for residual feed intake in nelore cattle. **Scientific Reports**, v. 8, n. 17072, p. 1-12, 2018

DOUGLAS, A. T.; HILL, R. E. Variation in vertebrate cis-regulatory elements in evolution and disease. **Transcription**, v. 5, n. 3, p. e28848, 2014.

FEHLMANN, T. et al. A review of databases predicting the effects of SNPs in miRNA genes or miRNA-binding sites. **Briefings in Bioinformatics**, p. 1-10, 2017.

FITZHUGH JR, H. A.; TAYLOR, C. S. Genetic analysis of the degree of maturity. **Journal of Animal Science**, v. 33, n. 4, p. 717-725, 1971.

GONDRET, F. et al. A transcriptome multi-tissue analysis identifies biological pathways and genes associated with variations in feed efficiency of growing pigs. **BMC Genomics**, v. 18, n. 244, p. 1-17, 2017.

HERNANDEZ-GARCIA, C. M.; FINER, J. J. Identification, and validation of promoters and cis-acting regulatory elements. **Plant Science**, v. 217, n. 218, p. 109-119, 2014.

HORODYSKA, J. et al. RNA-Seq of muscle from pigs divergent in feed efficiency and product quality identifies differences in immune response, growth, and macronutrient and connective tissue metabolism. **BMC Genomics**, v. 19, n. 1, p. 791, 2018.

HORODYSKA, J. et al. RNA-Seq of liver from pigs divergent in feed efficiency highlights shifts in macronutrient metabolism, hepatic growth, and immune response. **Frontiers in Genetics**, v. 10, n. 1, p. 17, 2019

KHIAOSA-ARD, R.; ZEBEL, Q. Cattle's variation in rumen ecology and metabolism and its contributions to feed efficiency. **Livestock Science**, v. 162, p. 66-75, abr. 2014.

KLEIBER, M. Problems involved in breeding for the efficiency of food utilization. **Journal of Animal Science**, v. 1936b, n. 1, p. 247-258, 1936.

KOCH, R. et al. Efficiency of feed use in beef cattle. **Journal of Animal Science**, v. 22, n. 2, p. 486-494, 1963.

LANGFELDER, P.; HORVATH S. WGCNA: an R package for weighted correlation network analysis. **BMC Bioinformatics**. v. 9, n. 1, p. 559, 2008

LEVY, S. E.; MYERS, R. M. Advancements in next-generation sequencing. **Annual Review of Genomics and Human Genetics**, v. 17, p. 95-115, 2016.

LIANG, Y.; KELEMEN, A. Dynamic modeling and network approaches for omics time course data: an overview of computational approaches and applications. **Briefings in Bioinformatics**, v. 19, n. 5, p. 1051-1068, 2017.

LIMA, Natália Ludmila Lins. Economic analysis, performance, and feed efficiency in feedlot lambs. **Revista Brasileira de Zootecnia**, v. 46, n. 10, p. 821-829, 2017

HERD, M. R.; BISHOP, S. C. Genetic variation in residual feed intake and its association with other production traits in British Hereford cattle. **Livestock Production Science**, v. 63, n. 2, p. 111-119, 2000

MARIANTE, A.S. et al. Resultados do controle de desenvolvimento ponderal. Raça Nelore. **EMBRAPA - CNPGC**. Campo Grande. Documentos, 25, 76 p. 1984.

MATTICK, J. S. Non-coding RNAs: the architects of eukaryotic complexity. **Embo Reports**, v. 2, n. 11, p. 986-991, 2001.

- MERTENS, D. R. Predicting intake, and digestibility using mathematical models of ruminal function. **Journal of Animal Science**, v. 64, n. 5, p. 1548-1558, 1987.
- MONTANO-BERMUDEZ et al. Energy requirements for maintenance of crossbred beef cattle with the different genetic potential for milk. **Journal of Animal Science**, v. 68, n. 8, p. 2279-2288, 1990.
- MOORE, S. S. et al. Molecular basis for residual feed intake in beef cattle. **Journal of Animal Science**, v. 87, n. suppl_14, p. E64-E71, 2009.
- NASCIMENTO M. L. et al. Feed efficiency indexes and their relationships with the carcass, non-carcass, and meat quality traits in Nelore steers. **Meat Science**, v. 116, p. 78-85, jun.2016.
- OLIVEIRA, de et al. **Nelore:base genética e evolução seletiva no brasileiro**. Embrapa Cerrados. Documentos, 2002. 50 p.
- OLIVIERI, B. F. et al. Genomic regions associated with feed efficiency indicator traits in an experimental Nelore cattle population. **Plos One**, v. 11, n. 10, p. e0164390, 2016.
- RHEE, S. Y.; MUTWIL, M. Towards revealing the functions of all genes in plants. **Trends in Plant Science**, v. 19, n. 4, p. 212-221, 2014.
- RØSOK, Ø; M., Sioud. Discovery of differentially expressed genes. In: **Technical considerations: Target discovery and validation reviews and protocols**. Methods in molecular biology. 360 ed. Humana Press, 2007. 115 p.
- SALLEH, S. M. et al. Gene co-expression networks from RNA sequencing of dairy cattle identifies genes and pathways affecting feed efficiency. **BMC Bioinformatics**, v. 19, n. 513, p. 1-15, 2018.
- SERIN, E. A. et al. Learning from co-expression networks: possibilities and challenges. **Frontiers in Plant Science**, v. 7, p. 444, 2016.
- TAMMEN, A. S. et al. Epigenetics: the link between nature and nurture. **Molecular Aspects of Medicine**, v. 34, n. 4, p. 753-764, 2013.
- TIZIOTO, P. C. et al. Global liver gene expression differences in nelore steers with divergent residual feed intake phenotypes. **BMC Genomics**, v. 16, n. 242, p. 1-14, 2015.
- TIZIOTO, P. C. et al. Gene expression differences in longissimus muscle of nelore steers genetically divergent for residual feed intake. **Scientific Reports**, v. 6, n. 39493, p. 1-12, 2016.
- UNITED STATES DEPARTMENT OF AGRICULTURE. LIVESTOCK AND POULTRY: **WORLD MARKETS AND TRADE**. Disponível em: <http://apps.fas.usda.gov/psdonline/circulars/livestock_poultry.pdf>. Acesso em: 28 mar. 2019.

VELAZCO, J. I. et al. Daily methane emissions and emission intensity of grazing beef cattle genetically divergent for residual feed intake. **Animal Production Science**, v. 57, n. 4, p. 627-635, 2016

VISSCHER, P. M. et al. 10 years of GWAS discovery: biology, function, and translation. **American Journal of Human Genetics**, v. 101, n. 1, p. 5-22., 2017.

WEBER, K. L. et al. Identification of gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-seq. **Plos One**, v. 11, n. 3, p. e0152274, 2016.

YUAN, Y.; WEIDHAAS, J. B. Functional microRNA binding site variants. **Molecular Oncology**, v. 13, n. 1, p. 4-8, 2019.

ZAREK et al. Differential expression of genes related to gain and intake in the liver of beef cattle. **BMC Research Notes**, v. 10, n. 1, p. 1, 2017

Chapter 2 The expression pattern of *PRUNE2* gene is related to feed efficiency in Nelore cattle

ABSTRACT

Feed efficiency improvement is beneficial to cattle production system, reducing the feeding costs and environmental impacts, and increasing the muscle in carcass. However, feed efficiency related-traits have polygenic nature as expected from their complex biological architectures. Previous studies linked the genes *CYP2B6*, *EGR1*, *COL1A1*, *CTGF*, *PRUNE2* and an isoform of the last gene to feed efficiency by comparing contrasting animals. In order to test the relationship of these gene's expression levels to feed efficiency in a continuous population distribution, we implemented an association analysis using a general linear mixed model for the RT-qPCR expression of these genes in the liver of 52 Nelore steers and their feed efficiency-related traits (ADG, BW, DMI, FCR, FE, KI, MBW, RFI, and RGR). This analysis revealed that the *PRUNE2* gene expression decreases the feed efficiency, which can be related to the fact that this gene may inhibit the AKT activity by interacting with BCL2 superfamily related to mitochondrial dysfunction. A new *PRUNE2* isoform found in RNA-Seq data from the same Nelore population was confirmed here. The expression of this new isoform increases feed efficiency revealing antagonism with *PRUNE2* gene expression in Nelore cattle.

Keywords: *Bos indicus*; gene expression; residual feed intake, *PRUNE2* isoform

2.1. INTRODUCTION

The Brazilian agribusiness represents 22% of national gross domestic product (GDP), being livestock responsible for 31% of this amount (ABIEC, 2018). In addition, the country is the first one exporter and the second producer of the world. Most of the Brazilian beef herd is composed of Nelore breed (*Bos indicus*) (USDA, 2018). Feed still corresponds to the most substantial monetary investment in beef cattle at a feedlot and is the crucial factor in determining profitability for producers (CLEMMONS et al., 2018; HILL et al., 2012; SILVA et al., 2016).

Feed efficiency is related to the increase in meat production per quantity of food offered, which reduces the costs in animal feed (NASCIMENTO et al., 2014). Efficient animals have been associated with a significant reduction on environmental impact from greenhouse gases emissions and the use of natural resources for meat production (BASARAB et al., 2003; KHIAOSA-ARD; ZEBELI, 2014; VELAZCO, 2017). Thus, improving feed efficiency may contribute to increasing sustainability and profitability in the beef cattle industry (BASARAB et al., 2003).

Variability in feed efficiency among individuals is commonly measured by the residual feed intake (RFI) approach, which is the difference between observed and predicted feed intake required for maintenance and growth (KOCH et al., 1963).

Feed efficiency has a polygenic nature, which results from complex biological mechanisms (MOORE et al., 2009). In this respect, different genes with significant effects on feed efficiency have been identified in beef cattle (ALEXANDRE et al., 2015; KERN et al., 2016; KONG et al., 2016; SALLEH et al., 2017; TIZIOTO et al., 2015; TIZIOTO et al., 2016; WEBER et al., 2016). Among these genes, cytochrome P450 subfamily 2B (*CYP2B6*), early growth response protein 1 (*EGRI*), connective tissue growth factor (*CTGF*), collagen type I alpha 1 chain (*COL1A1*), and prune homolog 2 (*PRUNE2*) have been related as possible biomarkers in cattle (FONSECA et al., 2015; KHANSEFID et al., 2017; SUN et al., 2018; TIZIOTO et al., 2015; TIZIOTO et al., 2016).

CYP2B6 gene has a role in oxidative stress metabolism of endogenous substrates, including steroids and fatty acids (TIZIOTO et al., 2015). *EGRI* gene exerts control of cell differentiation, growth, apoptosis and oxidative stress (PAGEL; DEINDL, 2011; TIZIOTO et al., 2016). *CTGF* gene has been related to increases in muscular tissue fibrosis (LIPSON et al., 2012). *COL1A1* expresses a protein constituent of the type I collagen, which is part of the structural component of the extracellular matrix (LI et al., 2016). Finally, *PRUNE2* gene has been related to apoptosis (LI et al., 2011; POTKIN et al., 2009).

Although these studies reported essential biological processes, considering different genes, belonging to distinct biological pathways by comparing divergent phenotypes, further studies investigating the gene-phenotype relationship with the use of continuous variables are needed to increase our understanding of the genetic mechanisms involved in the feed efficiency of Nelore cattle. By adopting continuous variation in the study of gene expression-phenotype relationships, one could avoid losses in phenotypic variability (SEO et al., 2016).

In this context, the objective of this study was to verify the association of the expression patterns of the genes *CYP2B6*, *EGR1*, *CTGF*, *COL1A1* and *PRUNE2* to feed efficiency related traits quantitative variation in Nelore cattle.

2.2. MATERIAL AND METHODS

2.2.1. Production of experimental animals and sampling

The experimental procedures were conducted in accordance with Institutional Animal Care and Use Committee Guidelines of the Brazilian Agricultural Research Corporation – EMBRAPA. Protocol CEUA 01/2013).

Our experiment used a subsample of the last year of feed efficiency trials (2009) from the Nelore steers population described in Nascimento et al. (2016). In summary, 52 animals stayed in the experiment for about 21 months, from birth to slaughter. Feed efficiency tests were carried out at Embrapa Southeast Livestock (São Carlos, SP, Brazil). The animals were allocated in individual pens, with an adaptation period of 28 days and feedlot for approximately 70 days, individually fed with 40% silage and 60% concentrate twice a day.

All the measures and evaluation for feed efficiency related-traits were described in de Oliveira et al. (2014) and Nascimento et al. (2016). In this study, we adopted average daily gain (ADG, kg/d), body weight (BW, kg), dry matter intake (DMI, kg/d), metabolic body weight (MBW, kg), feed conversion ratio (FCR, feed intake/gain; kg/kg), feed efficiency ratio (FE, gain/feed intake; kg/kg), Kleiber index (KI, ADG/MBW; kg/kg), residual feed intake (RFI, Kg/d) and relative growth rate (RGR, %/d). The animals were sent to slaughter according to the guidelines for the Humane Slaughter of Cattle. During slaughter, the carcasses were properly identified, and liver samples were collected and, immediately frozen in liquid nitrogen and kept at - 80°C.

2.2.2. Real-time PCR analysis (RT-qPCR)

Total RNA was extracted by homogenization of 100 mg of frozen liver in TRIzol® (Life Technologies Corporation, USA) following the manufacturer's instructions. The RNA concentration was measured by 260 nm UV absorbance (NanoDrop® ND-1000, Thermo Fisher Scientific) and the quality verified by the 260 nm/280 nm intensity ratio, followed by RNA integrity evaluation in agarose gel electrophoresis. Total RNA was further treated with DNase I (Deoxyribonuclease I - Invitrogen®) and submitted to cDNA synthesis by reverse transcription using SuperScript III (Invitrogen®), following the manufacturer's protocol.

Selection of the target genes was based on the results of previous differential gene expression analysis for divergent genetic group of residual feed intake (RFI) (10 samples low RFI and 10 samples high RFI), as described in Tizioto et al. (2015). Among the differentially expressed (DE) genes, *CYP2B6*, *EGR1*, *COL1A1*, *CTGF* and, *PRUNE2* were prioritized based on being enriched for biological processes related to the trait. The PCR primers for the targets (Table 2.1) were designed using Primer-Blast (YE et al., 2012), based on the transcript assembly from Tizioto et al. (2015). A new *PRUNE2* bovine isoform (*PRUNE2_isoform*) identified on that transcript assembly was also targeted in the RT-qPCR experiment. Furthermore, in order to obtain the *PRUNE2* total gene expression, regardless of isoform, we designed one pair of primers for targeting the common sequences between the known *PRUNE2* isoform (*PRUNE2_201*) and this new isoform (*PRUNE2_isoform*). In addition, the reference genes *HRPT1* and *YWHAZ* were used for relative quantification normalization. The quality control of the primers and specificity was performed using NetPrimer software (<http://www.premierbiosoft.com/netprimer/netprimer.html>).

The gene expression RT-PCR assays were performed in duplicate, randomized on the 96-Well Real-Time PCR and analyzed in a 7500 Real-Time PCR system® (Applied Biosystems, USA) equipment. cDNA samples (0.2-fold) were amplified by Power Up™ SYBR™ Green Master Mix (Thermo Fisher Scientific, USA). All the reactions were performed on the same amplification conditions: initiated with 2 min at 50 °C, cycle of 10 min at 95 °C, followed by 40 cycles of 15 sec at 95 °C, and 1 min at 60 °C, including the dissociation curve steps for 15 sec at 95 °C, 1 min at 60 °C, 95 °C at 30 sec, and 60 °C for 15 sec.

Table 2.1. Gene symbol, primers sequences, amplicon sizes for the RT-qPCR gene expression assays for five feed efficiency candidate genes and reference genes.

Gene	Sequence (5' - 3')	Amplicon size (bp)
<i>YWHAZ</i> ^a	F: GAACTCCCCTGAGAAAGCCT R: CCGATGTCCACAATGTCAAG	46
<i>HPRT1</i> ^a	F: TGCTGAGGATTTGGAGAAGG R: CAACAGGTCCGGCAAAGAAGT	105
<i>CTGF</i>	F: ATCTGTGCTTCTAACTGGGGA R: TGGGGTTGACGGACTATTC	91
<i>EGR1</i>	F: ACCTGACCGCAGAGTCCTTT R: TTGGCTGGGGTAACTCGTCT	77
<i>PRUNE2</i>	F: TGTCAACTGCTATTCACTTAGC R: ATGTCGGTATTATGTCTTAGAAG	184
<i>PRUNE2_isoform</i> ^b	F: TAACTGCTCCAGACAGGGGA R: AGCAATGGGACTCAGCATCC	143
<i>COL1A1</i>	F: TGGAAGAGCGGAGAATACTGG R: GGGTATACACAGGTCTCACCG	101
<i>CYP2B6</i>	F: TAACCACAGCCTCCCTTTGTC R: CTCACTGCATAGGGCTACTGG	140

F: Forward; R: Reverse; bp: bases pair; ^a constitutive internal control; ^bTargets only the new isoform of *PRUNE2*

2.2.3. Expression data normalization and association analysis

The environmental effects on the traits were tested using mixed models implemented on the MIXED procedure of SAS®. As our sample belonged to a single cohort in the original experiment, only age at slaughter was included as a covariate. The PCR efficiency for each sample was calculated by LinRegPCR (RUIJTER et al., 2009) and the cycle threshold (Ct) was adjusted for theoretical maximum PCR efficiency (100%) (BUSTIN et al., 2009), as described in Tizioto et al. (2013).

The BLUP (Best linear unbiased predictor) of Ct values for the random effects associated with the samples were obtained using a general linear mixed model according to the following equation:

$$y_{ij} = \mu + G_i + A_j + b_1(S_{ij} - \bar{S}) + \varepsilon_{ij}$$

Where:

y_{ij} is the Ct of the gene reference for the i^{th} reference gene, of the j^{th} sample;

μ is the average of Ct;

G_i is the fixed effect for the j^{th} reference gene ($j = \text{gene 1, gene 2}$);

A_j is the random effect associated with the j^{th} sample, taking account $A_j \sim \text{NID}(0, \sigma_a^2)$;

$b_1(S_{ijk} - \bar{S})$ is the regression coefficient associated with the animal's age at slaughter;

S_{ij} is the animal's age at slaughter; \bar{A} is the mean age at slaughter;

ε_{ijk} is the random residual effect, with $\varepsilon_{ijk} \sim \text{NID}(0, \sigma_e^2)$.

The degrees-of-freedom correction was performed using the Satterthwaite method (Satterthwaite, 1946), being also considered the variance for each reference gene.

The BLUP values were used to generate the adjusted cycle threshold (CtA). Pearson's correlation between the CtA values of each pair of target genes was estimated using the Corrplot R package (WEI; SIMBKO, 2017).

A general linear mixed model, including simultaneously all target genes' adjusted expression, was used to verify their association with all the feed efficiency related-traits (ADG, BW, DMI, FCR, FE, KI, MBW, RGR and, RFI) to avoid biases from collinearity among the variables. The degrees-of-freedom were corrected by the Satterthwaite method (Satterthwaite, 1946). The analysis was performed using the Statistical Analysis Software (SAS Institute, Cary, NC, USA, 2011) according to the following equation model:

$$y_{ijk} = \mu + b_1(S_{ij} - \bar{S}) + \sum_{k=1}^6 g_k(G_{ijk} - \bar{G}_K) + \varepsilon_{ijk}$$

Where:

y_{ij} is the i^{th} trait for the j^{th} animal;

μ is an overall mean;

$b_1(S_{ijk} - \bar{S})$ is the regression coefficient associated with the animal's age at slaughter;

S_{ij} is the animal's age at slaughter; \bar{A} is the mean age at slaughter;

G_{ijk} is the k^{th} gene expression adjusted, for the i^{th} trait, for the j^{th} animal;

\bar{G}_K is the mean for the k^{th} gene expression adjusted ($k = 1, 2, \dots, 6$);

ε_{ijk} is the random residual effect, with $\varepsilon_{ijk} \sim \text{NID}(0, \sigma_e^2)$.

2.2.4. Protein domain functional annotation for *PRUNE2* isoform

For the functional annotation of the isoform identified from liver RNA-Seq, according to the protein family and domains predicted, the sequences of amino acid residues were

predicted based on the transcript assembly from Tizioto et al. (2015), using the online ExPASy translate tool (<https://web.expasy.org/translate/>). Next, sequences of amino acid residues were analyzed by InterPro (Protein sequence analysis & classification) (FINN et al., 2017) and NCBI's Conserved Domain Database (CDD) (MARCHLER-BAUER et al., 2016).

