UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

DETERMINAÇÃO DO PERFIL TRANSCRIPTÔMICO ASSOCIADOS À EFICIÊNCIA ALIMENTAR DE BOVINOS NELORE (*BOS TAURUS INDICUS*)

Marta Serna García

2022

DETERMINAÇÃO DO PERFIL TRANSCRICIONAL ASSOCIADO À EFICIÊNCIA ALIMENTAR DE BOVINOS NELORE (*BOS TAURUS INDICUS*)

Discente: Marta Serna García

Orientador: Profa. Dra. Lucia Galvão de Albuquerque

Co-orientadores: Profa. Dra. Larissa Fernanda Simielli Fonseca, Prof. Dr. Mateus José Rodrigues Paranhos da Costa and Prof. Dr Julian Carretero Asuncion

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutora em Genética e Melhoramento Animal.

2022

G216d

García, Marta Serma

Determinação do perfil transcriptômico associados à eficiência alimentar de bovinos Nelore (Bos taurus indicus) / Marta Serma García. -- Jaboticabal, 2022

180 p. : il., tabs., fotos

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientadora: Lucia Galvão de Albuquerque

Coorientadora: Julian Carretero Asuncion

1. Bovino. 2. Car. 3. Genes. 4. Ncrna. 5. Splicing alternativo. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

unesp

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: DETERMINAÇÃO DO PERFIL TRANSCRIPTÔMICO ASSOCIADOS À EFICIÊNCIA ALIMENTAR DE BOVINOS NELORE (*BOS TAURUS INDICUS*)

AUTORA: MARTA SERNA GARCIA
ORIENTADORA: LUCIA GALVÃO DE ALBUQUERQUE
COORIENTADOR: MATEUS JOSÉ RODRIGUES PARANHOS DA COSTA
COORIENTADORA: LARISSA FERNANDA SIMIELLI FONSECA
COORIENTADOR: JULIAN CARRETERO ASUNCION

Aprovada como parte das exigências para obtenção do Título de Doutora em GENÉTICA E MELHORAMENTO ANIMAL, pela Comissão Examinadora:

Pós-doutoranda LARISSA FERNANDA SIMIELLI FONSECA (Participaçao Virtual)
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal

Pesquisadora Dra. ANA FABRÍCIA BRAGA MAGALHÃES (Participaçao Virtual)
Universidade Federal dos Vales do Jequitinhonha e Mucuri - Campus JK / Diamantina/MG

Profa. Dra. DANIELLY BERALDO DOS SANTOS SILVA (Participaçao Virtual)
Universidade José do Rosário Vellano-UNIFENAS / Alfenas/MG

Profa. Dra. NICLA FLACCO (Participaçao Virtual)
Universidad Europea de Valencia / Valencia/Espanha

Prof. Dr. GREGORIO MIGUEL FERREIRA DE CAMARGO (Participaçao Virtual)
Departamento de Zootecnia / Escola de Medicina Veterinária e Zootecnia da Universidade Federal da Bahia (UFBA) / Salvador/BA

Jaboticabal, 29 de abril de 2022

DADOS CURRICULARES DA AUTORA

Marta Serna García nasceu em 15 de fevereiro de 1991, em Valência (Espanha), filha de Ana Victoria García Tarín e Manuel Serna Bañuelos. Em setembro de 2016 finalizou a graduação em Engenharia Agroalimentar e Rural da *Universitat Politécnica de València*, UPV, Espanha. Durante a graduação de 2015-2016, recebeu uma bolsa da prefeitura de Valencia e da *Universitat Politécnica de València*, Espanha, tendo sido concedidas 45 bolsas, na modalidade. Foi orientada pelo Dr. Pascual Medina Besso na plataforma de Análise Multigênica *(Affymetrix*) da Unidade Central de Investigação da Faculdade de Medicina da *Universidad de València*, Espanha, continuando nesta plataforma até o ano de 2017.

Em maio de 2016, iniciou o curso mestrado no Programa de Pós-Graduação em Produção Animal da *Universitat Politècnica de València*, Espanha. Neste período recebeu uma bolsa Européia de mestrado concedida no âmbito do "*Erasmus Mundus Programme*" (Projeto EuroInkaNet) e obteve o título de Mestra em julho de 2018, sob orientação da Profa. Dra. María Antonia Santacreu Jerez e da Profa. Dra. Lucia Galvão de Albuquerque.