2.3. RESULTS

After removing samples with Ct variation between duplicates and null-expression values, we used the gene expression from 52 samples. Descriptive statistics of feed efficiency-related traits for these 52 Nelore steers are presented in Table 2.2. Three of these animals did not have records for RFI.

Table 2.2. Means, standard deviations (SD) and ranges of the measured feed efficiency-related traits of Nelore steers for residual feed efficiency.

Traits	N	Mean ± SD	Minimum	Maximum
Average daily gain (ADG, kg/d)	52	1.62 ± 0.23	0.99	2.17
Body weight (BW, kg)	52	346.6 ± 34.1	280.5	429.5
Dry matter intake (DMI, kg/d)	52	8.79 ± 1.04	6.4	11.2
Feed conversion ratio (FCR, feed intake/gain; kg/kg)	52	5.50 ± 0.76	3.51	6.88
Feed efficiency ratio (FE, gain/feed intake; kg/kg)	52	0.19 ± 0.03	0.15	0.29
Kleiber index (KI, ADG/MBW; kg/kg)	52	0.02 ± 0.002	0.014	0.027
Metabolic body weight (MBW, kg)	52	80.26 ± 5.93	68.5	94.3
Residual Feed Intake (RFI, Kg/d)	49	-0.014 ± 0.63	-1.29	1.53
Relative growth rate (RGR, %/d)	52	0.20 ± 0.03	0.13	0.27

N = number of animals with phenotype

In the general linear mixed model analysis between feed efficiency related-traits and the expression of each candidate gene (*CYP2B6*, *EGR1*, *COL1A1*, *CTGF*, *PRUNE2*, and *PRUNE2_isoform*), only *PRUNE2* total gene and *PRUNE2_isoform* expression significantly ($P \leq 0.05$) affected phenotypes (Table 2.3, Table-S1).

There was a negative relation between *PRUNE2* gene Ct and the measures of DMI (-0.3194 kg/d), FCR (-0.2427 kg/kg), RFI (-0.2557 kg/d), and a positive relation to FE (0.008361 kg/kg), which, given the inverse relationship between Ct and target mRNA

abundance, means that increasing the *PRUNE2* expression results in increasing DMI, FCR, RFI, and decreasing FE.

On the other hand, *PRUNE2_isoform* gene expression levels showed significant association ($P \leq 0.05$) only with FCR (0.2436) and FE (-0.00756), and both associations were in the opposite direction compared to *PRUNE2* gene expression (Table 2.3). These results suggest that the expression of *PRUNE2_isoform* is favorably associated with feed efficiency in Nelore animals.

Table 2.3. Summary of the significant association of target genes' expression levels to feed efficiency-related traits.

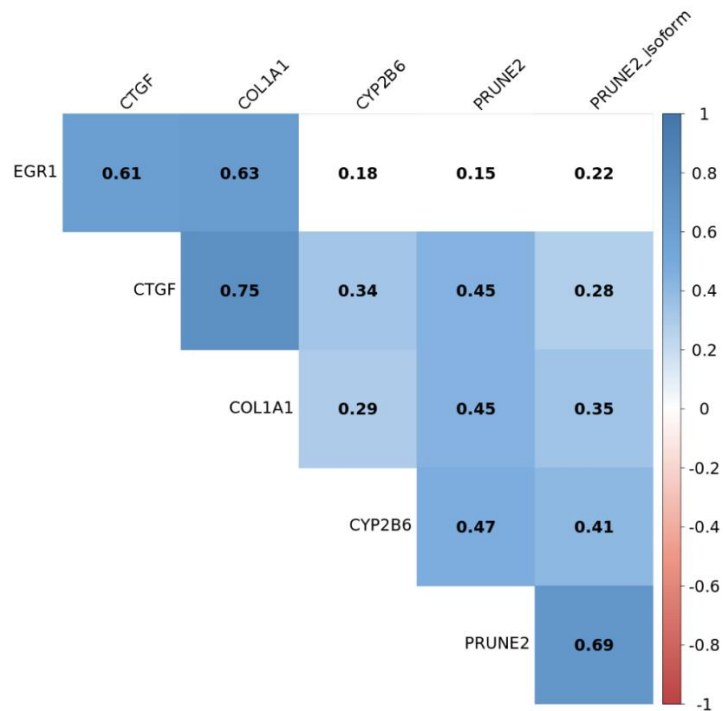
Traits	Genes	Estimated effect ^a ± SE	p-value
Dry matter intake (DMI)	<i>PRUNE2</i>	-0.3194 ± 0.1386	0.0259
Feed conversion ratio (FCR)	<i>PRUNE2</i>	-0.2427 ± 0.1017	0.0214
	<i>PRUNE2_isoform</i>	0.2436 ± 0.08551	0.0067
Feed efficiency ratio (FE)	<i>PRUNE2</i>	0.008361 ± 0.003827	0.0343
	<i>PRUNE2_isoform</i>	-0.00756 ± 0.003217	0.0233
Residual feed intake (RFI)	<i>PRUNE2</i>	-0.2557 ± 0,07907	0.0024

^ainversely proportional RNA quantification; SE: standard error.

PRUNE2 gene expression presented a significant correlation (0.69) with *PRUNE2_isoform* (Figure 2.1). Further, *PRUNE2* gene showed a moderate correlation with other genes (*COL1A1*, *CTGF*, and *CYP2B6*). However, the *PRUNE2_isoform* showed a low correlation with *COL1A1* and *CTGF* (Figure 2.1).

The global sequence alignment between the two *PRUNE2* isoforms indicated that they have a 95.64 % similarity. Furthermore, the total number of amino acid residues between the predicted protein products of the two isoforms were different, where *PRUNE2_201* has 323 residues, and *PRUNE2_isoform* has 294 residues.

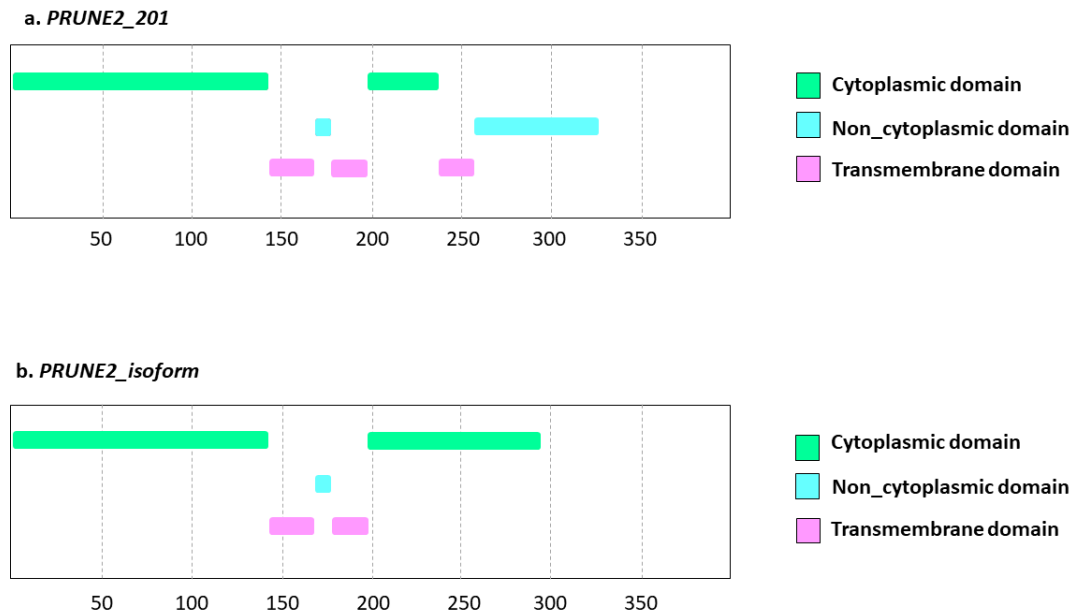
Figure 2.1. Pearson’s correlation matrix among the adjusted expression values (CtA) for all the target genes.



The colors and intensity indicate the strength and direction of the significant ($P < 0.05$) relationships; blue color indicates the positive correlation and red color indicates the negative correlation. White cells indicate non-significant correlations.

The functional domain annotation distinguished the isoforms’ products in one of the extremes of the proteins (Figure 2.2), the new isoform having a cytoplasmic terminal domain, in contrast to the *PRUNE2_201*, in which the terminal domain is non-cytoplasmatic.

Figure 2.2. mRNA sequence alignment and functional protein domain annotation for PRUNE2 gene isoforms.



(a) protein sequence described for bovine in the database for PRUNE2_201, (b) protein sequence identified for PRUNE2_isoform by RNA-Seq in liver tissue.

2.4. DISCUSSION

Herein, we investigated genes whose expression levels in Nelore liver were previously associated with feeding efficiency by a disruptive sampling approach (Tizioto et al., 2015). From the five genes (*CYP2B6*, *EGR1*, *COL1A1*, *CTGF*, and *PRUNE2*) that were found as differentially expressed between divergent phenotypes in RNA-Seq data of the same population studied in the present study (TIZIOTO et al., 2015; TIZIOTO et al., 2016), only the *PRUNE2* gene and its isoform were significantly associated to feed efficiency related-traits in the continuous sampling of the same population.

To better understand the significant RT-qPCR – phenotypes relationships observed herein, it is important to emphasize that Ct values are inversely proportional to target mRNA abundance. Thus, from the negative relation between *PRUNE2* Ct values and DMI, FCR and RFI, as well as the positive relationship between *PRUNE2* Ct values and FE, we can conclude that increasing the gene expression of *PRUNE2* is unfavorable for feed efficiency in Nelore animals. According to Anuppalle et al. (2017), Li et al. (2011), Li et al. (2012) and Potkin et al. (2009), this gene was implicated in the apoptosis pathway, exerting a proapoptotic function. This pathway was previously related to feed efficiency in *Bos taurus*

cattle (KHANSEFID et al., 2017). Also, a global view of gene expression differences between low and high RFI in chicken suggests that RFI can be explained by different metabolic pathways, including apoptosis (LEE et al., 2015).

Apoptosis is a process of cell death through the activation of genes necessary for cell destruction (DU et al., 2017; WIBLE; BRATTON, 2018). In the liver, it can be induced by oxidative stress, and contributing to mitochondrial dysfunction (DU et al., 2017; KUJOTH et al., 2005). A previous study performed by our research group indicated that oxidative stress is increased in inefficient animals and may be related to differences in mitochondrial function (TIZIOTO et al., 2015), thus suggesting that *PRUNE2* expression in the liver might be an indicator of oxidative stress-induced apoptosis in the liver.

Furthermore, a mitochondrial function has been indicated as a major factor that influences RFI (ZULKIFLI et al., 2007). In this respect, Fonseca et al. (2015) identified genes involved in this function and related to feed efficiency in Nelore cattle. Moreover, Casal et al. (2018) indicated that the highly efficient animals had a greater efficiency in hepatic nutrient metabolism, which was associated with superior hepatic mitochondrial function in *Bos taurus* cattle.

Tatsumi et al. (2015) suggest that *PRUNE2* interacts with BCL2 and inhibits the AKT phosphorylation, which can negatively affect the cell survival signaling.

As another critical pathway, PI3K/AKT is related to signaling cell survival and proliferation (FULDA, 2013), and is involved in insulin action on glucose metabolism (LU et al., 2017). Also, this pathway regulates changes involved in apoptosis and plays an antiapoptotic role in liver regeneration (VALIZADEH et al., 2019). Additionally, members of the PI3K may activate AKT (HUNG et al., 2018), and its interaction with AKT results in phosphorylation (FULDA, 2013). Thus, PI3K signaling regulates the apoptosis through the AKT activities, where AKT regulates the pro- and antiapoptotic factors, members of BCL2 superfamily (STILES et al., 2009). This family not only controls the apoptosis pathway but also regulates the mitochondrial membrane permeabilization (FULDA, 2013; STILES et al., 2009). According to Zhou et al. (2015), the up-regulated members of the PI3K complex are predicted to increase PI3K/AKT cascade activity in chicken with high feed efficiency. In this respect, considering the results of the present study one can postulate that up-regulated *PRUNE2* gene expression may result in activated BCL2 superfamily, which possibly inhibits the action of AKT and its interaction with PI3K, decreasing liver cell survival and regeneration under oxidative stress, thus resulting in lower feed efficiency.

However, studies associating *PRUNE2* gene with these phenotypic traits are sparse in the literature.

Interestingly, the isolated expression of *PRUNE2_isoform* had favorable effects on two feed efficiency related traits (FE and FCR), thus showing an antagonistic effect when compared to the total gene expression. In the Ensembl database for bovine, only one *PRUNE2* gene isoform (*PRUNE2_201*) was described (ZERBINO et al., 2018). Furthermore, four isoforms of Prune homolog two products have been reported in the human genome (LI et al., 2012; MACHIDA et al., 2006). To our knowledge, this is the first study to report a second *PRUNE2* isoform in the bovine genome and relate this gene with production traits as FE and FCR in cattle.

Although the two isoforms showed a relevant difference in the total number of amino acids residues, being the *PRUNE2_isoform* the smallest one, with 294 amino acid residues, it was revealed that both isoforms *PRUNE2_201* and *PRUNE2_isoform* have the same BNIP2 domain, which was related to the apoptosis pathway in humans (LI et al., 2011). The main predicted structural differences between this new isoform and the *PRUNE2_201*, were both the number of amino acids residues and the presence of a second cytoplasmatic domain in the one of the terminal protein, suggesting that these may have different functions in the cell.

These findings indicated that it is possible that the different transcripts of this gene may play different functions in the cell and could be contributing to the feed efficiency variation. Further, in vitro experiments need to be performed to investigate the functions of this isoform.

Thus, this study validated the association between *PRUNE2* gene expression and feed efficiency in Nelore cattle, reported in Tizioto et al. (2015), estimating a linear function between the level of gene expression and the traits. We also confirmed by RT-qPCR the expression of a new *PRUNE2* isoform, previously identified in the same liver RNA-Seq data, showing an antagonistic expression effect in feed efficiency related traits when compared to the total expression of *PRUNE2* gene. However, studies in different populations are important to confirm this hypothesis.

2.5. CONCLUSION

In this study, we verified the unfavorable effect of *PRUNE2* gene expression on DMI, FCR, FE, and RFI. Furthermore, a new isoform of this gene was confirmed in the liver, and showed a positive effect on FCR and FE traits.

2.6. REFERENCES

ABIEC -Associação Brasileira das Indústrias Exportadoras de Carne: **Exportação Brasileira de carne bovina**. Brazilian livestock profile: Annual Report. Disponível em: <<http://www.brazilianbeef.org.br/download/sumarioingles2018.pdf>>. Acesso em: mar. 2019.

ALEXANDRE, P. A. et al. Liver transcriptomic networks reveal the main biological processes associated with feed efficiency in beef cattle. **BMC Genomics**, v. 16, n. 1073, p. 1-13, 2015.

ANUPPALLE, M. et al. Expression patterns of *prune2* is regulated by a notch and retinoic acid signaling pathways in the zebrafish embryogenesis. **Gene Expression Patterns**, v. 23–24, p. 45-51, Jan. 2017.

BASARAB, J. A. et al. Residual feed intake and body composition in young growing cattle. **Canadian Journal of Animal Science**, v. 83, n. 2, p. 189-204, 2003

BUSTIN, S. A. et al. The miqe guidelines: minimum information for publication of quantitative real-time PCR experiments. **Clinical Chemistry**, v.55, n. 4, p. 611-622, 2009.

CASAL, A. et al. Hepatic mitochondrial function in Hereford steers with divergent residual feed intake phenotypes. **Journal of Animal Science**, v. 96, n. 10, p. 4431–4443, 2018.

CLEMMONS, David R. Role of IGF-1 in skeletal muscle mass maintenance. **Trends in Endocrinology & Metabolism**, v. 20, n. 7, p. 349-356, 2019.

CONNOR, E. E. Invited review: improving feed efficiency in dairy production: challenges and possibilities. **Animal**, v. 9, n. 3, p. 395-408, 2019.

De OLIVEIRA, P. S. et al. Identification of genomic regions associated with feed efficiency in nelore cattle. **BMC Genetics**, v. 15, n. 100, p. 1-10, 2014.

DU, X. et al. Elevated apoptosis in the liver of dairy cows with ketosis. **Cellular Physiology and Biochemistry**, v. 43, n. 2, p. 568–578, 2017.

FINN, R. D. et al. Interpro in 2017—beyond protein family and domain annotations. **Nucleic Acids Research**, v. 45, n. 4, p. D190–D199, 2016.

FONSECA, L. F. S. et al. Expression of genes related to mitochondrial function in Nelore cattle divergently ranked on residual feed intake. **Molecular Biology Reports**, v. 42, n. 2, p. 559–565, 2015.

- FULDA, S. Modulation of mitochondrial apoptosis by PI3K inhibitors. **Mitochondrion**, v. 13, n. 3, p. 195-198, 2013.
- HILL, R. A. **Feed efficiency in the beef industry**. 1 ed. Ames, IA: Wiley-Blackwell, 2012. 328p.
- HUANG, W. et al. global transcriptome analysis identifies differentially expressed genes related to lipid metabolism in wagyu and Holstein cattle. **Scientific Reports**, v. 7, p. 1-11, 2017.
- HUANG, X. et al. The PI3K/AKT pathway in obesity and type 2 diabetes. **International Journal of Biological Sciences**, v. 14, n. 11, p. 1483-1496, 2018.
- KERN, R. J. et al. Transcriptome differences in the rumen of beef steers with variation in feed intake and gain. **Gene**, v. 586, n. 1, p. 12-26, 2016.
- KHANSEFID, M. et al. Gene expression analysis of blood, liver, and muscle in cattle divergently selected for high and low residual feed intake. **Journal of Animal Science**, v. 95, n. 11, p. 4764-4775, 2017.
- KHIAOSA-ARD, R.; ZEBEL, Q. Cattle's variation in rumen ecology and metabolism and its contributions to feed efficiency. **Livestock Science**, v. 162, p. 66-75, 2014.
- KONG, R. S. G. et al. Transcriptome profiling of the rumen epithelium of beef cattle differing in residual feed intake. **BMC Genomics**, v. 17, n. 592, p. 1-16, 2016.
- KUJOTH, G. C. et al. Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. **Science**, v. 309, n. 5733, p. 481-484, 2005.
- LEE, J. et al. Transcriptomic analysis to elucidate the molecular mechanisms that underlie feed efficiency in meat-type chickens. **Molecular Genetics and Genomics**, v. 290, n. 5, p. 1673-1682, 2015.
- LI, Jun; DING, Yuemin; LI, Aiqing. Identification of COL1A1 and COL1A2 as candidate prognostic factors in gastric cancer. **World Journal of Surgical Oncology**, v. 14, n. 297, p. 1-5, 201, 2016.
- LI, S. et al. The expression and localization of prune2 mRNA in the central nervous system. **Neuroscience Letters**, v. 503, n. 3, p. 208-214, 2011.
- LI, S. et al. Olfaxin as a novel prune2 isoform predominantly expressed in the olfactory system. **Brain Research**, v. 1488, n. 7, p. 1-13, 2012.
- LIPSON, K. E. et al. CTGF is a central mediator of tissue remodeling and fibrosis, and its inhibition can reverse the process of fibrosis. **Fibrogenesis & Tissue Repair**, v. 5, n. 24, p. 1-8, 2012.
- LU, P. et. al. Insulin upregulates betatrophin expression via PI3K/AKT pathway. **Scientific Reports**, v. 7, n. 5594, p. 1-9, 2017.

MACHIDA, T. et al. Increased expression of proapoptotic *bmcc1*, a novel gene with the *bnip2* and *cdc42gap* homology (*bch*) domain, is associated with favorable prognosis in human neuroblastomas. **Oncogene**, v. 25, p. 1931–1942, 2006.

MARCHLER-BAUER, Aron. Cdd/sparcle: Functional classification of proteins via subfamily domain architectures. **Nucleic Acids Research**, v. 45, p. D200–D203, 2016.

MONTAÑO-BERMUDEZ, M.; NIELSEN, M. K; DEUTSCHER, G. H. Energy requirements for maintenance of crossbred beef cattle with different genetic potential for milk. **Jornal of Animal Science**, v. 68, n. 8, p. 2279-88, 1990.

MOORE, S. S.; MUJIBI, F. D.; SHERMAN, E. L. Molecular basis for residual feed intake in beef cattle. **Journal of Animal Science**, v. 87, p. E41–E47, 2009.

NASCIMENTO M. L. et al. Feed efficiency indexes and their relationships with carcass, non-carcass, and meat quality traits in Nellore steers. **Meat Science**, v. 116, p. 78-85, 2016.

PAGEL J. I.; DEINDL E. Early growth response 1- a transcription factor in the crossfire of signal transduction cascades. **Indian Journal of Biochemistry & Biophysics**, v. 48, n. 4, p. 226-35, 201, 2011.

POTKIN, Steven G. et al. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. **Plos One**, v. 4, n. 8, p. 1-15, 2009.

RUIJTER, J. M. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. **Nucleic Acids Research**, v. 37, n. 6, p. 45, 2009.

SALLEH, M. S. et al. 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**, v. 18, n. 258, p. 1-17, 2017.

SATTERTHWAITE, F. E. An approximate distribution of estimates of variance components. **Biometrics Bulletin**, v. 2, n. 6, p. 110-114, 194, 1946.

SEO, Minseok. Rna-seq analysis for detecting quantitative trait-associated genes. Scientific reports. **Scientific Reports**, v. 6, n. 24375, p. 1-12, 2016.

STILES, B. L. PI-3-K, and AKT: onto the mitochondria. **Advanced drug delivery reviews**, v.61, n. 14, p. 1276-1282, nov.2009.

SUN, H. et al. Landscape of multi-tissue global gene expression reveals the regulatory signatures of feed efficiency in beef cattle. **Bioinformatics**, 2018.

TATSUMI, Y. *Bmcc1*, which is an interacting partner of *BLC2*, attenuates AKT activity, accompanied by apoptosis. **Cell Death & Disease**, v. 6, p. e1607, Jan. 2015.

TIZIOTO, P. C. et al. Identification of *KCNJ11* as a functional candidate gene for bovine meat tenderness. **Physiological Genomics**, v. 45, n. 24, p. 12, 2013.

TIZIOTO, P. C. et al. Global liver gene expression differences in nelore steers with divergent residual feed intake phenotypes. **BMC Genomics**, v. 16, n. 242, p. 1-14, 2015.

TIZIOTO, P. C. et al. Gene expression differences in longissimus muscle of nelore steers genetically divergent for residual feed intake. **Scientific Reports**, v. 6, n. 39493, p. 1-12, 2016.

UNITED STATES DEPARTMENT OF AGRICULTURE. LIVESTOCK AND POULTRY: **WORLD MARKETS AND TRADE**. Disponível em: <http://apps.fas.usda.gov/psdonline/circulars/livestock_poultry.pdf>. Acesso em: 28 mar. 2019.

VALIZADEH, A. et al. The roles of signaling pathways in liver repair and regeneration. **Journal of Cellular Physiology**, p. 1-9, 2019.

VELAZCO, J. I. et al. Daily methane emissions and emission intensity of grazing beef cattle genetically divergent for residual feed intake. **Animal Production Science**, v. 57, n. 4, p. 627-635, 2016.

WEBER, Kristina L. et al. Identification of gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-seq. **Plos One**, v. 11, n. 3, p. e015227, 2016.

WEI, T.; SIMKO, V. R package "corrplot": **Visualization of a Correlation Matrix** (Version 0.84). Available from <https://github.com/taiyun/corrplot>. 2017

YE, Ji. Et al. Blast: a tool to design target-specific primers for a polymerase chain reaction. **BMC Bioinformatics**, v. 13, n. 134, p. 2-11, 2012.

ZAREK, C. M. et al. Differential expression of genes related to gain and intake in the liver of beef cattle. **BMC Research Notes**, v. 10, n. 1, p. 1-8, 2017.

Zerbino D. R. et al. ensembl 2018. **Nucleic Acids Research**, v. 46, n. 1, p. D754–D761, 2017.

ZHOU, N.; LEE, W. R; ABASHT, B. Messenger RNA sequencing, and pathway analysis provide novel insights into the biological basis of chickens' feed efficiency. **BMC Genomics**, v. 16, n. 195, p. 1-20, 2015.

Zulkifli M. N. et al. Cattle residual feed intake candidate genes. **J Anim Breed Genetics**, v. 18, n. 1, p. 668-671, 2007.

Chapter 3 Potential biomarkers for feed efficiency related-traits in Nelore cattle identified by co-expression network

ABSTRACT

Feed efficiency is an important economic trait for the beef production system, and its improvement can contribute to reducing the environmental impacts, the feeding costs, and to increasing carcass quality. This study aimed to identify potential biomarkers for feed efficiency related-traits in Nelore cattle based on a co-expression approach, based on the information from 180 RNA-Seq samples from muscle (*Longissimus thoracis*) tissue. We identified 391 potential biomarkers (hub genes) associated to feed efficiency related-traits. These hub genes were acting in pathways related to protein synthesis, muscle growth, and immune response processes. The gene expression quantitative variation of *CCDC80*, *FBLN5*, *SERPINF1* and *OGN* was significantly associated (q-value < 0.05) with the feed efficiency-related traits average daily gain (ADG), feed conversion ratio (FCR), feed efficiency ratio (FE), Kleiber index (KI), and relative growth ratio (RGR) under a linear model. These genes are related to glucose homeostasis, osteoblastogenesis, and oxidative stress pathways. Moreover, within the hub genes, we identified six potential gene regulators through enrichment analysis for transcription factor binding sites (TFBS). Among them, is the transcription factor *TCF4*, which can regulate genes previously associated with feed efficiency, and that may have a role in muscle growth. Finally, we integrated the hub genes with previous studies performed by our research group and identified two potential cis-regulatory elements that could regulate two hub genes identified in this study. These are potential functional variants related to feed efficiency variation.

Keywords: Nelore, Feed efficiency, *Bos indicus*, SNPs, WGCNA.

3.1. INTRODUCTION

Animal feeding is related to the energy of maintenance (ARTHUR; HERD, 2008) and represents up to 60% of the total production costs (CONNOR, 2015; MONTAÑO-BERMUDEZ et al., 1990). Also, feed efficient animals can contribute to reduced environmental impact due to greenhouse gases emissions and to the use of natural resources for meat production (BASARAB et al., 2003; KHIAOSA-ARD; ZEBELI, 2014; VELAZCO, 2016). Thus, improving feed efficiency may contribute to increasing sustainability and profitability in the beef cattle industry (BASARAB et al., 2003). However, feed efficiency-related traits are expensive to measure and be adopted in breeding programs, therefore justifying the search for biomarkers.

In this sense, phenotypic and genomic variation for feed efficiency were studied in the same population of Nelore cattle adopted here (DE OLIVEIRA et al., 2014, NASCIMENTO et al., 2016, TIZIOTO et al., 2015, TIZIOTO et al., 2016). The first reported genomic regions associated with feed efficiency related-traits and described genomic heritability ranging from 0.18 to 0.57. Also, it identified candidate genes on BTA 24, associated with dry matter intake (DMI) and residual feed intake (RFI). The Histamine receptor H4 (*HRH4*) gene, located in this region, is related to inflammation and immunity, gastric acid secretion, food allergies, appetite regulation and metabolism in humans. Another example is the quantitative trait loci (QTLs) associated with RFI and DMI on BTA 7, related to proteolysis, an essential biological process in feed efficiency (DE OLIVEIRA et al., 2014).

Other studies highlighted the oxidative stress as pivotal for feed efficiency (TIZIOTO et al., 2015; TIZIOTO et al., 2016). In these studies, the authors indicated that the Early growth response 1 (*EGR1*), a transcription factor (TF) that interacted with several other DE genes related to oxidative stress, was differentially expressed (DE) between divergent residual feed intake group, thus indicating this TF as a candidate regulator in this biological process.

The number of genes and biological processes described as associated to feed efficiency so far are in agreement with the polygenic nature and biological complexity of these traits. Consequently, more investigations are needed in order to understand their complexity.

Nowadays, integrative tools contribute to elucidate the biological mechanisms related to phenotypic variation. In this context, integrative analysis, as co-expression networks, allows the investigation of expression pattern similarity in thousands of genes (SERIN et al., 2016). This information can be used to explain biological processes related to phenotypes.

Gene co-expression network has been helpful to identify highly connected genes (hub genes) (DAM et al., 2018), which may play essential functions in biological processes (GUILIETTI et al., 2018) and has been fruitful to dissect complex phenotypes (ALEXANDRE et al., 2015; DE OLIVEIRA et al., 2018; DINIZ ET al., 2019; FONSECA et al., 2018; KONG et al., 2016; SALLEH et al., 2018; WEBER et al., 2016). In some of these studies, the immune response pathway was associated with feed efficiency (ALEXANDRE et al., 2015; DE OLIVEIRA et al., 2018; WEBER et al., 2016;). Also, the Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit (*EZH2*), and Dolichol Kinase (*DOLK*), which are related to inflammation response (ALEXANDRE et al., 2015) were described as regulators of feed efficiency.

Although shading light to pathways related to feed efficiency, these co-expression studies were performed comparing contrasting phenotypic groups for this trait. Thus, it is important to explore regulators and essential genes (hub genes) in biological processes and metabolic pathways contributing to the continuous feed efficiency variation, as this will be the main substrate of genomic selection. Therefore, we aimed to identify potential biomarkers (hub genes) based on population level co-expression network for feed efficiency related-traits, as well as predicting functional variants that could explain the traits' variation.

3.2. MATERIALS AND METHODS

3.2.1. Animals and phenotypic traits

The experimental procedures were conducted in accordance to Institutional Animal Care and Use Committee Guidelines of the Brazilian Agricultural Research Corporation – EMBRAPA (CEUA Protocol 01/2013).

The expression data were collected from 192 Nelore steers as previously described elsewhere (CESAR et al., 2018). The phenotypes collection were described on Nascimento et al. (2016) and de Oliveira et al., (2014). Briefly, the experiment lasted three years, with animals born in 2007, 2008 and 2009 at two different feedlot places, Embrapa Southeast Livestock (São Carlos, SP, Brazil- feedlot 1) and Embrapa Beef Cattle (Campo Grande, MS, Brazil- feedlot 2). The adaptation period was of at least 28 days, and the animals stayed on trial for at least 70 days under the same formulated diet, which contained 40% silage and 60% concentrate, twice a day.

The animals were evaluated for growth and feed efficiency-related traits as described by de Oliveira et al. (2014), including average daily gain (ADG, kg/d), body weight (BW,

kg), dry matter intake (DMI, kg/d), feed conversion ratio (FCR, feed intake/gain; kg/kg), feed efficiency ratio (FE, gain/feed intake; kg/kg), Kleiber index (KI, ADG/MBW; kg/kg), metabolic body weight (MBW, kg), residual feed intake (RFI, Kg/d), and relative growth rate (RGR, %/d).

3.2.2. Library preparation, RNA sequencing, and data processing

The RNA-sequencing, data quality control, alignment, and quantification were previously described in Cesar et al. (2018). In summary, total mRNA was extracted using the Trizol reagent (Life Technologies, Carlsbad, CA) from *Longissimus thoracis* muscle of 192 steers were collected at slaughter. The mRNA integrity was estimated using the Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA). A total of 2µg of each RNA sample was used for library preparation according to the TruSeq RNA Sample Preparation kit v2 guide (Illumina, San Diego, CA). Paired-end (PE) sequencing was performed using the Illumina HiSeq 2500 platform. Samples were multiplexed with unique six-mer barcodes and ran on multiple lanes to obtain 2 x 100 bp reads.

Data quality control (QC) of raw reads was performed by FastQC v.0.11.2 (ANDREWS, 2015) and MultiQc v.1.4 (EWELS et al., 2016). The paired-end (PE) reads were filtered using the Seqclean v1.4.13 package (<https://bitbucket.org/izhbannikov/seqclean>), which removed all the reads with a mean quality under 24 Phred score. Additionally, reads shorter than 65 base pairs (bp), as well as primers and vector contaminants, were removed using the UniVec database (<http://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/>).

To obtain the normalized expression values (TPM- Transcripts Per Kilobase Million), alignment and quantification were carried out using RSEM (RNA-Seq by Expectation Maximization) based on the *Bos taurus* genome assembly (UMD3.1).

3.2.3. Co-expression Network and module trait association (MTA) analyses

We applied a co-expression network approach based on WGCNA framework as previously described by Langfelder; Horvath (2008). To this end, we carried out a quality control step filtering out samples with missing phenotypes for all traits and genes with low expression. The TPM values were logarithmic transformed (\log_2 (TPM+1)) and a linear model was fitted to adjust the gene expression data (TPM values) for batch effect (lane effect) by applying Limma R-package v.3.36.2 (RITCHIE et al., 2015).

From the adjusted expression data, we constructed a signed co-expression network based on Spearman's correlation. Considering the scale-free topology criteria, we chose a soft thresholding power ($\beta = 16$). Clustering genes into modules were carried out based on the Topological Overlap Measure (TOM) by applying the dynamic tree cut v.1.63.1 package. The module eigengene (ME), the value of the first principal component of each module, was estimated and used to associate the modules to each trait. To this end, we fitted a linear model which included the contemporary group (CG) as a fixed effect and animal's age at slaughter as a covariate. The CG was described in de Oliveira et al. (2014) and included feedlot location, year of the experiment, animal origin, and pen type. The model can be represented as follows:

$$y_{ijk} = \mu + CG_i + A_j + T_k + \varepsilon_{ijk},$$

where:

y_{ijk} are the ME values for the i^{th} animal for the j^{th} module (n =25);

μ is the mean;

CG_i is the fixed effect of the contemporary group;

A_j is the animal's age at slaughter as a covariate;

T_i is the observation for the animal i^{th} for the k^{th} phenotype;

ε_{ijk} is the random residual effect [$\sim N(0, \sigma^2_e)$].

Modules significantly associated ($p \leq 0.05$) were selected for further analyses.

3.2.4. Hub genes identification, functional annotation and pathway over-representation analyses

To identify putative regulatory genes and their effect on feed efficiency related-traits, we selected the genes highly interconnected within the associated modules, as these genes likely have more effect on the phenotypic variation. We selected the hub genes based on the module membership ($MM \geq 0.8$) (LANGFELDER; HORVATH, 2008) and associated them to the traits by applying a linear model.

The hub genes' expression was also adjusted to the CG and animal's age, besides the lane effect. The association analysis between hub genes and traits followed the overmentioned model for MTA. However, the phenotypes were taken as the dependent variable. We considered as significantly associated those genes with a p-value ≤ 0.05 after multiple correction tests by the false discovery ratio (FDR) (BENJAMINI; HOCHBERG (1995) available in R package.

To gain biological insights on the feed efficiency pathways, we carried out an over-representation analysis for the identified hub genes using ClueGO v.2.5.1 (BINDEA et al., 2009). We adopted the Cytoscape software (SHANNON et al., 2003) for network construction and data visualization.

3.2.5. Enrichment analysis to identify TFBS in hub genes

To identify potential regulators (hub-TFs) among the trait-associated hub genes, we carried out a TFBS enrichment analysis using the Rcis Target software (AIBAR et al., 2017) considering a human database. Moreover, we identified the hub genes described as coding transcription factors (TF) based on the TF database for *Bos taurus* (SOUZA et al., 2018). The data visualization was performed on Cytoscape software (SHANNON et al., 2003).

3.2.6. Data integration for identification of regulatory regions and functional variants

Different approaches have been applied in the Nelore population used here to dissect the underlying genes acting on feed efficiency traits. To identify the overlapping genes among these studies and the hub genes described herein, we intersected our hub genes list with those genes identified on muscle differential expression for RFI (TIZIOTO et al., 2016), genome-wide association (GWAS) (DE Oliveira et al., 2014), allelic expression (ASE) (SOUZA, under review), and eQTL analysis (CESAR et al., 2018). Genes harbored in the QTL regions reported by de Oliveira et al. (2014) were retrieved using Biomart version 3.5 (R/Bioconductor package) (DURINCK et al., 2009). For the eQTLs, we considered the list of genes that were regulated by the eQTLs described in Cesar et al. (2018). Cytoscape (SHANNON et al., 2003) was used for data visualization.

From the list of cis-eQTL that were associated to variation in hub genes expression, we predicted the presence of TFBS within a 51 bp window around the eQTL SNP, according to the sequence on the Ensembl database, using the online software LASAGNA (LEE; HUANG et al., 2013).

3.3. RESULTS

Herein, by analyzing the gene co-expression pattern in *Longissimus thoracis* and integrating structural and functional data from previous experiments, according to the following steps: i) co-expression network construction and trait association; ii) hub genes association to feed efficiency related-traits; iii) hub genes enrichment analysis to identify

TFBS; and iv) Hub genes, QTL, DE genes, ASE and eQTL data integration for the identification of regulatory regions and potential variants. We identified potential central genetic elements underlying feed efficiency-related traits in Nelore cattle .

3.3.1. Background of animals and expression data

After data quality control from the *Longissimus thoracis* muscle expression, we kept 180 samples out of 192 and 8,622 genes out of 24,000 with at least two counts in 90% of the Nelore steers samples. The descriptive statistics for the feed efficiency related-traits in this sample are described in Table 3.1.

Table 3.1. Mean, standard deviation (SD), minimum (Min) and maximum (Max) observed for feed efficiency related-traits in Nelore steers.

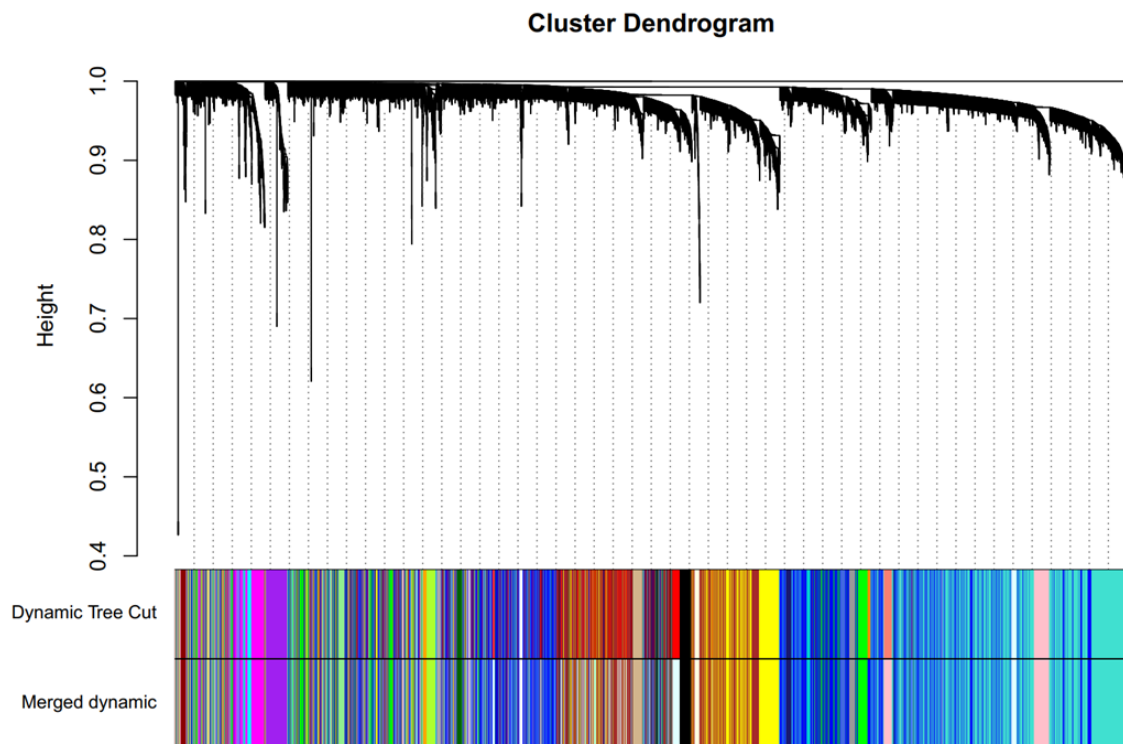
Traits	Mean \pm SD	Min	Max
Average daily gain (ADG, kg/d)	1.42 \pm 0.29	0.80	2.33
Body weight (BW, kg)	382.98 \pm 51	280.5	515.5
Dry matter intake (DMI, kg/d)	8.56 \pm 1.23	5.4	12.18
Feed conversion ratio (FCR, feed intake/gain; kg/kg)	6.24 \pm 1.40	3.74	11.5
Feed efficiency ratio (FE, gain/feed intake; kg/kg)	0.17 \pm 0.035	0.087	0.27
Kleiber index (KI, ADG/MBW; kg/kg)	0.016 \pm 0.004	0.008	0.027
Metabolic body weight (MBW, kg)	86.43 \pm 8.63	68.54	108.19
Residual feed intake (RFI, Kg/d)	-0.028 \pm 0.67	-1.71	1.8
Relative growth rate (RGR, %/d)	0.16 \pm 0.04	0.092	0.285

3.3.2. Weighted gene co-expression network analysis (WGCNA) and module trait association (MTA)

We used the WGCNA framework to carry out a co-expression network analysis based on the muscle expression of the 8,622 genes. We constructed a signed network using a soft thresholding power ($\beta = 16$) to reach the scale-free topology criteria. We identified 25 modules, including the MEGrey (Figure 3.1.). The estimated module eigengene (ME) was able to explain between 34% (MEblue) and 67% (MEgrey60) of the expression variation (Table-S2).