Em agosto de 2018 o iniciou o curso de doutorado no Programa de Pós-Graduação em Genética e Melhoramento Animal da Universidade Estadual "*Júlio de Mesquita Filho*" campus de Jaboticabal (FCAV- UNESP), sob orientação do Prof. Dr. Mateus José Rodrigues Paranhos da Costa. Em 2021 passou a ser orientada pela Profa. Dra. Lucia Galvão de Albuquerque. Foi bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), tendo como linha de pesquisa genética molecular e seleção genômica. Nesse período, colaborou com a *Universidad de Valencia*, Espanha (UV) sob orientação Prof. Dr. Jose Viña e realizou 219 horas de docência universitária durante este período.

Desde setembro de 2021 é professora, em tempo integral, na *Universidad Europea de València*, Espanha.

"*Se debe aprender a navegar en océanos de incertidumbres a través de archipiélagos de certeza*."

(Edgar Morin, 1921)

"Aos meus pais (Ana e Manuel), pelo apoio e amor. À minha irmã Eva, por compartilhar a ciência e ser a minha inspiração e referência"

"A todos os cientistas e à ciência por sua luta, especialmente nestes tempos e ao ser humano mais criativo graças à incerteza"

Dedico e Ofereço.

AGRADECIMENTOS

Agradeço à toda minha família e aos amigos por me apoiarem, em especial ao Jethro pela ajuda, paciência e muito amor.

Agradeço ao Joaquim Panadero por ter sido meu guia, por fazer ilusionar e amar ainda mais a ciência, por fazer o difícil fácil, pela inumerável paciência e delicadeza, por ser uma pessoa maravilhosa e por descrever o brilhantismo e profissionalismo com humildade.

Agradeço à Profª. Drª. Lucia Galvão de Albuquerque pela orientação, apoio e pela confiança. Obrigada por me receber, ser um exemplo de profissionalismo e dedicação.

Agradeço ao Profª. Dr. Mateus Paranhos da Costa pela coorientação, pela confiança e atenção.

Agradeço à Profª. Drª. Larissa Fonseca pela coorientação, pela ajuda e apoio profissionalmente e mostrar qualidade humana.

Agradeço ao Profª. Drª. Julian Carretero Asunción pela coorientação, pela ajuda no processo.

Agradeço ao Prof. Dr. Jesus Ferro, Dra Rosa Peiró pelas contribuições bem como à Drª. Maria Antonia Santacreu Jerez por me ensinar à amar o melhoramento animal.

Agradeço ao Dr. José Viña pela oportunidade ao me receber em seu grupo de pesquisa na Espanha, pelas contribuições e ensino.

Agradeço à banca examinadora (qualificação e defesa) pelas contribuições e apontamentos que foram de fundamental importância.

A todos os amigos do Departamento de Zootecnia, especialmente aos colegas de sala e aos meus queridos amigos: Ana, Andres, Bruna, Danielle, Diogo, Flavio, Gabriela, Malane, Natalia, Patricia, Thaise. Obrigada não só pelo auxílio nas pesquisas, como também pela amizade, pelos momentos e apoio moral.

A todos os professores, de Brasil e Espanha. Também aos técnicos administrativos e demais funcionários da Universidade.

À Universidad Europea de Valencia, en especial Dra. Nicla Flacco pelo apoio e ajuda.

À FCAV-UNESP e ao Programa de Pós-graduação em Genética e Melhoramento Animal pelo apoio logístico.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP nº dos processos #2009/16118-5, #2017/10630-2 e #2018/20026-8)

Summary

DETERMINAÇÃO DO PERFIL TRANSCRICIONAL ASSOCIADO À EFICIÊNCIA ALIMENTAR DE BOVINOS NELORE (BOS TAURUS INDICUS)