A total of six modules were selected for further analysis based on the module trait association (MTA, $P\text{-value} \leq 0.05$) (Figure 3.2). The Mebrown and Metan gathered the highest (1,125) and smallest (92) number of genes, respectively. Regarding the MTA, the Mebrown and Meyellow showed the highest number of associations, which includes ADG, FCR, FE, KI, and RGR. Nonetheless, the FCR and RGR had the greater number of associated modules, four and five, respectively (Table 3.2).

Figure 3.1. Clustering dendrogram of genes in modules. Each color represents one module eigengene (ME).



We identified a positive MTA for ADG, FE, KI, and RGR, whereas FCR and RFI were negatively associated (Figure 3. 2).

3.3.3. Hub genes and their association with feed efficiency

From the six selected modules, that were associated with at least one feed efficiency trait, we identified 391 hub genes based on the module membership ($MM \geq 0.8$). We then evaluated the effect of these hub genes on every phenotype using a linear model analysis between the gene expression and the trait. We identified hub genes associated with ADG (5), FCR (186), FE (147), KI (137), and RGR (278) (adjusted $p\text{-value} \leq 0.05$) Furthermore, the

genes *CCDC80*, *FBLN5*, *SERPINF1* and *OGN* were pointed out as simultaneously impacting all the traits (Figure 3.3).

Table 3.2. Description of significantly feed efficiency related-traits associated modules in Nelore cattle.

Module	Traits	Genes *	ME variation (%)
MEblack	RGR	309	52.9
MEbrown	FCR, FE, ADG, KI, RGR	1125	34.8
MElightcyan	FCR, KI, RGR	507	40.4
MEpurple	RFI	224	35.1
MEtan	FCR, FE, RGR	92	62.8
MEyellow	FCR, FE, ADG, KI, RGR	732	39.6

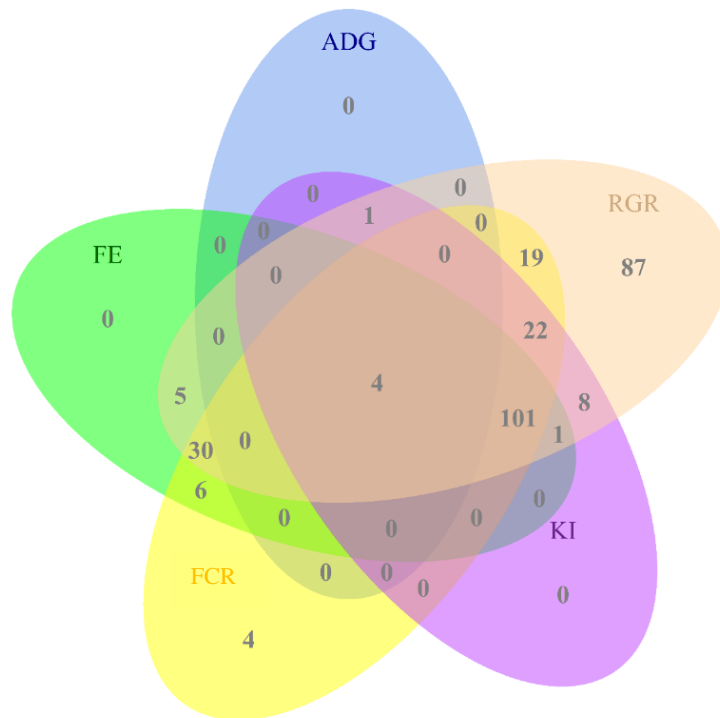
*Genes clustered into the module; ME: module eigengene; ME variation: expression variation explained by the module eigengene.

Figure 3.2. Linear module trait association for all FE related-traits. Beta coefficients are displayed on top and p-values in parenthesis.

MTA- Module trait association

Module	BW	MBW	ADG	DMI	FCR	FE	KI	RGR	RFI
MEdarkred	-1e-05 (1)	-9e-05 (0.9)	-0.02 (0.3)	1e-04 (1)	0.006 (0.3)	-0.2 (0.3)	-2 (0.3)	-0.3 (0.1)	0.005 (0.6)
MElightgreen	2e-05 (0.9)	1e-04 (0.9)	-3e-04 (1)	0.004 (0.5)	0.002 (0.7)	-0.1 (0.6)	-0.3 (0.9)	-0.08 (0.7)	0.008 (0.3)
MEcyan	-1e-04 (0.5)	-7e-04 (0.5)	-4e-05 (1)	-4e-04 (0.9)	-6e-05 (0.9)	-0.03 (0.2)	0.2 (0.9)	0.1 (0.5)	0.004 (0.6)
MEdarkgreen	-1e-04 (0.4)	-9e-04 (0.4)	0.007 (0.8)	0.002 (0.7)	-0.002 (0.7)	-0.06 (0.8)	0.9 (0.7)	0.2 (0.4)	0.01 (0.2)
MEpink	-7e-05 (0.7)	-5e-04 (0.6)	-0.002 (0.9)	0.004 (0.4)	0.003 (0.6)	-0.2 (0.4)	0.1 (0.9)	0.1 (0.6)	0.01 (0.1)
MEturquoise	-1e-04 (0.5)	-6e-04 (0.5)	0.01 (0.7)	0.002 (0.7)	-0.001 (0.9)	9e-04 (1)	2 (0.5)	0.3 (0.1)	0.009 (0.3)
MEdarkturquoise	-5e-05 (0.8)	-3e-04 (0.8)	0.03 (0.2)	0.004 (0.5)	-0.008 (0.1)	0.2 (0.5)	3 (0.2)	0.4 (0.1)	0.007 (0.4)
MEmidnightblue	-8e-05 (0.6)	-5e-04 (0.6)	0.01 (0.3)	0.004 (0.4)	-0.004 (0.4)	0.002 (1)	2 (0.4)	0.2 (0.4)	0.01 (0.2)
MElightyellow	1e-04 (0.4)	8e-04 (0.4)	0.03 (0.6)	0.003 (0.6)	-0.007 (0.2)	0.2 (0.3)	2 (0.3)	0.4 (0.08)	-0.004 (0.6)
MEgreen	3e-04 (0.09)	0.002 (0.09)	0.03 (0.3)	0.01 (0.06)	-0.002 (0.9)	-0.1 (0.6)	0.7 (0.6)	-0.01 (1)	0.008 (0.4)
MEblack	-1e-04 (0.5)	-7e-04 (0.5)	0.03 (0.2)	-7e-04 (0.9)	-0.008 (0.2)	0.3 (0.2)	3 (0.1)	0.5 (0.02)	-7e-04 (0.9)
MElightcyan	-8e-05 (0.6)	-4e-04 (0.6)	0.04 (0.9)	0.001 (0.8)	-0.01 (0.04)	0.3 (0.1)	4 (0.05)	0.5 (0.01)	-4e-04 (1)
MEroyalblue	1e-04 (0.5)	6e-04 (0.5)	0.03 (0.4)	0.005 (0.4)	-0.006 (0.3)	0.2 (0.5)	3 (0.2)	0.3 (0.2)	0.001 (0.9)
MEblue	7e-05 (0.7)	4e-04 (0.7)	0.02 (0.3)	0.006 (0.3)	-0.003 (0.5)	-0.02 (0.9)	2 (0.5)	0.3 (0.3)	0.008 (0.4)
MEgrey60	9e-05 (0.6)	5e-04 (0.6)	0.04 (0.1)	0.006 (0.3)	-0.005 (0.4)	0.1 (0.6)	3 (0.2)	0.4 (0.07)	0.004 (0.6)
MEtan	-2e-04 (0.3)	-9e-04 (0.4)	0.03 (0.5)	-0.003 (0.5)	-0.01 (0.04)	0.4 (0.05)	4 (0.07)	0.5 (0.02)	-0.006 (0.5)
MEbrown	-4e-05 (0.8)	-2e-04 (0.8)	0.05 (0.05)	-0.001 (0.8)	-0.01 (0.01)	0.5 (0.02)	5 (0.03)	0.6 (0.01)	-0.009 (0.3)
MEyellow	-5e-05 (0.8)	-3e-04 (0.8)	0.06 (0.02)	-0.002 (0.7)	-0.02 (0.002)	0.6 (0.005)	6 (0.01)	0.6 (0.005)	-0.01 (0.2)
MEwhite	2e-04 (0.3)	0.001 (0.3)	0.02 (0.3)	0.004 (0.4)	-2e-04 (1)	0.07 (0.7)	1 (0.5)	0.3 (0.2)	-0.003 (0.8)
MEdarkgrey	6e-05 (0.7)	4e-04 (0.7)	0.007 (0.8)	-4e-04 (0.9)	-4e-04 (0.9)	0.05 (0.8)	0.2 (0.9)	0.2 (0.4)	-0.005 (0.6)
MEgreenyellow	4e-05 (0.8)	2e-04 (0.8)	-0.02 (0.5)	0.002 (0.8)	0.007 (0.2)	-0.2 (0.4)	-1 (0.5)	-0.09 (0.7)	0.005 (0.6)
MEorange	-7e-05 (0.7)	-4e-04 (0.4)	0.02 (0.4)	0.003 (0.6)	-2e-04 (1)	0.1 (0.5)	3 (0.2)	0.4 (0.09)	0.007 (0.4)
MEmagenta	3e-05 (0.8)	2e-04 (0.8)	-0.01 (0.7)	-0.002 (1)	-2e-04 (0.8)	-0.07 (0.4)	-2 (0.4)	-0.2 (0.3)	-0.006 (0.5)
MEpurple	-5e-05 (0.8)	-3e-04 (0.8)	3e-04 (1)	-0.007 (0.2)	-0.01 (0.07)	0.3 (0.2)	0.3 (0.9)	-0.06 (0.8)	-0.02 (0.05)
MEgrey	1e-04 (0.4)	9e-04 (0.4)	-0.02 (0.4)	0.003 (0.6)	0.007 (0.2)	-0.3 (0.1)	-3 (0.1)	-0.6 (0.006)	0.004 (0.7)

Figure 3.3. Venn diagram showing the overlapped hub genes with linear expression effect on more than one trait.



3.3.4. Hub genes' pathway over-representation analyses

The pathway over-representation analysis was performed for 391 hub genes identified based on the MM (Table-S3). The main pathways identified are described in Figure 3.4 and include adherens junction, leukocyte transendothelial migration, regulation of actin cytoskeleton, relaxin signaling pathway, platelet activation, PI3K-AKT signaling pathway, sphingolipid signaling pathway, and tight junction. We identified several hub genes involved in more than one pathway (Figure 3.5).

Figure 3.4. Number of hub genes per pathway identified by functional annotation and enrichment analysis performed by ClueGO software.

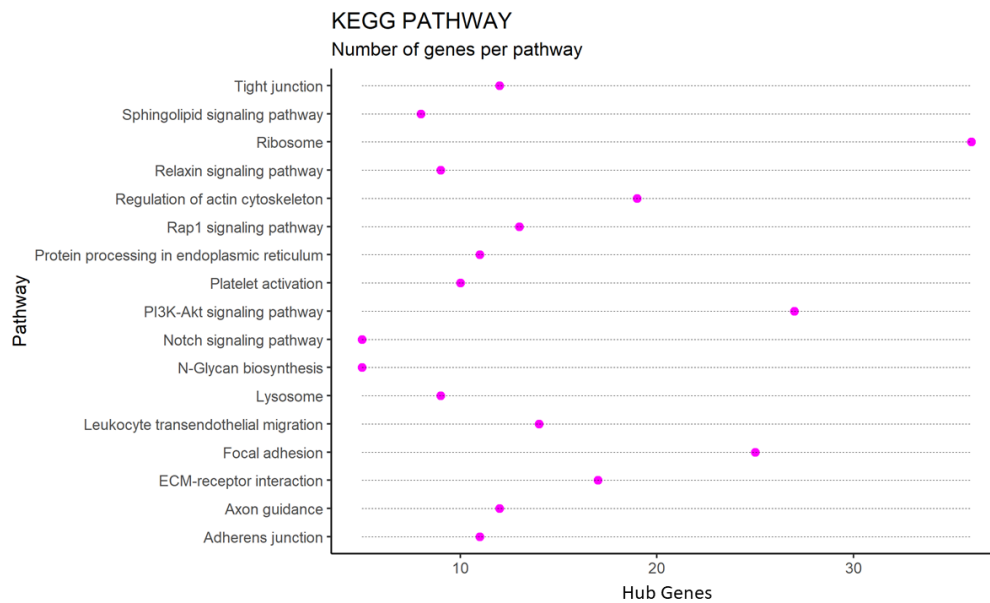
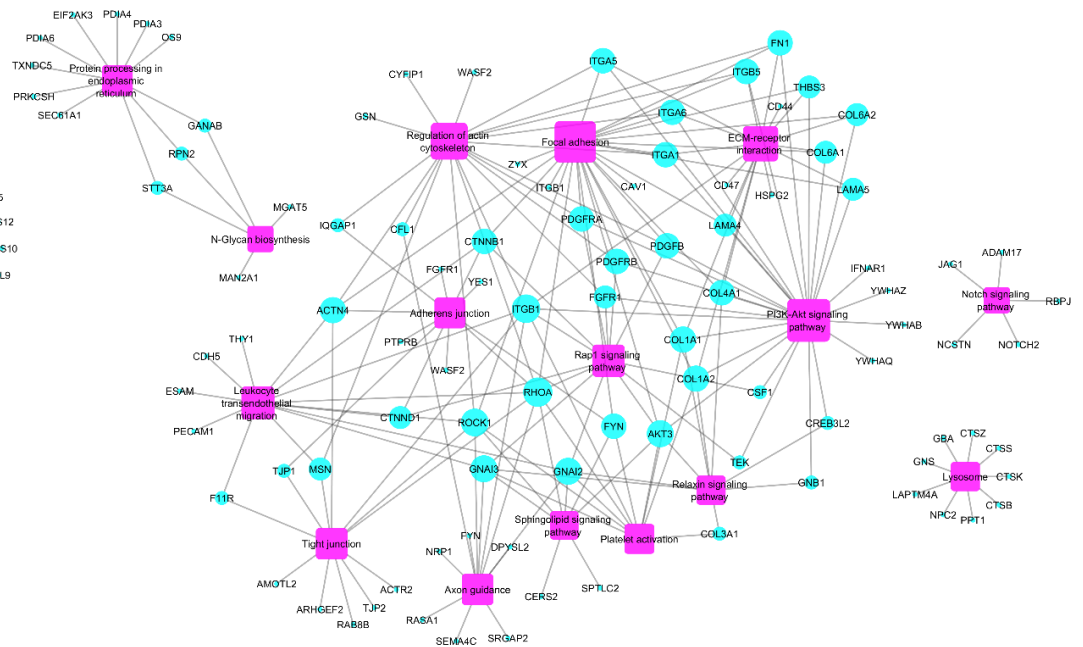


Figure 3.5. Hub genes and the enriched pathways interaction.

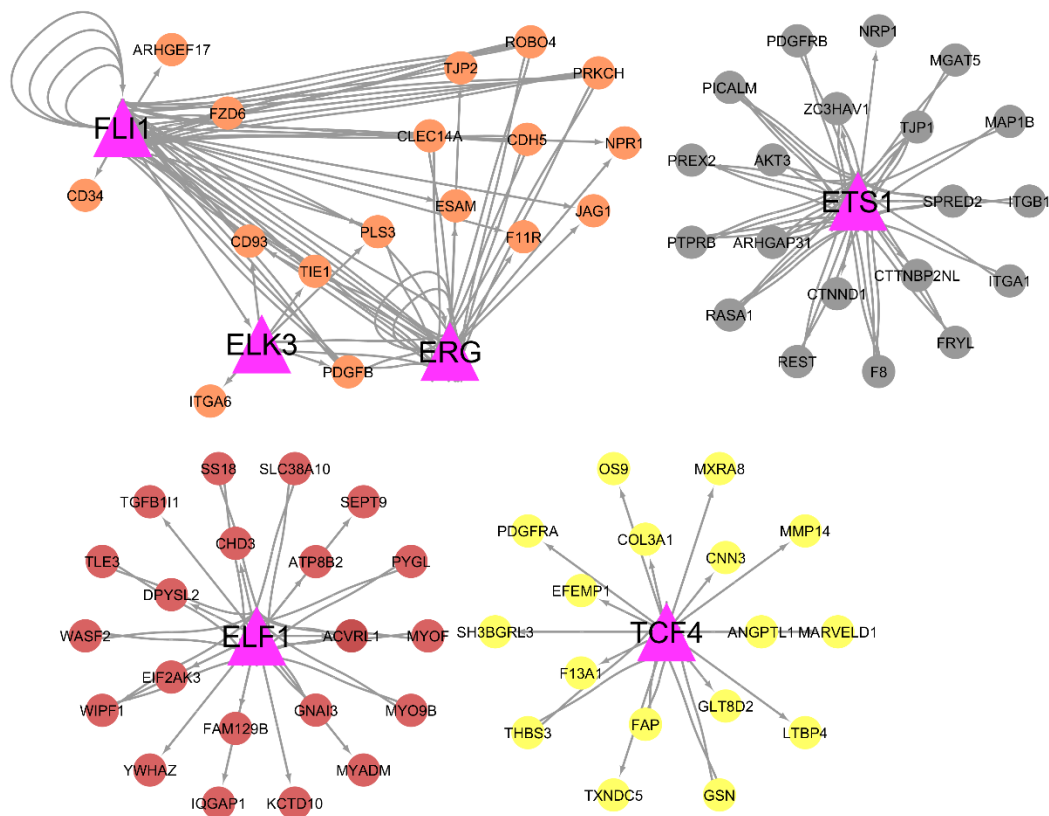


The purple squares are the pathways. The size of the squares represents the connectivity intensity of the nodes (number of hub genes connections with the pathways). The biggest turquoise circle nodes were connected with more pathways.

3.3.5. Enrichment analysis to identify TFBS in hub genes

TFBS enrichment analysis in the hub genes subset showed binding sites for 333 transcription factors (TFs). Furthermore, 13 hub genes out of the 391 pointed out in this study were themselves TFs described for bovine (hub-TFs). From these, six were potential regulators for the set of 391 hub genes, as revealed by the enrichment analysis of TFBS. These TFs were *ELF1*, associated to RGR, *ELK3*, associated to FCR, FE, KI and RGR, *ERG* (did not show linear association with feed efficiency related-traits), *ETS1*, associated to FCR, FE, KI and RGR, *FLI1*, associated to FCR and RGR, and *TCF4*, associated to FE, KI and RGR (Figure 3.6).

Figure 3.6. Integration of the TFs (hub-TFs) with their targets.



The colors of the nodes refer to the hub genes which belong to the module, and the triangles are the hub-TFs.

3.3.6. Data integration for identification of regulatory regions and functional variants

We carried out a data analysis integration of the hub genes list with those previously associated through GWAS or RNAseq with feed efficiency (TIZIOTO et al., 2016, DE OLIVEIRA et al., 2014) and regulatory regions (CESAR et al., 2018; SOUZA et al., 2018; SOUZA et al. under review) identified by our group in the same population evaluated here. We identified several hub genes overlapping with the previously identified genes (Figure 3.7) and a total of six hub genes overlapped in more than one study (*COL4A1*, *EFEMP1*, *EFLI*, *ETSI*, *PECAMI*, *MARVEL1*). Additionally, 26 hub genes are likely affected by eQTL regions, including the *PCDH18* and *SPARCL1*, which were affected by cis-eQTL.

To characterize the potential functional variants in cis-eQTLs related to the hub genes *PCDH18* and *SPARCL1*, we performed a prediction of transcription factor binding site (TFBS) analysis in the flanking sequence of the cis-eQTL and found significant alterations in TFBS for the two cis-eQTL alleles (Table 3.3).

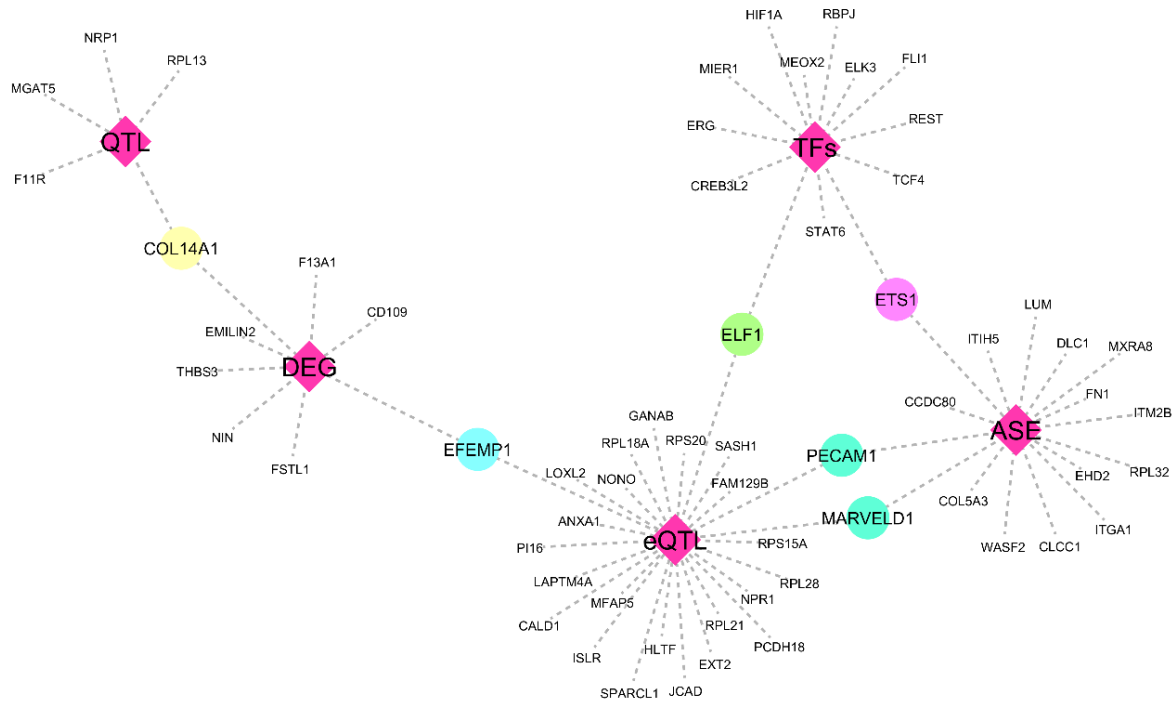
Furthermore, the data integration revealed 16 hub genes that showed allelic-specific expression (ASE), including the genes *ETSI*, *MARVEL1* and *WASF2*, whose ASE-SNP were located in the three prime untranslated regions (3' UTR) (SOUZA et al., under review).

Table 3.3. Description of the putative transcription factors binding sites in the cis-eQTLs for the hub genes *PCDH18* and *SPARCL1*.

Gene	eQTL SNP	Allele	TFs	Sequence ^a	Strand
			<i>NFE2L2</i>		
<i>PCDH18</i>	rs134107947	G/A	(MA0150.1)	GCGACTGAGCa	+
			<i>NR3C1</i>		
<i>SPARCL1</i>	rs43486035	T/G	(MA0113.1)	CAGAGCaAAATGTGCTTG	-

^aTFBS motif sequence found in the 51 bp window surrounding the eQTL SNP.

Figure 3.7. Integrative network of the hub genes identified in the present co-expression network, with regulatory elements (ASE, TFs, eQTLs) and genes previously associated with FE (DEGs and QTLs).



The pink diamonds are the information from previous studies, and the color nodes are the hub genes from the present work that were also identified in more than one study.

3.4. DISCUSSION

We investigated genetic elements with potential central role in the muscle processes related to feed efficiency variation in Nelore steers by applying co-expression and data integration approaches. Our main findings pointed out pathways related to protein synthesis, muscle growth and immune response. Moreover, we identified hub genes associated with feed efficiency, as well as potential functional variants in these genes, and these could be used as biomarkers for feed efficiency related-traits.

3.4.1. Module-trait association analysis

Based on co-expression networks, we identified six clusters of highly co-expressed genes (modules) linearly associated to feed efficiency related-traits (Figure 3.2). Considering that, ADG, KI and FE traits showed a positive association with the modules. In contrast, RFI and FCR, which are inversely related to efficiency (KOCH et al., 1963), showed a negative

association with the modules (Figure 3. 2). These results indicate that modules have a favorable association to feed efficiency variation.

3.4.2. Hub genes identification and their linear expression association with feed efficiency

Hub genes are supposed to have a pivotal role in biological processes related to the phenotype when compared with other genes in the module due to their highest connectivity (DAM et al., 2018). We identified several hub genes into the modules affecting ADG, FCR, FE, KI, and RGR traits. Although the genes gathered into the MEpurple were associated with RFI, the hub genes were not individually associated with RFI (Figure 3.2).

Our analyses revealed hub genes that individually affected feed efficiency variation and showed the same direction of expression association found for their MTA, both favorable for the feed efficiency variation (Figure 3.2).

From the hub genes, four of them were simultaneously associated with ADG, FCR, FE, KI, and RGR. The *CCDC80* has been reported affecting the glucose homeostasis (TREMBLAY et al., 2012), body weight regulation, fat mass, and energy homeostasis (GRILL et al., 2016). Furthermore, this gene was pointed out as a potential negative regulator of adipogenesis, since its dysfunction can lead to excessive body fat (GRILL et al., 2016). In agreement, Tremblay et al. (2009) suggested the *CCDC80* gene as an adipocyte differentiation regulator. Efficient animals are related to improved carcass quality and reduced fat deposition (NASCIMENTO et al., 2014). Based on that, we suggested that this gene might be involved in carcass fat deposition reduction in more efficient animals.

The *FBLN5* gene is a putative antagonist regulator of angiogenesis and inhibits endothelial cells proliferation (ALLAN; SCHIEMAN, 2004). This gene might regulate the production of reactive oxygen species (ROS) in the vascular wall and may protect from endothelial dysfunction and thrombotic responses (SPENCER et al., 2005). Supporting this, Yanagisawa et al. (2009) suggested that, by controlling the excess production of ROS, *FBLN5* gene can control proliferation and death of the endothelial cells. In a previous study, inefficient animals demonstrated overrepresentation of genes related to oxidative stress (TIZIOTO et al., 2016).

The *SERPINF1* gene encodes the pigment epithelium-derived factor (PEDF) (LI et al., 2016). PEDF acts inhibiting endothelial cells and can suppress ROS in the cells by NADPH oxidase-mediated, having potent anti-inflammatory, anti-oxidant, anti-angiogenic, and anti-

thrombotic effects (LIU et al., 2013; LI et al., 2018; MA et al., 2018). Due to their role in the negative regulation of fat mass (MA et al., 2018), *SERPINF1* can contribute to the explanation of the relationship between feed efficiency related-traits and fat mass. Furthermore, Bogan et al. (2013) reported that mouse with non-functional PEDF had the accumulation of unmineralized bone matrix and that this protein might be related to the proper development of bone, controlling the matrix mineral deposition.

The *OGN* gene is considered a putative humoral anabolic bone factor produced on muscle tissue (KAJI et al., 2014). This gene has a role in bone formation by osteoblasts at well-differentiated stage (KEN-ICHIRO et al., 2012) and acts in the balance between osteoblastogenesis and adipogenesis with a positive effect on bone mass (CHEN et al., 2017). Bone weight has been positively correlated with feed efficiency and negatively with subcutaneous and intramuscular fat in cattle (MADER et al., 2008).

3.4.3. Pathway over-representation analysis for the hub genes

3.4.3.1. Protein synthesis metabolism

Animal feed efficiency is directly related to energy usage towards protein turnover and fat metabolism (HERD; ARTHUR, 2009). Among the biological mechanisms acting and regulating these processes, we identified ribosome and PI3K/AKT pathways overrepresented.

The efficiency of the translation process affects protein synthesis rate, and the number of ribosomes is pivotal on that (BUSH et al., 2003; NADER, 2014). Besides to be negatively associated with RFI, we observed that most of the hub genes gathered into this module (ME purple) was also part of the ribosome pathway. Based on the assumption that inefficient animals require more maintenance energy, we suggest that an increased expression of ribosomal genes may be favorable to feed efficiency improvement. In agreement, Gondret et al. (2017) reported a higher ribosomal genes expression acting on the translation process in efficient pigs. Horodyska et al. (2018) highlighted that protein synthesis is essential for pig efficiency, once it impacts muscle growth.

The PI3K pathway, together with AKT and mTOR, modulates muscle hypertrophy based on the regulation of protein translation, where ribosome activity is essential (DAVIS et al., 2012). Bottje et al. (2014) reported the role of these pathways on efficient broilers. Among the hub genes acting in the PIK3/AKT pathway, we identified *AKT3* and *FGFRI* (Fibroblast Growth Factor), which plays an essential role in the regulation of cell

proliferation, differentiation, and migration. It is important to highlight that the overmentioned pathways are nutritionally regulated, insulin-mediated, and the AKT activation lead to an increased number of myofibrils (DAVIS et al.,2012).

3.4.3.2. Immune system

We detected enriched pathways related to immune response (Figure 3.5), e.g. adherens junctions, leukocyte transendothelial migration, platelets activation, actin cytoskeleton regulation, relaxin pathway, sphingolipid signaling pathway, and tight junctions. Based on these pathways and the direction of the association between hub gene expression and the traits in this study, we hypothesized that a better immune response is favorable to feed efficiency variation.

The *RhoA* and *ROCK* genes are acting in several pathways identified as linked to feed efficiency related-traits (Figure 3.5) The Rho-GTPases, as RhoA, and their activity is associated with ROCK, Rho-associated kinases (ROCKs), which are downstream targets of RhoA (MOKADY; MEIRI, 2015; NOMA et al., 2006).

Among the pathways that these genes participate, there is the actin cytoskeleton regulation pathway. The actin cytoskeleton plays a role in regulating the stability of endothelial cell junctions and vascular permeability (RADEVA; WASCHKE, 2017), and it is crucial to several processes related to leukocyte migration (MARELLI-BERG; JANGANI, 2018). Also, in actin cytoskeleton regulation pathway, the regulation of RhoA occurs through ROCK interaction in the myosin light chain and leads to the formation of stress fibers when phosphorylated (MOKADY; MEIRI, 2015; NOMA et al., 2006; SCHOFIELD; BERNARD, 2013).

The regulation of actin cytoskeleton occurs by the turnover of F-actin polymers (FALZONE et al., 2013). That way, in leukocyte transendothelial migration (TEM), occurs the formation of endothelial membrane protrusions rich in F-actin filamentous. It is regulated by the remodeling of the F-actin cytoskeleton in endothelial cells hat coated these vessels (SCHIMMAEL et al., 2017). This process depends on the extracellular signal, which activates the small Rho-GTPases, like RhoA (SCHIMMAEL et al., 2017). Cerutti and Ridley (2017) reported that RhoA activity increased leukocyte migration.

Semi-permeable endothelial barriers separating the internal vessels space from tissues are mediated by endothelial cell-cell adhesions, as tight junctions and adherens junctions, which are connected to the actin cytoskeleton (CERUTTI; RIDLEY, 2017). Since leukocyte

migration for inflammation sites occurs between endothelial cells junctions, the leukocytes overcome the mechanical cytoskeletal barriers (ALON; BUUL, 2017).

The remodeling of endothelial cell-cell junctions occurs in response to extracellular signals, regulating the endothelial transition and permeability (CERUTTI; RIDLEY, 2017). Thus, pro-inflammatory stimuli increase permeability, being actin cytoskeleton regulation crucial to junction stability and changes in vascular permeability (CERUTTI; RIDLEY, 2017).

The *PECAMI* gene was enriched in this study and participated in the TEM pathway. It is related to cellular adhesion (PRIVRATSKY; NEWMAN, 2014), and in vascular barrier formation (PRIVRATSKY et al., 2010). The increase of *PECAMI* expression may decrease the vascular permeability via inhibition of RhoA (ROLFE et al., 2005) and, may act in response to inflammatory stimuli.

Sphingolipid signaling pathway was also enriched in this work. Into this pathway, we can pinpoint *GNAI2* and *GNAI3* genes. Both are members of Guanine nucleotide-binding protein (G proteins) family and belong to subunit $G_{\alpha i}$ (ZHANG et al., 2019). The S1P receptors (S1P1-5) (NIKOLOVA-KARAKASHIAN; REID, 2011) are G-coupled receptors to $G_{\alpha i}$ (ARGRAVES et al., 2008; REINHARD et al., 2017). The SP1 (sphingosine-1-phosphate) activity can mediate the regulation in the endothelial barrier and may involve the $G_{\alpha i}$ receptors (REINHARD et al., 2017).

Sphingolipids are part of the lipid membranes in cells (KHAVANDGAR; MURSHED, 2015) and can potentially influence muscle function (NIKOLOVA-KARAKASHIAN; REID, 2011). Biosynthetic and degradative pathways related to sphingolipids can result in an intermediated product, sphingosine (SMITH et al., 2000). Phosphorylated sphingosine by sphingosine kinase can lead to the formation of SP1 (NIKOLOVA-KARAKASHIAN; REID, 2011), that can act as a growth factor for muscle fibers (CORDEIRO et al., 2018; NIKOLOVA-KARAKASHIAN; REID, 2011), and it is essential for immune system response (RIVERA et al., 2008). Moreover, SP1 has a role in muscle regeneration with positive modulation on the growth of regenerating fibers (GERMINARIO et al., 2012). The extracellular stimuli increase permeability through RhoA activity, as well as increase S1P generation, leading to an increase in the endothelial barrier (CERUTTI; RIDLEY, 2017; REINHARD et al., 2017; ZHANG et al., 2016).

The explanation of the proinflammatory cytokines probably suppressing muscle growth factors can be due to energy prioritizing for the immune and homeostatic pathways (GABLER; SUPRLOCK, 2008). Elevation of inflammatory status can increase insulin

resistance since the fat tissue contributes to proinflammatory cytokines formation, interfering in the insulin pathway (STIENSTRA et al., 2007). The adipose tissue could contribute to the maintenance of the immune system (GABLER; SUPRLOCK, 2008). Starvation or excessive fat amounts can lead to issues in immune function and results in diseases (GABLER; SUPRLOCK, 2008).

Based on these findings, we speculate that actin cytoskeleton regulation may have an essential role in the regulation of leukocyte transendothelial migration and endothelial permeability through RhoA/ROCK activity. The RhoA activity may increase the endothelial permeability in response to proinflammatory stimuli, and this can lead to enhanced leukocyte migration. However, the increase of endothelial permeability is related to fluid losses across endothelial cells, and this might compromise organs function (RODRIGUES; GRANGER, 2015). The increase in RhoA activity can increase the *PECAMI* involved in decreased permeability and increase the S1P product from sphingolipids metabolism related to an increase in the endothelial barrier. Moreover, the *GNAI2* and *GNAI3* genes might be involved in the endothelial barrier function mediated by S1P activity.

We also find an enrichment of genes related to the activation platelets pathway (Figure 3.5). Platelets play a role in the cell and vessel integrity (SOPOVA et al., 2012), participating in inflammation response (KASPERSKA-ZAJAC; ROGALA, 2007), and contain several growth factors and cytokines (CÁCERES et al., 2008). Their activity is dependent on their shape change (CIMMINO; GOLINO, 2013). RhoA activity is directly related to their morphological alteration (AKBAR et al., 2016; LI et al., 2010). Platelets activation can modulate immune responses (ASSINGER, 2014) and it is related to leukocyte migration as well (KRAL et al., 2016). Platelet activation is a pathway not yet related to feed efficiency, but its relationship with immune response can explain its enrichment in our analysis.

Horodyska et al. (2018a) reported that highly efficient animals are related to higher immune response. Additionally, in lean pigs, efficient animals have an increase in the number of leukocytes (CLAPPERTON et al., 2006), and higher immune response (ADLER et al., 2013). Weber et al. (2016) reported an increase in immune response in low-RFI cattle. Zarek et al. (2017) and Gondret et al. (2017) identified genes related to improvement in response to inflammation in more efficient pigs. In agreement, Horodyska et al. (2019) identified genes in liver tissue related to the immune response in more efficient pigs. The better immune response, based on adaptive immunity, suggesting the efficient animals destine less feed to support the immune system as compared to the inefficient animals (HORODYSKA et al.

2018). This more efficient energy conserving can be utilized for other important biological processes.

3.5. Enrichment analysis to identify TFBS in hub genes

In this study, we identified six hub-TFs potentially regulating the genes identified here (Figure 3.6). Among these, *EST1* gene is a member of the ETs transcription factors (GARRIT-SHINA, 2013) family and was previously related to feed efficiency in pigs (WEBER et al., 2016).

Previous studies linked *EST1* gene with glucose metabolism and cell growth. Zhang et al. (2017) reported that an increased expression of *EST1* related to cell energy metabolism through glycolysis in cancer cells. The same authors stated that the knockout of *ETS1* gene decreases the GLUT-1, which is involved in aerobic glycolysis and expression and leading to a reduced level of glucose uptake and ATP production.

Verschoor et al. (2010) reported that the overexpression of *ETS1* in cancer cells promoted glucose deprivation leading to growth decrease, and might be involved in oxidative stress

Another hub-TF, *TCF4*, might regulates the *F13A1*, *EFEMP1*, and *THSB3* genes. These genes were previously identified as DE for RFI in muscle samples in our population (TIZIOTO et al., 2016). This TF is related to the Wnt signaling pathway (FENG et al., 2017), which has a role in myogenic development (ANAKWE et al., 2003) and vascular smooth muscle cell proliferation (ZHUANG et al., 2015). *TCF4* can have a role in TGF- β signaling pathway (CONTRERAS et al., 2016; FORREST et al., 2013). This pathway can contribute negatively to muscle mass (ALEXANDRA et al., 1997; LEE et al., 2010).

3.6. Data integration for identification of regulatory regions and functional variants

Data integration analysis revealed potential regulatory elements (i.e., miRNA binding site, TFs, and TFBS) linked to hub genes identified here. We identified 3' UTRs variants in hub genes showing allelic-specific expression (ASE) (SOUZA et al. - under review), potentially targeting miRNAs. One of these cases, is the hub-TF *ETS1*, a potential regulator for other hub genes identified in this study (Figure 3.6)

Among the 26 hub genes being affected by eQTLs, the *PCDH18*, and *SPARCL1* were under control of one cis-regulatory element (cis-eQTL) each. (CESAR et al., 2018), The *PCDH18* is a member of the cadherin family and plays a role in cell adhesion (AAMAR;

DAWID, 2008), and it is related to the immune system (VAZQUEZ-CINTRON et al., 2012). Transcription start site (TSS) of *PCDH18* is 0.367 Mb distant from the cis-eQTL affecting it. In the presence of the alternative allele, a site for the *NFE2L2* is created in the TFBS. The *NFE2L2* is an important TF, which can mediate the antioxidant pathway in response to oxidative stress in more efficient chicken (KONG et al., 2017).

The *SPARCLI*, is related to adipogenesis in humans, acting in differentiation capacity of adipose tissue (MEISSBURGUER et al., 2016). The TSS of this gene is 0.0381 Mb distant from the cis-eQTL affecting it. The alternative allele disrupts a binding site for the *NR3C1* TF. *NR3C1* encodes the glucocorticoid receptor (NIU et al., 2009), associated with ADG in pigs and potentially involved in protein catabolism (PILCHER et al., 2015). Both TFs, *NFE2L2*, and *NR3C1* were expressed in muscle in our samples.

3.7. CONCLUSION

Our co-expression network approach indicated putative central genes, such as *AKT3*, *ROCK1*, and *RhoA*, modulating feed efficiency related-traits. Pathway enrichment analysis reinforced the role of muscle proteins synthesis and immune response in feed efficiency. Furthermore, we identified potential functional variants in TFBS and miRNA binding site of the hub genes reported.

3.8. REFERENCES

AAMAR, E.; DAWID., I. B. Protocadherin-18a has a role in cell adhesion, behavior, and migration in zebrafish development. **Developmental Biology**, v. 318, n. 2, p. 335-346, 2008.

ADLER, M. et al. PBMCmc transcription profiles of pigs with divergent humoral immune responses and lean growth performance. **International Journal of Biological Sciences**, v. 9, p. 907, 2013.