RESUMO – A população mundial deverá atingir 9,8 milhões pessoas em 2050. Essa tendência ameaça a sustentabilidade dos sistemas e preocupa quanto a capacidade de suprir necessidades alimentares do planeta. Devido à crescente demanda dos produtos pecuários, é necessário estar preparados para os impactos na produção, mas fazê-lo de maneira inclusiva e sustentável exigirá grandes transformações. O consumo alimentar residual (CAR) permite identificar e selecionar animais mais eficientes sem prejuízo do crescimento e reprodução, ou da qualidade da carne. Além disso, o aumento da eficiência alimentar diminui o impacto ambiental, uma vez que animais mais eficientes tendem a consumir menos e produzir mais quando comparados aos menos eficientes. A característica é de difícil mensuração e a busqueda pelo entendimento genético e biológico se tornou uma alternativa na investigação de animais superiores, e dentre as técnicas moleculares, o RNA-Seq, tem sido uma das mais utilizadas, pois permite o sequenciamento e a quantificação dos transcritos com uma grande resolução, detecção de eventos de *splicing* alternativo e ncRNAs (*non-coding* RNA), que podem ter diversas funções na arquitetura regulatória. Visando entender melhor os processos genéticos envolvidos na expressão do Consumo Alimentar Residual (CAR), foram analisados dois grupos genéticos (G1 e G2) de bovinos Nelore. No G1, animais (n=60) do mesmo grupo de contemporâneos foram submetidos a uma avaliação do CAR e dentre eles, 24 animais extremos para a característica foram abatidos. Amostras de tecido hepático foram coletadas e sequenciadas por meio da plataforma HiSeq 2500 System (Illumina). Para G2, foram utilizados dados públicos, provenientes do sequenciamento de 20 amostras de tecido hepático de animais divergentes para CAR, também do mesmo grupo de contemporâneos, obtidos por meio HiSeq 2000 System (Illumina). Os objetivos do estudo foram identificar os genes diferencialmente expressos, descrever eventos de *splicing* alternativo, traçar um perfil dos transcritos de ncRNA comparando duas populações de Nelore (*Bos taurus indicus*), classificados de forma divergente para CAR, em tecido hepático usando RNA-Seq. Foram encontrados 1.811 genes de expressão diferencial (DEGs) e 2.054 DEGs (p-valor 0,05) para G1 e G2, respectivamente. Quanto aos DEG comumente identicados nas duas populações, foram encontrados 88, dos quais 33 estavam envolvidos na resposta imune e no bloqueio do estresse oxidativo. Foram encontrados sete (B2M, ADSS, SNX2, TUBA4A, ARHGAP18, MECR, ABCF3) possíveis biomarcadores, considerando AUC > 0,70. O gene B2M foi mais expresso no grupo mais eficiente [CAR negativo] comparado com o grupo ineficiente [CAR positivo] atua regulando o *turnover* proteico, metabolismo lipídico e inibe a morte celular. Em ambos os grupos genéticos, o grupo CAR negativo apresentou genes com capacidade antioxidante, maior capacidade de proliferação celular e proteção contra a morte. Os genes BoLA-DRA, BoLA-DRB3 e CD74 foram preditos como *hub*, e estão envolvidos na melhoria da resposta imune. Os fatores de transcrição LRRFIP1 e NR4A2 foram detectados e estão envolvidos na promoção da angiogênese e diminuição da apoptose. A análise de redes de interação física mediante dados públicos (BioGRID) detectou 7 possíveis nodes; DDB2, ACTB, PAK1, VIM, PRC1, RPS6KA1 e PKM regulando positivamente proliferação celular, lipólise, organização do citoesqueleto e sobrevivência celular no grupo CAR negativo. No capítulo 4, com os mesmos animais do capítulo 2 (fenotipicamente divergentes para RFI), foram identificados eventos de *splicing* alternativos diferencialmente expressos (DAS) a partir da abordagem *exon-centric*. Os DAS encontrados foram transcritos para 67 e 75 genes (p-value ≤ 0,05) respectivamente para G1 e para G2. Identificamos 9 comuns genes DAS, no geral, desempenham um papel no metabolismo lipídico e do sistema imunológico, destacando SAT1 como gene hub. Os RNAs não codificantes significativos encontrados foram; MIR25 and SNORD16 para G1 RNase_MRP e SCARNA10 e para G2 em RLFI. Destacamos o MIR25 sendo capaz de atuar bloqueando a citotoxicidade e o estresse oxidativo e o RNase_MRP como bloqueador do dano mitocondrial. Os resultados relatados neste estudo indicam genes candidatos e mecanismos moleculares reguladores do CAR em dois grupos geneticos em bovinos Nelore, o que irá melhorar a compreensão da eficiência alimentar, auxiliar em futuras abordagens dos métodos quantitativos de melhoramento para aumentar da produtividade e a diminuição do impacto ambiental da pecuária através da identificação de animais com alta eficiência alimentar com rapidez e menor custo.