ANDREWS, S. FastQC: a quality control tool for high throughput sequence data. Disponível em :<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>, 2010

AKBAR, H. et al. RhoA, and rac1 GTPases differentially regulate agonist-receptor-mediated reactive oxygen species generation in platelets. **Plos One**, v. 11, n. 9, p. e0163227, 2016.

ALBIG A. R; SCHIEMANN., W. P. Fibulin-5 antagonizes vascular endothelial growth factor (VEGF) signaling and angiogenic sprouting by endothelial cells. **DNA and Cell Biology**, v. 23, n. 6, p. 367-379, 2004.

ALEXANDRE, P. A. et al. Liver transcriptomic networks reveal the main biological processes associated with feed efficiency in beef cattle. **BMC Genomics**, v. 16, n. 1073, p. 2-13, 2015.

ANAKWE K. et al. Wnt signaling regulates myogenic differentiation in the developing avian wing. **Development**, v. 130, p. 3503-3514, 2003.

ARGRAVES K. M. et al. High-density lipoprotein-associated sphingosine 1-phosphate promotes endothelial barrier function. **Journal of Biological Chemistry**, v. 286, n. 36, p. 25074-25081, 2008.

ARTHUR, P.F.; HERD, R.M. Residual feed intake in beef cattle. **Revista Brasileira de Zootecnia**, Viçosa, v. 37, n. esp. 37, p. 269-279, 2008.

ASSINGER, A. Platelets, and infection – an emerging role of platelets in viral infection. **Frontiers in Immunology**, v. 5, p. 649, 2014.

BASARAB, J. A. et al. Residual feed intake and body composition in young growing cattle. **Canadian Journal of Animal Science**, v. 83, n. 2, p. 180-204, 2015.

BENJAMINI; Y.; HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society: Series b (methodological)**, v. 57, n. 1, p. 289-300, 1995

BINDEA, G. et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. **Bioinformatics**. v.23, n.8. p. 1091-3, 2009

BOGAN, R. et al. A mouse model for human osteogenesis imperfecta type VI. **Journal of Bone and Mineral Research**, v. 28, n. 7, p. 1531–1536, 2013.

BUSH, J. A. et al. Translational control of protein synthesis in muscle and liver of growth hormone-treated pigs. **Endocrinology**, v. 144, n. 4, 2003

CACERES M et al. Effect of platelet-rich plasma on cell adhesion, cell migration, and myofibroblastic differentiation in human gingival fibroblasts. **Journal of Periodontology**, v. 79, n. 4, p. 714-720, 2008.

CARNAGARIN, R. et al. Molecular aspects of glucose homeostasis in skeletal muscle – a focus on the molecular mechanisms of insulin resistance. **Molecular and Cellular Endocrinology**, v. 417, p. 52-62, 2015.

CESAR, Al. S. M. et al. Identification of putative regulatory regions and transcription factors associated with intramuscular fat content traits. **BMC Genomics**, v. 19, n. 499, p. 2-20, 2018.

CERUTTI, C.; RIDLEY, J. A. Endothelial cell-cell adhesion, and signaling. **Experimental Cell Research**, v. 358, n. 1, p. 31-38, 2017.

CIMMINO, G; GOLINO, P. Platelet biology and receptor pathways. **Journal of Cardiovascular Translational Research**, v. 6, n. 3, p. 299-309, 2013.

CHEN, X. et al. Effects of osteoglycin (OGN) on treating senile osteoporosis by regulating mscs. **Musculoskeletal Disorders**, v. 18, n. 423, p. 1-10, 2017.

CLAPPERTON M et al. Selection for lean growth and food intake leads to correlated changes in innate immune traits in large white pigs. **Animal Science**, v. 82, n. 6, p. 867-876, 2006.

CONTRERAS, O. et al. Connective tissue cells expressing fibro/adipogenic progenitor markers increase under chronic damage: relevance in fibroblast-myofibroblast differentiation and skeletal muscle fibrosis. **Cell and Tissue Research**, v. 364, n. 3, p. 647-660, 2016

CONNOR, E.E. Invited review: improving feed efficiency in dairy production: challenges and possibilities. **Animal**, v. 9, n. 3, p. 395-408, 2015.

CORDEIRO A. V. et al. The role of sphingosine-1-phosphate in skeletal muscle: physiology, mechanisms, and clinical perspectives. **Journal of Cellular Physiology**, v. 234, n. 7, p. 10047-10059, 2019.

DAM, S. V. et al. Gene co-expression analysis for functional classification and gene-disease predictions. **Briefings in Bioinformatics**, v. 19, n. 4, p. 575–592, 2018.

Davis M. E. et al. **review: biological determinants of between-animal variation in feed efficiency of growing beef cattle: in hormonal regulation of feed efficiency:** Hormonal Regulation of Feed Efficiency. 12 ed.,: John Wiley & Sons, 2012, s321-s335 p.

DE OLIVEIRA, P.S.N. et al. Identification of genomic regions associated with feed efficiency in nelore cattle. **BMC Genetics**, v. 15, n. 1, p. 1-10, 2014.

DE OLIVEIRA, P. S. N. et al. An integrative transcriptome analysis indicates regulatory mRNA-miRNA networks for residual feed intake in nelore cattle. **Scientific Reports**, v. 8, n. 17072, p. 1-12, 2018

- DINIZ, W. J. S. et al. Detection of co-expressed pathway modules associated with mineral concentration and meat quality in nelore cattle. **Frontiers in Genetics**, v. 10, n. 210, p. 1-12, 2019
- DOBIN, A. et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics**. v. 29, p. 15–21, 2013.
- DURINCK, S. et al. Mapping identifiers for the integration of genomic datasets with the r/Bioconductor package biomart. **Nature Protocols**, Cidade, v. 4, p. 1184–1191, 2009.
- EWEL, P. et al. M. MultiQC: summarize analysis results for multiple tools and samples in a single report. **Bioinformatics**, v. 32, n.19, p. 3047-3048, 2016.
- FALZONE, T. T. et al. Actin assembly factors regulate the gelation kinetics and architecture of f-actin networks. **Biophysical Journal**, v. 104, p. 1709-1719, 2013.
- FENG, W. et al. Growth retardation induced by avian leukosis virus subgroup j associated with downregulated wnt/ β -catenin pathway. **Microbial Pathogenesis**, v. 104, p. 48-55, 2017
- FONSECA, L. D. et al. Liver proteomics unravel the metabolic pathways related to feed efficiency in beef cattle. **Scientific Reports**, v. 9, n. 5364, p. 1-11, 2019.
- FULDA, S. Modulation of mitochondrial apoptosis by PI3K inhibitors. **Mitochondrion**, v. 13, p. 195-198, 2013
- GABLER N. K; SPURLOCK, M. E. Integrating the immune system with the regulation of growth and efficiency. **Journal of Animal Science** **86**,, v. 86, n. suppl_14, p. E64-E74, 2008.
- GARRETT-SINHA L. A. Review of ets1 structure, function, and roles in immunity. **Cellular and Molecular Life Science**, v. 70, n. 18, p. 3375-3390, 2013.
- GERMINARIO E. et al. S1P2 receptor promotes mouse skeletal muscle regeneration. **Journal of Applied Physiology**, v. 113, n. 5, p. 707-713, 2012.
- GIULIETTI, M. et al. Emerging biomarkers in bladder cancer identified by network analysis of transcriptomic data. **Frontiers in Oncology**, v. 8, p. 450, 2018.
- GONDRET, F. et al. A transcriptome multi-tissue analysis identifies biological pathways and genes associated with variations in feed efficiency of growing pigs. **BMC Genomics**, v. 18, n. 244, p. 1-17, 2017.
- GRILL J. I. et al. Loss of dro1/ccdc80 results in obesity and promotes adipocyte differentiation. **Molecular and Cellular Endocrinology**, v. 439, p. 286-296,2017.
- FORREST, M. P. et al. Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. **Plos One**, v. 8, p. e73169, 2013

HEEMSKERK, N. et al. Rho-GTPase signaling in leukocyte extravasation an endothelial point of view. **Cell adhesion & Migration**, v. 8, n. 2, p. 67-75, 2014.

HERD, R M.; ARTHUR P. F. Physiological basis for residual feed intake. **Journal of Animal Science**, v. 87, n. Issue suppl_14, p. E64–E71,2009.

HOPKINS, B. D. et al. Suppression of insulin feedback enhances the efficacy of PI3K inhibitors. **Nature**, v. 560, n. 7719, p. 499, 2018.

HORODYSKA, J. et al. Analysis of meat quality traits and gene expression profiling of pigs divergent in residual feed intake. **Meat Science**, v. 137, p. 265-274, 2018.

HORODYSKA, J. et al. RNA-Seq of muscle from pigs divergent in feed efficiency and product quality identifies differences in immune response, growth, and macronutrient and connective tissue metabolism. **BMC Genomics**, v. 19, n. 1, p. 791, 2018a.

HORODYSKA, J. et al. RNA-Seq of liver from pigs divergent in feed efficiency highlights shifts in macronutrient metabolism, hepatic growth, and immune response. **Frontiers in Genetics**, v. 10, n. 1, p. 17, 2019

HUANG, X. et al. The PI3K/AKT pathway in obesity and type 2 diabetes. **International Journal of Biological Sciences**, v. 14, n. 11, p. 1483-1496, 2018.

KAJI, H. Interaction between muscle and bone. **Journal of Bone Metabolism**, v. 21, n. 1, p. 29-40, 2014.

KASPERSKA-ZAJAC A.; ROGALA, B. Platelet activation during allergic inflammation. **Inflammation**, v. 30, n. 5, p. 161-6, 2007.

KEN-ICHIRO, T. et al. Role of osteoglycin in the linkage between muscle and bone. **The Journal of Biological Chemistry**, v. 287, n. 15, p. 11616–11628, 2012.

KHAVANDGAR, Z.; MURSHED, M. Sphingolipid metabolism and its role in the skeletal tissues. **Cellular and Molecular Life Sciences**, v. 72, n. 5, p. 959-969, 2015.

KHIAOSA-ARD, R.; ZEBELI, Q. Cattle's variation in rumen ecology and metabolism and its contributions to feed efficiency. **Livestock Science**, v. 162, p. 66-75, 2014.

KLIP A. et al. Signal transduction meets vesicle traffic: the software and hardware of GLUT4 translocation. **American journal of Physiology-Cell Physiology**, v. 306, n. 10, p. C879-C886, 2014.

KOCH, R. et al. Efficiency of feed use in beef cattle. **Journal of Animal Science**, v. 22, n. 2, p. 486-494, 1963.

KONG B. W. et al. Rna sequencing for global gene expression associated with muscle growth in a single male modern broiler line compared to a foundational barred Plymouth rock chicken line. **bmc genomics**, v. 18, n. 1, 2017.

KRAL J. B. et al. Platelet interaction with innate immune cells. **Transfusion Medicine and Chemotherapy**, v. 43, n. 2, p. 78-88, 2016.

LANGFELDER, P.; HORVATH S. WGCNA: an R package for weighted correlation network analysis. **BMC Bioinformatics**. v. 9, n. 1, p. 559, 2008

LAURINO, L. et a. PI3K activation by igf-1 is essential for the regulation of membrane expansion at the nerve growth cone. **Journal of Cell Science**, v. 118, n. 16, p. 3653-3662, 2005.

LI, F. et al. Pigment epithelium-derived factor upregulate expression of vascular endothelial growth factor by human mesenchymal stem cells: possible role in pedf regulated matrix mineralization. **Biochemical and Biophysical Research Communications**, v. 478, n. 3, p. 106–1110, 2016.

LIU, J. et al. The association study of plasma levels of pigment epithelium-derived factor with the acute coronary syndrome in the Chinese han population. **Cardiology**, v. 127, p. 31-37, 2013.

LI, M. et al. Correlation between pigment epithelium-derived factor (PEDF) level and degree of coronary angiography and the severity of coronary artery disease in a Chinese population. **Clinical research**, v. 24, p. 1751, 2018.

MA, S. et al. The effects of pigment epithelium-derived factor on atherosclerosis: putative mechanisms of the process. **Lipids in Health and Disease**, v. 17, n. 240, p. 1-11, 2018.

MADER C. J. et al. Relationships among measures of growth performance and efficiency with carcass traits, visceral organ mass, and pancreatic digestive enzymes in feedlot cattle. **Journal of Animal Science**, v. 87, n. 4, 2009.

MARELLI-BERG, F. M.; J, Maryam. Metabolic regulation of leukocyte motility and migration. **Journal of Leukocyte Biology**, v. 104, n. 2, p. 285-293, 2018.

MEISSBURGER, Bettina. Regulation of adipogenesis by paracrine factors from adipose stromal-vascular fraction - a link to fat depot-specific differences. **Biochimica et Biophysica Acta (BBA) - molecular and cell biology of lipids**, v. 1861, n. 9, p. 1121-1131, 2016

MONTANO-BERMUDEZ et al. Energy requirements for maintenance of crossbred beef cattle with the different genetic potential for milk. **Journal of Animal Science**, v. 68, n. 8, p. 2279-2288, 1990.

MOKADY D.; MEIRI, D. Rhogtpases – a novel link between cytoskeleton organization and cisplatin resistance. **Drug Resistance Updates**, v. 19, p. 22-32, 2015

MOORE, S. S. et al. Molecular basis for residual feed intake in beef cattle. **Journal of Animal Science**, v. 87, n. suppl_14, p. E64-E71, 2009.

MULLER W. A. Mechanisms of leukocyte transendothelial migration. **Annual Review of Pathology: Mechanisms of Disease**, v. 6, p. 323-344, 2011.

NADER, G. A. Ribosomes ‘muscle up’ postnatal muscle growth. **The Journal of Physiology**, v. 592, n. 23, p. 5143, 2014.

NASCIMENTO, M.L. et al. Feed efficiency indexes and their relationships with the carcass, non-carcass, and meat quality traits in Nelore steers. **Meat Science**, v. 116, p. 78-85, 2016.

NIU, N. et al. Human glucocorticoid receptor α gene (NR3C1) pharmacogenomics: gene resequencing and functional genomics. **The Journal of Clinical Endocrinology & Metabolism** v. 94, n. 8, p. 3072-84, 2009

NOMA, K. et al. Physiological role of ROCKSs in the cardiovascular system. **American Journal of Physiology-Cell Physiology**, v. 290, n. 3, p. C661-C668., 2006.

PILCHER C. M. et al. Transcript profiles in longissimus dorsi muscle and subcutaneous adipose tissue: a comparison of pigs with different postweaning growth rates. **Journal of Animal Science**, v. 93, n. 5, p. 2134-2143, 2015

PRIVRATSKY J. R. PECAM-1: conflicts of interest in inflammation. **Life Sciences**, v. 87, p. 69-82, 2010.

PRIVRATSKY J. R.; NEWMAN., P. J. PECAM-1: a regulator of endothelial junctional integrity. **Cell and Tissue Research**, v. 355, n. 3, p. 607-619, 2014.

RADEVA, M. Y.; WASCHKE, J. Mind the gap: mechanisms regulating the endothelial barrier. **Acta Physiologica**, v. 222, n. 1, p. e12860, 2017.

REINHARD N. R et al. The balance between $gai-cdc42/rac$ and $ga12/13-RhoA$ pathways determines endothelial barrier regulation by sphingosine-1-phosphate. **Molecular Biology of the Cell**, v. 28, n. 23, p. 3371-3382, 2017.

RIVERA J. et al. The alliance of sphingosine-1-phosphate and its receptors in immunity. **Nature Reviews Immunology**, v. 8, n. 10, p. 753, 2008.

RITCHIE, M.E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. **Nucleic Acids Research**. V. 20, n.7, p. e47-, 2015

RODRIGUES, S F; GRANGER, D N. Blood cells and endothelial barrier function. **Tissue Barriers**, v. 3, p. e978720, 2015.

ROLFE, B. E. et al. Rho and vascular disease. **Atherosclerosis**, v. 183, p. 1-6, 2005

TIZIOTO, P. C. et al. Global liver gene expression differences in nelore steers with divergent residual feed intake phenotypes. **BMC Genomics**, v. 16, n. 242, p. 1-13, 2015.

TIZIOTO, P. C. et al. Gene expression differences in longissimus muscle of nelore steers genetically divergent for residual feed intake. **Scientific Reports**, v. 6, n. 39493, p. 1-15, 2016.

TREMBLAY, F. et al. Bidirectional modulation of adipogenesis by the secreted protein ccdc80/dro1/urb. **Journal of Biological Chemistry**, v. 284, n. 12, p. 8136-8147, 2009

TREMBLAY, F. et al. Loss of coiled-coil domain containing 80 negatively modulates glucose homeostasis in diet-induced obese mice. **Endocrinology**, v. 153, n. 9, p. 4290-4303, 2012.

SCHIAFFINO, S. et al. Mechanisms regulating skeletal muscle growth and atrophy. **The febs journal**, v. 280, n. 17, p. 4294-4314, 2013

SCHIMMEL, L. et al. Leukocyte transendothelial migration: a local affair. **Small GTPases**, v. 8, n. 1, p. 1-15, 2017.

SCHOFIELD, A. V.; BERNARD, O. Rho-associated coiled-coil kinase (rock) signaling and disease. **Critical Reviews in Biochemistry and Molecular Biology**, v. 48, n. 3, p. 301-316, 2013.

SERIN, E. A. et al. Learning from co-expression networks: possibilities and challenges. **Frontiers in Plant Science**, v. 7, p. 444, 2016.

SHANNON, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome Research**. v.13, n. 11, p. 2498-50, 2003

SMITH, L. R. et al. Systems analysis of biological networks in skeletal muscle function. **Wiley Interdisciplinary Reviews: Systems Biology and Medicine**, v. 5, n. 1, p. 55-71, 2013.

SOPOVA K et al. Platelets and platelet interaction with progenitor cells in vascular homeostasis and inflammation. **Current Vascular Pharmacology**, v. 10, n. 5, p. 555-562, 2012.

SOUZA, M. M. et al. A comprehensive manually-curated compendium of bovine transcription factors. **Scientific Reports**, v. 8, n. 13747, 2018.

SPENCER J. A. et al. Context-specific effects of fibulin-5 (dance/evec) on cell proliferation, motility, and invasion. **The Journal of Biological Chemistry**, v. 277, n. 30, p. 2946-2951, 2002.

SPENCER J. A. et al. Altered vascular remodeling in fibulin-5-deficient mice reveals a role of fibulin-5 in smooth muscle cell proliferation and migration. **Proceedings of the National Academy of Sciences**, v. 102, n. 8, p. 2946-2951, 2005.

STIENSTRA R. et al. Ppars, obesity, and inflammation. **PPARa Research**, 2007.

VAZQUEZ-CINTRON, E. J. et al. Protocadherin-18 is a novel differentiation marker and an inhibitory signaling receptor for cd8+ effector memory t cells. **Plos One**, v. 7, p. e36101, 2012

VELAZCO, J. I et al. Daily methane emissions and emission intensity of grazing beef cattle genetically divergent for residual feed intake. **Animal Production Science**, v. 57, n. 4, p. 627-635, 2016.

VERSCHOOR. M. L. Ets-1 regulates energy metabolism in cancer cells. **Plos One**, v. 5, n. 10, p. e13565, 2010.

YANAGISAWA, H.; SCHLUTERMAN, M. K.; BREKKEN, R. A. Fibulin-5, an integrin-binding matricellular protein: its function in development and disease. **Journal of Cell Communication and Signaling**, v. 3, p. 337–347, 2009.

WEBER, K. L. et al. Identification of gene networks for residual feed intake in Angus cattle using genomic prediction and RNA-seq. **PIOs One**, v. 11, n. 3, p. e0152274, 2016.

WEN, Y. et al. Ribosome biogenesis is necessary for skeletal muscle hypertrophy. **Exercise and Sport Sciences Reviews**, v. 44, n. 3, p. 110, 2016.

ZAREK et al. Differential expression of genes related to gain and intake in the liver of beef cattle. **BMC Research Notes**, v. 10, n. 1, p. 1, 2017

ZHANG X. E. et al. Activation of RhoA, but not rac1, mediates early stages of s1p-induced endothelial barrier enhancement. **Plos One**, v. 11, n. 5, p. e0155490, 2016.

ZHANG X et al. Ets-1: a potential target of glycolysis for metabolic therapy by regulating glucose metabolism in pancreatic cancer. **International Journal of Oncology**, v. 50, n. 1, p. 232-240, 2017.