Palavras-Chave: bovino, CAR, genes, *Hub* gene, ncRNA, *nodes*, *splicing* alternativo

DETERMINATION OF THE TRANSCRIPTOME PROFILE ASSOCIATED WITH THE FEED EFFICIENCY OF NELORE CATTLE (BOS TAURUS INDICUS)

Abstract - The world population is expected to reach 9.8 million by 2050. This trend threatens the sustainability of the systems in place and concerns the ability to meet the requirements of our planet. Due to the growing demand for livestock products, it is necessary to be prepared for the impact on production but doing so in an inclusive and sustainable way will require major transformations. Residual Feed Intake (RFI) indicates how to identify and select more efficient animals without detriment to growth and reproduction or compromising meat quality. In addition, increased feed efficiency reduces environmental impact, as more efficient animals tend to consume less and produce more compared to less efficient ones. The search for genetic and biological understanding of this type of characteristics has become an alternative in the investigation of more efficient animals and, among molecular techniques, RNA-seq is one of the most used. It allows sequencing and quantifying transcripts with high resolution, detection of alternative splicing events, and ncRNAs (non-coding RNAs), which may have multiple roles in regulatory architecture. To gain a better understanding of the genetic processes involved in RFI expression, two genetic groups (G1 and G2) of Nelore cattle were analyzed. In G1, animals (n = 60) from the same contemporaneous group underwent RFI evaluation. Twenty-four extreme animals were selected for this population, low RFI (N = 12) and high RFI (N = 12); they were subsequently slaughtered. Liver tissue was collected and sequenced using the HiSeq 2500 System (Illumina). Public data was used for G2, which includes RNA-seq data from 20 samples of liver tissue from animals divergent for RFI. These animals belong to the same contemporary group and the data were obtained by the HiSeq 2000 system (Illumina). The study aimed to identify differentially expressed genes, alternative splicing events and the ncRNA transcript profile by comparing two populations of Nelore cattle (*Bos taurus indicus*) classified divergently for RFI using RNA-seq data. We found 1,811 differentially expressed genes (DEGs) and 2,054 DEGs (p-value 0.05) in low RFI (LRFI) group compared with high RFI (HRFI) group for G1 and G2, respectively. Analyses found 88 common genes between the genetic groups, and 33 of these were involved in immune response and blocking of oxidative stress. Seven possible markers (B2M, ADSS, *SNX2, TUBA4A, ARHGAP18, MECR, ABCF3*) were found, considering AUC > 0.70. The *B2M* gene was overexpressed in the efficient group (LRFI) when compared with the inefficient group (HRFI); this gene acts by regulating protein turnover, lipid metabolism and inhibition of cell death. For both genetic groups, the LRFI group includes genes that could have an antioxidant capacity, increased cell proliferation capacity and protection against death. The *BoLA-DRA, BoLA-DRB3* and *CD74* have been predicted to be hub genes and are involved in enhancing the immune response. Transcription factors *LRRFIP1* and *NR4A2* are involved in the promotion of angiogenesis and decreasing apoptosis. Physical interaction network analysis using public data (BioGRID) detected seven possible nodes: *DDB2, ACTB, PAK1, VIM, PRC1, RPS6KA1* and *PKM.* In chapter 4, with the same animals selected in chapter 2 (phenotypically divergent for RFI), differentially expressed alternative splicing events (DAS) were identified from the exon-centric approach. The DAS found were transcribed for 67 and 75 genes, respectively for G1 and for G2. We identified 9 common DAS genes, in general, play a role in lipid metabolism and the immune system, highlighting SAT1 as the hub gene. The significant non-coding RNAs found were MIR25 and SNORD16 for G1 RNase_MRP and SCARNA10 and for G2 in RLFI group. We highlight MIR25 being able to act by blocking cytotoxicity and oxidative stress and RNase_MRP as a blocker of mitochondrial damage.These genes were positively regulated for cell proliferation, lipolysis, cytoskeletal organization and cell survival in the efficient group. The results reported in this study indicate candidate genes and regulatory molecular mechanisms for RFI in the two genetic groups of Nelore cattle, which will improve the understanding of feed efficiency, help in future approaches of quantitative breeding methods to increase productivity and reduce the environmental impact of livestock through rapid identification of animals with high feed efficiency and low cost.

Keywords: bovine, genes, Hub gene, ncRNA, nodes, RFI, splicing alternative

CHAPTER 1 – General Considerations

1. Introduction

The world's population is constantly growing and will be over 11 billion by the end of the century. Moreover, between 720-811 million people faced hunger in 2020 (FAO, 2019). Improving food production systems and promoting the sustainable use of natural resources are critical to achieve Sustainable Development Goal 2 (SDG2): a world without hunger, food insecurity and malnutrition and SDG 13, "Climate Action" (UN, 2020). Humanity faces the challenge of minimizing global climate change, mainly caused by greenhouse gas (GHG) emissions, which are currently the highest in history (PORTER et al., 2014). There is a global trend of a rise in demand for livestock products (WRIGHT et al., 2012; GERBER et al., 2013; ROJAS-DOWNING et al., 2017) which are fundamental to global food security (ROSEGRANT et al., 2009).