ZHANG, Y. et al. Guanine and nucleotide binding protein 3 promotes odonto/osteogenic differentiation of apical papilla stem cells via jnk and erk signaling pathways. **International Journal of Molecular Medicine**, v. 43, n. 1, p. 382-392, 2019.

ZHUANG, Y. Hyperlipidemia induces vascular smooth muscle cell proliferation involving wnt/b-catenin signaling. **Cell Biology International**, v. 40, n. 2, p. 121-130, 2015.

4. GENERAL CONCLUSION

In this study, we identified several genes in pathways related to feed efficiency and can be useful as biomarkers. Also, we identified potential functional variants for this trait in Nelore cattle. However, it is important more studies based on other populations to confirm these findings.

5. SUPPLEMENTARY INFORMATION

Table-S1. Results from the linear association of the target genes expression levels among feed efficiency related-traits (ADG, BW, DMI, FCR, FE, KR, MBW, RFI and, RGR).

Genes	Estimated effect \pm SE ¹	p-value
ADG - Average daily gain		
<i>CTGF</i>	-0.05463 \pm 0.03615	0.1379
<i>EGR1</i>	0.006249 \pm 0.02985	0.8352
<i>COL1A1</i>	0.04015 \pm 0.03465	0.2528
<i>PRUNE2</i>	0.01765 \pm 0.03172	0.5807
<i>PRUNE2_isoform</i>	-0.05107 \pm 0.02667	0.0620
<i>CYP2B6</i>	0.01868 \pm 0.02562	0.4699
BW - Body weight		
<i>CTGF</i>	-4.8956 \pm 5.1061	0.3429
<i>EGR1</i>	-0.9382 \pm 4.2168	0.8250
<i>COL1A1</i>	6.0785 \pm 4.8945	0.2209
<i>PRUNE2</i>	-3.0165 \pm 4.4808	0.5043
<i>PRUNE2_isoform</i>	-1.7563 \pm 3.7669	0.6433
<i>CYP2B6</i>	-2.5767 \pm 3.6192	0.4802
DMI - Dry matter intake		
<i>CTGF</i>	-0.01882 \pm 0.1579	0.9057
<i>EGR1</i>	-0.07655 \pm 0.1304	0.5602
<i>COL1A1</i>	0.1239 \pm 0.1514	0.4175
<i>PRUNE2</i>	-0.3194 \pm 0.1386	0.0259
<i>PRUNE2_isoform</i>	0.1052 \pm 0.1165	0.3713
<i>CYP2B6</i>	-0.05125 \pm 0.1119	0.6492
FCR - Feed conversion ratio		
<i>CTGF</i>	0.1642 \pm 0,1159	0.1636
<i>EGR1</i>	-0.05661 \pm 0.09572	0.5573
<i>COL1A1</i>	-0.06448 \pm 0.1111	0.5646
<i>PRUNE2</i>	-0.2427 \pm 0.1017	0.0214
<i>PRUNE2_isoform</i>	0.2436 \pm 0.08551	0.0067

<i>CYP2B6</i>	-0.1074 ± 0.08215	0.1978
FE - Feed efficiency		
<i>CTGF</i>	-0.00633 ± 0.004361	0.1540
<i>EGR1</i>	0.002492 ± 0.003601	0.4925
<i>COL1A1</i>	0.003166 ± 0.004180	0.4528
<i>PRUNE2</i>	0.008361 ± 0.003827	0.0343
<i>PRUNE2_isoform</i>	-0.00756 ± 0.003217	0.0233
<i>CYP2B6</i>	0.002903 ± 0.003091	0.3527
KI - Kleiber ratio		
<i>CTGF</i>	-0.00049 ± 0.000390	0.2176
<i>EGR1</i>	0.000119 ± 0.000322	0.7137
<i>COL1A1</i>	0.000269 ± 0.000374	0.4759
<i>PRUNE2</i>	0.000330 ± 0.000343	0.3403
<i>PRUNE2_isoform</i>	-0.00056 ± 0.000288	0.0569
<i>CYP2B6</i>	0.000349 ± 0.000277	0.2143
MBW- Metabolic body weight		
<i>CTGF</i>	-0.8380 ± 0.8882	0.3506
<i>EGR1</i>	-0.1719 ± 0.7335	0.8158
<i>COL1A1</i>	1.0565 ± 0.8514	0.2212
<i>PRUNE2</i>	-0.5250 ± 0.7794	0.5041
<i>PRUNE2_isoform</i>	-0.3110 ± 0.6553	0.6374
<i>CYP2B6</i>	-0.4478 ± 0.6296	0.4807
RFI - Residual feed intake		
<i>CTGF</i>	0.003541 ± 0,09381	0.9701
<i>EGR1</i>	-0.04633 ± 0,07558	0.5434
<i>COL1A1</i>	0.1294 ± 0,08864	0.1521
<i>PRUNE2</i>	-0.2557 ± 0,07907	0.0024
<i>PRUNE2_isoform</i>	0.08431 ± 0,07458	0.265
<i>CYP2B6</i>	-0.00709 ± 0,06651	0.9157
RGR - Relative growth ratio		
<i>CTGF</i>	-0.00622 ± 0.004205	0.1463
<i>EGR1</i>	0.001202 ± 0.003473	0.7308
<i>COL1A1</i>	0.003954 ± 0.004031	0.332
<i>PRUNE2</i>	0.004352 ± 0.003690	0.2446
<i>PRUNE2_isoform</i>	-0.00541 ± 0.003102	0.0884
<i>CYP2B6</i>	0.002760 ± 0.002981	0.3596

1.SE = Standard error

Table-S2. Explained variance of the ME

Module	expression variation
MEblack	0.529161792
MEblue	0.341166502
MEbrown	0.348460323
MEcyan	0.455369844
MEdarkgreen	0.474324797
MEdarkgrey	0.471251943
MEdarkred	0.422467663
MEdarkturquoise	0.546128967
MEgreen	0.371211896
MEgreenyellow	0.470073785
MEgrey	0.120013449
MEgrey60	0.672276362
MElightcyan	0.404563372
MElightgreen	0.378664464
MElightyellow	0.61420966
MEmagenta	0.34104693
MEmidnightblue	0.506621523
MEorange	0.535249708
MEpink	0.481340867
MEpurple	0.351808253
MEroyalblue	0.631023845
MEtan	0.628122526
MEturquoise	0.495467429
MEwhite	0.516652131
MEyellow	0.396791351

Table-S3. Hub gene annotation

Ensembl_id	Strand	Name	chr	Initial (pb)	End (pb)	TSS
ENSBTAG00000000057	1	THBS3	3	15470302	15480443	15470302
ENSBTAG00000000078	1	GLIPR2	8	60797959	60816816	60797959
ENSBTAG00000000099	1	CERS2	3	19835017	19843638	19835017
ENSBTAG00000000113	1	ARHGEF2	3	14843522	14868193	14843522
ENSBTAG00000000137	-1	FRYL	6	68829579	68974140	68974140
ENSBTAG00000000184	1	EIF2AK3	11	47302536	47384700	47302536
ENSBTAG00000000210	1	ARHGAP31	1	64747665	64795985	64747665
ENSBTAG00000000215	1	GNB1	16	52106960	52184132	52106960
ENSBTAG00000000215	1	GNB1	16	52106960	52184132	52158465
ENSBTAG00000000236	1	YWHAZ	14	65584487	65617329	65584487
ENSBTAG00000000283	-1	CSF1	3	33607139	33621030	33621030
ENSBTAG00000000310	1	MFAP5	5	101647629	101659377	101647629
ENSBTAG00000000653	1	PPP1R16B	13	68258627	68366080	68258627
ENSBTAG00000000711	1	NDRG1	14	9109762	9165926	9109762
ENSBTAG00000000742	1	AMOTL2	1	136053624	136071002	136053624
ENSBTAG00000000781	-1	HIP1	25	34393880	34528240	34528240
ENSBTAG00000000925	-1	GLT8D2	5	67994626	68036627	68036627
ENSBTAG00000001004	-1	ESAM	29	28601863	28610589	28610589
ENSBTAG00000001041	1	SAO	19	43555808	43559784	43555808
ENSBTAG00000001097	1	FKBP7	2	18413884	18421599	18413884
ENSBTAG00000001108	1	GMCL1	11	68149829	68205598	68149829
ENSBTAG00000001117	1	ANKRD50	17	33496005	33529948	33496005
ENSBTAG00000001141	1	ADAM17	11	87898074	87940943	87898074
ENSBTAG00000001182	-1	SEPT7	4	61611420	61710127	61710127
ENSBTAG00000001204	-1	JCAD	13	35331913	35368126	35368126
ENSBTAG00000001360	1	RPS12	9	71974860	71978200	71974860
ENSBTAG00000001483	1	SRGAP2	16	4023323	4101731	4023323
ENSBTAG00000001509	1	ELK3	5	60823240	60891359	60823240
ENSBTAG00000001523	-1	YES1	24	35973659	36038238	36038238
ENSBTAG00000001585	1	WIPF1	2	22253662	22294719	22253662
ENSBTAG00000001589	1	TM9SF2	12	80317273	80373389	80317273
ENSBTAG00000001597	1	PITPNM2	17	54716441	54799043	54716441
ENSBTAG00000001648	-1	RPL21	12	32852826	32859542	32859542
ENSBTAG00000001657	1	PICALM	29	9613012	9665840	9613012
ENSBTAG00000001745	-1	LUM	5	21037443	21044658	21044658
ENSBTAG00000001814	1	PLXND1	22	56774467	56824220	56774467
ENSBTAG00000001826	1	SASH1	9	86852607	86998307	86852607
ENSBTAG00000001928	1	PDIA6	11	86834898	86857648	86834898
ENSBTAG00000001961	-1	MAP1B	20	9330175	9419040	9419040
ENSBTAG00000001987	-1	SWAP70	15	43509670	43585976	43585976
ENSBTAG00000002038	1	RPL14	22	13336573	13339983	13336573
ENSBTAG00000002060	1	RPL19	19	40332947	40334703	40332947
ENSBTAG00000002068	1	TAGLN2	3	9878839	9886647	9878839
ENSBTAG00000002092	1	PII6	23	10836790	10848749	10836790
ENSBTAG00000002092	1	PII6	23	10836790	10848749	10836838
ENSBTAG00000002108	1	YWHAQ	11	87842280	87873674	87842280
ENSBTAG00000002286	-1	DNAJC18	7	52305244	52337268	52337268

ENSBTAG00000002326	1	LGALS3	10	67843328	67861113	67843328
ENSBTAG00000002329	-1	ASAP2	11	88012967	88104077	88104077
ENSBTAG00000002341	-1	ETS1	29	32358407	32430222	32430222
ENSBTAG00000002391	1	TGFB1I1	25	27760811	27766954	27760811
ENSBTAG00000002411	1	CTNND1	15	82283934	82329276	82283934
ENSBTAG00000002468	1	RPS28	7	18198992	18200124	18198992
ENSBTAG00000002608	1	SEPT2	3	120968024	120988257	120968024
ENSBTAG00000002633	-1	SEPT9	19	55118592	55153750	55153750
ENSBTAG00000002646	-1	DYNC1I2	2	24807357	24860776	24860776
ENSBTAG00000002648	1	RPS18	23	7388703	7393361	7388703
ENSBTAG00000002683	-1	PFKP	13	45500599	45553520	45553520
ENSBTAG00000002697	1	KCTD10	17	65950166	65979498	65950166
ENSBTAG00000002736	-1	DNMT1	7	15913446	15956917	15956917
ENSBTAG00000002804	-1	PDGFRB	7	63496230	63532702	63532702
ENSBTAG00000002996	-1	SHROOM4	X	93577262	93628877	93628877
ENSBTAG00000003061	1	LAMA5	13	55379959	55433278	55379959
ENSBTAG00000003109	1	ITM2B	12	18114553	18139506	18114553
ENSBTAG00000003124	1	MFSD14B	8	30529	228065	30529
ENSBTAG00000003166	1	AXL	18	50677199	50709510	50677199
ENSBTAG00000003229	-1	RPL23	19	40075000	40079360	40079360
ENSBTAG00000003238	-1	MEOX2	4	23943520	24019359	24019359
ENSBTAG00000003265	1	ADD3	26	30840113	30971458	30840113
ENSBTAG00000003276	1	PRKCH	10	73694668	73927120	73694668
ENSBTAG00000003312	-1	CHST15	26	44047737	44081511	44081511
ENSBTAG00000003418	-1	MSN	X	100070406	100162454	100162454
ENSBTAG00000003505	-1	DCN	5	21080013	21119087	21119087
ENSBTAG00000003530	-1	DDX31	11	102698014	102771662	102771662
ENSBTAG00000003585	-1	CD47	1	53103996	53169038	53169038
ENSBTAG00000003598	-1	P3H3	5	103972418	103985413	103985413
ENSBTAG00000003602	1	RBPJ	6	47318340	47429147	47318340
ENSBTAG00000003745	1	WDR48	22	12416919	12469477	12416919
ENSBTAG00000003777	-1	TIE1	3	103278178	103297710	103297710
ENSBTAG00000003825	-1	PTPN12	4	43834354	43884954	43884954
ENSBTAG00000003832	1	MFAP2	2	136187693	136192151	136187693
ENSBTAG00000003880	1	EMILIN2	24	37496731	37556313	37496731
ENSBTAG00000003902	-1	ZNF512	11	72067123	72097173	72097173
ENSBTAG00000004014	-1	FBLN2	22	58990282	59038424	59038424
ENSBTAG00000004094	-1	SPARCL1	6	104149824	104202396	104202396
ENSBTAG00000004190	1	ARHGAP29	3	49392814	49474286	49392814
ENSBTAG00000004207	-1	CD93	13	42244268	42247239	42247239
ENSBTAG00000004238	1	TACC1	27	33597619	33659545	33597619
ENSBTAG00000004261	-1	SPON2	6	109224627	109228237	109228237
ENSBTAG00000004279	1	RHOA	22	51277867	51323093	51277867
ENSBTAG00000004334	-1	NCSTN	3	9453300	9468330	9468330
ENSBTAG00000004356	-1	ROBO4	29	28719989	28734705	28734705
ENSBTAG00000004383	-1	FNBP1L	3	50082610	50146984	50146984
ENSBTAG00000004427	-1	OSBPL8	5	5823987	5903115	5903115
ENSBTAG00000004553	-1	TPM4	7	7923143	7948265	7948265

ENSBTAG00000004757	1	LTBP4	18	50173268	50202104	50173268
ENSBTAG00000004757	1	LTBP4	18	50173268	50202104	50177282
ENSBTAG00000004885	-1	DDR2	3	6672018	6847222	6847222
ENSBTAG00000004937	-1	SEC61A1	22	60217808	60229341	60229341
ENSBTAG00000004950	1	BRB	10	26288223	26289852	26288223
ENSBTAG00000005043	1	TIMP1	X	91232235	91236073	91232235
ENSBTAG00000005161	-1	ATP8B2	3	16269720	16289004	16289004
ENSBTAG00000005250	1	BGN	X	39639906	39653687	39639906
ENSBTAG00000005296	1	RPL13A	18	56394558	56398082	56394558
ENSBTAG00000005443	-1	MIER1	3	78663151	78730912	78730912
ENSBTAG00000005481	1	ADAM10	10	51598073	51739153	51598073
ENSBTAG00000005483	-1	ESYT2	4	120230771	120306448	120306448
ENSBTAG00000005620	1	RPS3	15	55370367	55375306	55370367
ENSBTAG00000005960	-1	EPB41L2	9	69916605	70027346	70027346
ENSBTAG00000006007	1	SH3GL1	7	20923762	20951051	20923762
ENSBTAG00000006024	-1	ISLR	21	34903758	34906887	34906887
ENSBTAG00000006126	-1	GTF3C1	25	25276351	25356558	25356558
ENSBTAG00000006130	-1	CLEC14A	21	48843729	48845622	48845622
ENSBTAG00000006214	-1	LOXL2	8	71278583	71349786	71349786
ENSBTAG00000006234	-1	NPR1	3	16747641	16761172	16761172
ENSBTAG00000006322	1	DENND5A	15	44009023	44054311	44009023
ENSBTAG00000006335	1	STAT6	5	56658077	56670842	56658077
ENSBTAG00000006346	1	DAP	20	62630506	62697437	62630506
ENSBTAG00000006487	-1	RPS9	18	63381416	63388728	63388728
ENSBTAG00000006747	-1	LTBP3	29	44380356	44400578	44400578
ENSBTAG00000006837	-1	UBA6	6	85050210	85137078	85137078
ENSBTAG00000006877	1	MMP16	14	76757695	77111545	76757695
ENSBTAG00000006995	1	SPTBN1	11	37030009	37241384	37030009
ENSBTAG00000007152	-1	OS9	5	56070693	56104924	56104924
ENSBTAG00000007153	-1	C1QA	2	130792855	130795743	130795743
ENSBTAG00000007153	-1	C1QA	2	130792855	130795743	130795255
ENSBTAG00000007173	1	PDGFRA	6	71373513	71421283	71373513
ENSBTAG00000007203	1	STT3A	29	29374406	29395597	29374406
ENSBTAG00000007268	1	F13A1	23	48633936	48776698	48633936
ENSBTAG00000007356	-1	ELF1	12	11183378	11269367	11269367
ENSBTAG00000007374	-1	LHFPL2	10	9500429	9527328	9527328
ENSBTAG00000007390	-1	VAT1	19	43687403	43694870	43694870
ENSBTAG00000007421	1	CDH5	18	34260148	34274741	34260148
ENSBTAG00000007806	-1	MTPN	4	100550308	100626568	100626568
ENSBTAG00000007808	1	ANTXR1	11	67334898	67590481	67334898
ENSBTAG00000007814	1	WWTR1	1	119437952	119583580	119437952
ENSBTAG00000007909	1	NOTCH2	3	23307157	23478243	23307157
ENSBTAG00000008022	-1	PTGFRN	3	26369811	26414580	26414580
ENSBTAG00000008140	1	FAP	2	34297726	34376828	34297726
ENSBTAG00000008202	1	PRKCSH	7	17034684	17048121	17034684
ENSBTAG00000008283	1	FLI1	29	32664912	32725664	32664912
ENSBTAG00000008300	-1	FN1	2	103881402	103950562	103950562
ENSBTAG00000008403	1	ROCK1	24	35445610	35519082	35445610

ENSBTAG00000008411	-1	PLEKHO2	10	45281371	45303690	45303690
ENSBTAG00000008483	1	CALCRL	2	8901432	9030989	8901432
ENSBTAG00000008585	1	ARHGEF10	27	219420	264986	219420
ENSBTAG00000008733	-1	MAGED1	X	95708020	95714844	95714844
ENSBTAG00000008792	-1	RNASE6	10	26402507	26404000	26404000
ENSBTAG00000008817	1	LAMA4	9	38644310	38810902	38644310
ENSBTAG00000009020	1	CRIM1	11	18835047	19041904	18835047
ENSBTAG00000009237	-1	KDELR2	25	38871560	38886919	38886919
ENSBTAG00000009502	1	LUZP1	2	130360072	130363164	130360072
ENSBTAG00000009513	1	TGFBI	7	49059242	49095515	49059242
ENSBTAG00000009565	1	RASA1	7	89281002	89391377	89281002
ENSBTAG00000009580	-1	SH3BGRL3	2	127429585	127431204	127431204
ENSBTAG00000009615	1	ANXA2	10	49860062	49904536	49860062
ENSBTAG00000009665	1	UTRN	9	82762003	83311756	82762003
ENSBTAG00000009705	1	SERPINF1	19	23422880	23430667	23422880
ENSBTAG00000009717	1	FGL2	4	44236254	44240215	44236254
ENSBTAG00000009761	1	ACTR2	11	63552997	63591980	63552997
ENSBTAG00000009886	1	KDELR3	5	110660391	110673555	110660391
ENSBTAG00000009998	1	GALNT16	10	81396104	81494769	81396104
ENSBTAG00000010050	1	COL16A1	2	122557458	122611951	122557458
ENSBTAG00000010179	-1	COL5A3	7	15769155	15811269	15811269
ENSBTAG00000010395	-1	DOCK9	12	79691579	79870108	79870108
ENSBTAG00000010529	-1	FZD6	14	63358656	63392777	63392777
ENSBTAG00000010562	1	CD34	16	77367502	77389361	77367502
ENSBTAG00000010587	-1	SH3BGRL	X	70399116	70527059	70527059
ENSBTAG00000010719	-1	ANGPTL1	16	61657545	61681733	61681733
ENSBTAG00000010726	1	F8	X	38838455	38982287	38838455
ENSBTAG00000010793	-1	CCDC80	1	57819993	57855630	57855630
ENSBTAG00000010830	-1	LRP1	5	56562309	56641157	56641157
ENSBTAG00000010888	-1	VSIR	28	28054082	28079127	28079127
ENSBTAG00000010899	1	TIMP2	19	54079297	54131052	54079297
ENSBTAG00000011001	-1	ERG	1	152379404	152515109	152515109
ENSBTAG00000011125	-1	MYO9B	7	5804898	5913540	5913540
ENSBTAG00000011146	-1	RAB8B	10	46837292	46907334	46907334
ENSBTAG00000011215	1	ACTN4	18	48668482	48741185	48668482
ENSBTAG00000011226	1	SLC4A2	4	114438006	114450606	114438006
ENSBTAG00000011256	1	MYO1B	2	80166217	80369680	80166217
ENSBTAG00000011284	1	SLC39A1	3	16530784	16534816	16530784
ENSBTAG00000011383	-1	SNX4	1	70504978	70557887	70557887
ENSBTAG00000011400	-1	DBN1	7	40312121	40325975	40325975
ENSBTAG00000011425	-1	PTPRA	13	52600919	52771800	52771800
ENSBTAG00000011454	1	FKBP10	19	42650885	42655321	42650885
ENSBTAG00000011494	-1	PYGL	10	43800152	43840994	43840994
ENSBTAG00000011531	1	SS18	24	31020509	31096441	31020509
ENSBTAG00000011559	1	RPL7A	11	104311808	104315125	104311808
ENSBTAG00000011578	1	CD44	15	66454331	66541790	66454331
ENSBTAG00000011578	1	CD44	15	66454331	66541790	66454413
ENSBTAG00000011578	1	CD44	15	66454331	66541790	66489338

ENSBTAG00000011613	1	PLS3	X	71709928	71806536	71709928
ENSBTAG00000011613	1	PLS3	X	71709928	71806536	71709936
ENSBTAG00000011770	1	TJP2		8 45598159	45696886	45598159
ENSBTAG00000011770	1	TJP2		8 45598159	45696886	45649371
ENSBTAG00000011789	1	REST		6 73838854	73863778	73838854
ENSBTAG00000011802	1	COL6A1		1 147404396	147423644	147404396
ENSBTAG00000011824	1	OGN		8 85453019	85468721	85453019
ENSBTAG00000011824	1	OGN		8 85453019	85468721	85453132
ENSBTAG00000011851	1	FYN		9 39107807	39259883	39107807
ENSBTAG00000011963	-1	RPS19		18 51689627	51697161	51697161
ENSBTAG00000012022	-1	TBC1D2B		21 30902155	30968114	30968114
ENSBTAG00000012044	1	RPL13		18 14533161	14535556	14533161
ENSBTAG00000012066	-1	PECAM1		19 49175892	49238414	49238414
ENSBTAG00000012088	1	FBLN1		5 116616200	116695692	116616200
ENSBTAG00000012152	-1	NONO	X	84662118	84675255	84675255
ENSBTAG00000012152	-1	NONO	X	84662118	84675255	84675194
ENSBTAG00000012191	1	WWC2		27 13139716	13204496	13139716
ENSBTAG00000012247	1	MXRA8		16 52424102	52428456	52424102
ENSBTAG00000012305	-1	ADGRL2		3 62643353	62805720	62805720
ENSBTAG00000012305	-1	ADGRL2		3 62643353	62805720	62727738
ENSBTAG00000012342	1	LIMA1		5 29804689	29898282	29804689
ENSBTAG00000012387	1	PAM		7 104321898	104501791	104321898
ENSBTAG00000012387	1	PAM		7 104321898	104501791	104321915
ENSBTAG00000012442	1	CTSB		8 7414945	7423429	7414945
ENSBTAG00000012505	1	ARHGEF17		15 53582241	53643983	53582241
ENSBTAG00000012817	-1	JAG1		13 3832286	3876681	3876681
ENSBTAG00000012847	-1	FAM129B		11 98232719	98251825	98251825
ENSBTAG00000012849	-1	COL4A1		12 88876125	89009422	89009422
ENSBTAG00000013004	1	ITIH5		13 16278275	16404464	16278275
ENSBTAG00000013016	-1	GNAI3		3 33969546	34013930	34013930
ENSBTAG00000013060	-1	IQGAP1		21 22530902	22614701	22614701
ENSBTAG00000013093	1	ALDH3B1		29 46188321	46203015	46188321
ENSBTAG00000013093	1	ALDH3B1		29 46188321	46203015	46188329
ENSBTAG00000013103	1	COL1A1		19 37088246	37104998	37088246
ENSBTAG00000013222	1	CD109		9 13421249	13551857	13421249
ENSBTAG00000013363	-1	CAP1		3 106638795	106667878	106667878
ENSBTAG00000013367	1	PPT1		3 106622151	106638462	106622151
ENSBTAG00000013369	1	COL14A1		14 83892853	84109620	83892853
ENSBTAG00000013461	-1	RPL24		1 46415223	46420721	46420721
ENSBTAG00000013472	1	COL1A2		4 11624470	11661163	11624470
ENSBTAG00000013478	1	MARVELD1		26 18732857	18740515	18732857
ENSBTAG00000013527	-1	PGD		16 44043496	44058198	44058198
ENSBTAG00000013530	1	DDAH2		23 27399306	27402825	27399306
ENSBTAG00000013662	1	COL8A1		1 43541936	43717619	43541936
ENSBTAG00000013745	1	ITGA5		5 25778012	25799053	25778012
ENSBTAG00000013755	-1	ITGB5		1 69801844	69899676	69899676
ENSBTAG00000013773	-1	PKP4		2 37815459	37929695	37929695
ENSBTAG00000013824	-1	WWC3	X	142845107	142966304	142966304

ENSBTAG00000013843	-1	ACVRL1	5	28097837	28106733	28106733
ENSBTAG00000013866	-1	RPS27	3	16505732	16507247	16507247
ENSBTAG00000013953	1	CALD1	4	99475015	99580189	99475015
ENSBTAG00000014059	1	ATP2B4	16	1272937	1313257	1272937
ENSBTAG00000014191	1	QSOX1	16	62804447	62845817	62804447
ENSBTAG00000014208	-1	RPL35A	1	70788697	70792140	70792140
ENSBTAG00000014226	-1	RPL34	6	17828130	17832711	17832711
ENSBTAG00000014324	-1	ANTXR2	6	96346732	96526016	96526016
ENSBTAG00000014377	1	CHD3	19	28185822	28206741	28185822
ENSBTAG00000014451	-1	YEATS2	1	83927442	84030074	84030074
ENSBTAG00000014471	1	AKAP2	8	101361501	101406330	101361501
ENSBTAG00000014518	-1	RPL9	6	60210361	60215120	60215120
ENSBTAG00000014648	1	RPN2	13	66799465	66853779	66799465
ENSBTAG00000014665	1	ADAMTS2	7	1956352	2165242	1956352
ENSBTAG00000014713	1	RARRES1	1	109661962	109704520	109661962
ENSBTAG00000014782	-1	STAB1	22	48884238	48909562	48909562
ENSBTAG00000014824	-1	MMP14	10	21806054	21814533	21814533
ENSBTAG00000014841	1	GBA	3	15445102	15463930	15445102
ENSBTAG00000014933	-1	TRAK2	2	90355099	90397117	90397117
ENSBTAG00000014933	-1	TRAK2	2	90355099	90397117	90380410
ENSBTAG00000015147	1	S100A10	3	18799612	18810545	18799612
ENSBTAG00000015283	1	RPL32	22	56985249	56989012	56985249
ENSBTAG00000015285	-1	RPS8	3	101816844	101818956	101818956
ENSBTAG00000015296	1	PTPRB	5	43114782	43239445	43114782
ENSBTAG00000015327	1	SPTAN1	11	99131533	99179460	99131533
ENSBTAG00000015363	1	CDC42SE1	3	19768810	19776435	19768810
ENSBTAG00000015388	-1	RPL18	18	55710193	55713956	55713956
ENSBTAG00000015398	-1	TJP1	21	28934158	28992880	28992880
ENSBTAG00000015405	1	DCHS1	15	47042652	47064122	47042652
ENSBTAG00000015438	-1	RRBP1	13	38289297	38317466	38317466
ENSBTAG00000015457	-1	FGFR1	27	33250534	33291989	33291989
ENSBTAG00000015473	1	RPS27A	11	37823446	37825428	37823446
ENSBTAG00000015527	1	MYO1D	19	17665144	18023500	17665144
ENSBTAG00000015541	1	DLC1	27	22814893	22993901	22814893
ENSBTAG00000015549	1	PCDH18	17	20602272	20616261	20602272
ENSBTAG00000015580	-1	TLE3	10	16901655	16949714	16949714
ENSBTAG00000015598	-1	RPS10	23	8434516	8441522	8441522
ENSBTAG00000015739	1	MRC2	19	47721711	47747246	47721711
ENSBTAG00000015802	-1	CREB3L2	4	102401898	102529709	102529709
ENSBTAG00000015831	-1	RPL18A	7	5206112	5209504	5209504
ENSBTAG00000015910	1	ITGB1	13	20248945	20292114	20248945
ENSBTAG00000015910	1	ITGB1	13	20248945	20292114	20248978
ENSBTAG00000015978	1	ANXA1	8	49624473	49642916	49624473
ENSBTAG00000016152	1	DAB2	20	35018908	35079162	35018908
ENSBTAG00000016224	-1	RPS7	8	112896493	112901718	112901718
ENSBTAG00000016278	1	RPL30	14	68467563	68470952	68467563
ENSBTAG00000016420	1	CTNNB1	22	13842703	13889468	13842703
ENSBTAG00000016462	-1	TCF4	24	55056433	55435276	55435276

ENSBTAG00000016462	-1	TCF4	24	55056433	55435276	55161459
ENSBTAG00000016467	1	SLC22A17	10	21409116	21414758	21409116
ENSBTAG00000016469	-1	TNKS1BP1	15	81792138	81813276	81813276
ENSBTAG00000016525	-1	ITGA1	20	26116747	26227530	26227530
ENSBTAG00000016593	1	FMNL3	5	30320681	30372554	30320681
ENSBTAG00000016720	-1	RAB1A	11	63429725	63458565	63458565
ENSBTAG00000016720	-1	RAB1A	11	63429725	63458565	63458558
ENSBTAG00000016740	-1	ACLY	19	42691369	42735634	42735634
ENSBTAG00000016750	-1	PTBP3	8	103339338	103383299	103383299
ENSBTAG00000016751	1	MYO6	9	15670077	15763485	15670077
ENSBTAG00000016813	1	SH3D19	17	6636917	6684756	6636917
ENSBTAG00000016822	1	PPIB	10	45874978	45880918	45874978
ENSBTAG00000016846	1	YWHAB	13	74018886	74040819	74018886
ENSBTAG00000016894	-1	CYFIP1	2	996692	1104749	1104749
ENSBTAG00000016918	-1	MYOF	26	14667438	14848281	14848281
ENSBTAG00000016956	-1	GANAB	29	41664315	41679641	41679641
ENSBTAG00000016984	1	PTPN9	21	33697444	33766711	33697444
ENSBTAG00000017079	1	MFAP3	7	67331187	67338606	67331187
ENSBTAG00000017122	1	HSPG2	2	131517579	131587498	131517579
ENSBTAG00000017129	1	CLCC1	3	34511153	34544280	34511153
ENSBTAG00000017135	1	CTSS	3	20024302	20047228	20024302
ENSBTAG00000017143	-1	PDIA4	4	112989514	113006466	113006466
ENSBTAG00000017165	-1	MATN2	14	68478180	68651684	68651684
ENSBTAG00000017196	1	PDIA3	21	55924230	55946434	55924230
ENSBTAG00000017266	-1	ITGA6	2	24131486	24217715	24217715
ENSBTAG00000017339	1	RUNX1T1	14	74642806	74787356	74642806
ENSBTAG00000017339	1	RUNX1T1	14	74642806	74787356	74726178
ENSBTAG00000017349	1	PCDHGC3	7	54152475	54281944	54152475
ENSBTAG00000017349	1	PCDHGC3	7	54152475	54281944	54157792
ENSBTAG00000017349	1	PCDHGC3	7	54152475	54281944	54248918
ENSBTAG00000017382	1	P3H1	3	104065149	104084575	104065149
ENSBTAG00000017448	-1	EFEMP1	11	38338744	38408288	38408288
ENSBTAG00000017465	1	GNS	5	49283646	49333344	49283646
ENSBTAG00000017753	1	APP	1	9607382	9921004	9607382
ENSBTAG00000017788	1	AKT3	16	34132648	34404652	34132648
ENSBTAG00000017830	-1	RBMS2	5	57167137	57219971	57219971
ENSBTAG00000017846	1	F11R	3	8483556	8508122	8483556
ENSBTAG00000017869	-1	CAV1	4	52173110	52208687	52208687
ENSBTAG00000017970	1	ZYX	4	107598856	107607834	107598856
ENSBTAG00000018013	1	EMP3	18	55482512	55486850	55482512
ENSBTAG00000018052	1	PTPRS	7	20191965	20256722	20191965
ENSBTAG00000018123	-1	FBLN5	21	57153110	57246389	57246389
ENSBTAG00000018152	1	MYADM	18	62018419	62024004	62018419
ENSBTAG00000018271	1	SLC38A10	19	52045287	52085223	52045287
ENSBTAG00000018373	1	DPYSL2	8	75089346	75168436	75089346
ENSBTAG00000018374	1	WASF2	2	126399312	126469335	126399312
ENSBTAG00000018463	1	VIM	13	31945012	31952941	31945012
ENSBTAG00000018744	-1	MGAT5	2	63118922	63362056	63362056

ENSBTAG00000018784	1	CTSZ	13	57889707	57899205	57889707
ENSBTAG00000019147	-1	RPS20	14	24955079	24956324	24956324
ENSBTAG00000019237	-1	TXNDC5	23	47333403	47348212	47348212
ENSBTAG00000019251	-1	EPB4IL3	24	39312037	39404644	39404644
ENSBTAG00000019269	1	COL6A2	1	147542825	147572432	147542825
ENSBTAG00000019269	1	COL6A2	1	147542825	147572432	147570799
ENSBTAG00000019338	-1	IL13RA1	X	3060168	3100688	3100688
ENSBTAG00000019339	-1		16	30390092	30490348	30490348
ENSBTAG00000019456	-1	SPRED2	11	63639742	63672178	63672178
ENSBTAG00000019494	1	RPL10A	23	9391523	9394013	9391523
ENSBTAG00000019526	-1	CMTM6	22	7001971	7024065	7024065
ENSBTAG00000019612	-1	RNASE4	10	26423874	26445617	26445617
ENSBTAG00000019627	-1	THY1	15	30509504	30515399	30515399
ENSBTAG00000019630	1	RGL1	16	66355358	66439715	66355358
ENSBTAG00000019644	1	EXT2	15	75074940	75265705	75074940
ENSBTAG00000019704	1	HLTF	1	120013765	120075520	120013765
ENSBTAG00000019718	1	RPS15	7	45465834	45467519	45465834
ENSBTAG00000019733	1	ADGRL4	3	65921137	66062039	65921137
ENSBTAG00000019755	1	REEP3	28	19709521	19806763	19709521
ENSBTAG00000019866	1	NRP1	13	19911857	20056980	19911857
ENSBTAG00000019877	1	DOCK7	3	83494253	83658175	83494253
ENSBTAG00000019915	1	GSN	8	112578066	112639758	112578066
ENSBTAG00000019915	1	GSN	8	112578066	112639758	112609393
ENSBTAG00000020046	-1	CLMP	15	34226369	34332243	34332243
ENSBTAG00000020139	-1	RPL7	14	38741347	38744856	38744856
ENSBTAG00000020148	-1	TEK	8	17040335	17143857	17143857
ENSBTAG00000020281	-1	NIN	10	43640454	43692975	43692975
ENSBTAG00000020345	1	CNN3	3	48763975	48794136	48763975
ENSBTAG00000020421	1	SUPT16H	10	25827154	25861850	25827154
ENSBTAG00000020480	-1	SPTLC2	10	89756991	89852261	89852261
ENSBTAG00000020528	-1	PCOLCE	25	36490771	36495789	36495789
ENSBTAG00000020645	-1	GNAI2	22	50670852	50691007	50691007
ENSBTAG00000020717	1	CXHXorf36	X	103969315	104013030	103969315
ENSBTAG00000020733	-1	RPS15A	25	16531761	16537583	16537583
ENSBTAG00000020795	-1	RPS21	13	55358720	55359979	55359979
ENSBTAG00000020894	1	LAPTM4A	11	78862495	78880461	78862495
ENSBTAG00000020905	-1	RPL11	2	129791372	129795563	129795563
ENSBTAG00000020935	1	HIF1A	10	74095881	74139364	74095881
ENSBTAG00000020944	-1	PIEZO1	18	13984761	14002517	14002517
ENSBTAG00000021035	1	CTSK	3	19994998	20007861	19994998
ENSBTAG00000021093	-1	RPS16	18	49393725	49396191	49396191
ENSBTAG00000021191	1	EHD2	18	55071102	55087454	55071102
ENSBTAG00000021307	1	BNIP3L	8	74924184	74947549	74924184
ENSBTAG00000021338	1	OAF	15	31312383	31330638	31312383
ENSBTAG00000021381	1	DAAM2	23	13775149	13829110	13775149
ENSBTAG00000021455	-1	CFL1	29	44638896	44642280	44642280
ENSBTAG00000021457	-1	EFEMP2	29	44650628	44657896	44657896
ENSBTAG00000021466	-1	COL3A1	2	7318227	7356937	7356937

ENSBTAG00000021466	-1	COL3A1	2	7318227	7356937	7344001
ENSBTAG00000021467	-1	IGFBP6	5	27044009	27047853	27047853
ENSBTAG00000021543	-1	MDFIC	4	53756424	53854749	53854749
ENSBTAG00000021602	-1	CTTNBP2NL	3	30987502	31026962	31026962
ENSBTAG00000021617	-1	ZC3HAV1	4	103472028	103520426	103520426
ENSBTAG00000021675	-1	PJA2	7	111045732	111095494	111095494
ENSBTAG00000021697	-1	PDGFB	5	111180170	111200490	111200490
ENSBTAG00000021697	-1	PDGFB	5	111180170	111200490	111200409
ENSBTAG00000021746	1	ANXA5	6	3542635	3575330	3542635
ENSBTAG00000021771	-1	PTTG1IP	1	145057828	145075459	145075459
ENSBTAG00000021778	-1	SELENON	2	127851648	127869417	127869417
ENSBTAG00000021799	1	RCN3	18	56422333	56430760	56422333
ENSBTAG00000021819	-1	IFNAR1	1	1467704	1496151	1496151
ENSBTAG00000021911	-1	PTPRG	22	39175038	40360572	39502562
ENSBTAG00000021911	-1	PTPRG	22	39175038	40360572	40360572
ENSBTAG00000021919	1	NAV1	16	49447984	49564506	49447984
ENSBTAG00000021919	1	NAV1	16	49447984	49564506	49540151
ENSBTAG00000021920	-1	SEMA4C	11	2784990	2793644	2793644
ENSBTAG00000021945	-1	NID2	10	44894657	44986659	44986659
ENSBTAG00000021955	-1	NPC2	10	86170653	86179237	86179237
ENSBTAG00000021977	-1	PRRC1	7	27685132	27723230	27723230
ENSBTAG00000022155	-1	FSTL1	1	65742633	65802423	65802423
ENSBTAG00000022169	1	PREX2	14	34040796	34340573	34040796
ENSBTAG00000022278	1		X	85898898	85899424	85898898
ENSBTAG00000022777	-1	CDC42BPA	16	30708032	31000733	31000733
ENSBTAG00000023343	-1	RPL28	18	62547220	62549950	62549950
ENSBTAG00000023652	-1	PROS1	1	37803108	37866950	37866950
ENSBTAG00000023907	1	COL18A1	1	146989244	147040968	146989244
ENSBTAG00000024081	-1	ECM2	8	85540501	85579683	85579683
ENSBTAG00000024909	1	H3F3B	19	56453856	56455637	56453856
ENSBTAG00000025029	1	MAN2A1	7	111396329	111604883	111396329
ENSBTAG00000026327	-1	RPL8	14	1505030	1507633	1507633
ENSBTAG00000027020	1	COL5A2	2	7139738	7298551	7139738
ENSBTAG00000027684	1	FOLR2	15	52602819	52605536	52602819