In the last 50 years, GHG emissions have almost doubled and will continue increasing (EDENHOFER et al., 2014). In fact, 57% of global GHG emissions are from the production of animal-based food, 25% being bovine meat production (FAO, 2019). South and Southeast Asian and South America were the largest emitters of production-based GHG emissions (XU et al., 2021). Cattle emit most of the enteric methane (77%) of global emissions (POORE; NEMECEK, 2018).
Sustainability of food systems covering the needs of the global population will be an enormous challenge that we have to face (HOLDEN et al., 2018). For all the above, we need to maintain a balance between productivity, food security and environmental preservation (WRIGHT et al., 2012). One of the alternatives is the search for more efficient animals.

Individual feed efficiency of cattle that induces differences in muscle growth and development (MEHRBAN et al., 2021) is controlled by multiple genes and/or isoforms (polygenic). A common proposal to increase this efficiency is selection through residual food intake (RFI) (BRANCO et al., 2009; GRION et al., 2014; CEACERO et al., 2016). Selection for RFI does not affect animal growth, reproduction or the meat quality but leads to animals with lower consumption and lower maintenance requirements (LANNA; DE ALMEIDA, 2003; KOCH et al., 1963; BASARAB et al., 2003; CARBERRY et al., 2012). As a result, meat production costs and environmental impact are reduced, since there is a decrease in pasture area for livestock production and in enteric methane emission (HEGARTY et al., 2007).

Despite the importance of RFI for animal production in zebu cattle, studies with

Nelore, especially using RNA-seq (ribonucleic acid sequencing), are limited because this trait is difficult and costly to measure and is influenced by different biological processes. The RNA-seq tool is powerful, robust and adaptable for measuring genome-wide transcription (CORCHETE et al., 2020). However, there is no consensus on the most appropriate methodology to guarantee the validity of the results in terms of robustness, accuracy and reproducibility, so this field is in continuous development (HRDLICKOVA; TOLOUE; TIAN, 2017; GÓMEZ-GALLEGO et al., 2018; CORCHETE et al., 2020; GERACI; SAHA; BIANCHINI, 2020).

There are several studies using RNA-seq for RFI in Nelore cattle, such as in muscle tissue (FIDELIS et al., 2017; MENEZES et al., 2020), blood (NASCIMENTO et al., 2015) and liver tissue (TIZIOTO et al., 2015; AL-HUSSEINI et al., 2016; CARVALHO et al., 2019; MUKIIBI et al., 2019, 2020). To our knowledge, this is the first study carried out to integrate RNA-seq data from different Nelore populations and it may allow us to provide a broader and more precise vision of the action mechanisms involved in the RFI trait. The RNA-seq technique also provides the possibility of studying alternative splicing events, non-coding RNAs and co-expression network analysis that can facilitate the detection of important biological pathways and the identification of central genes (hub or key players) (LANGFELDER; MISCHEL; HORVATH, 2013). Results from this work could help future approaches in quantitative methods of animal husbandry and management and allow the use of possible markers for selecting more efficient animals.

2. Objectives

2.1. General objective

To study liver tissue transcriptome in two genetic groups of Nelore cattle (*Bos taurus indicus*) classified by their RFI records, using the RNA-seq technique. In addition, to identify biological and cellular processes underlying the RFI trait assuming it may help to highlight possible biomarkers to be used in future studies and further explorations into the genetic architecture of this trait.

2.2. Specific objective

- To identify differentially expressed genes in liver tissue of Nelore cattle common to two genetic groups associated with the RFI trait and to understand the biology of the liver better.
- To conduct gene co-expression network analysis, prediction of hub genes and analysis of the nodes of physical interactions with a public database

(BioGRID) of Nelore cattle RFI trait in two genetic groups.

- To identify spliced genes differentially expressed in liver tissue of the Nelore breed (*Bos taurus indicus*) for the RFI trait in two genetic groups using data obtained by RNA-seq.
- To study non-coding RNA expressed in liver tissue of animals from both Nelore genetic groups divergent for the RFI trait using data obtained by RNA-seq.

3. Background

3.1. Residual Feed Intake (RFI)

The residual feed intake (RFI) methodology aims to assess feed efficiency. It has been used as a selection criterion for cattle, focusing on increasing the efficiency of individual feeding (BRANCO et al., 2009; GRION et al., 2014; CEACERO et al., 2016) without detriment to reproduction or deterioration of meat quality.

RFI is the difference between observed consumption and estimated consumption, through adjustments for average metabolic weight (BMR) and weight gain rate (AWG, kg/day) (KOCH et al., 1963). This is a difficult trait to measure compared to other determinations (MARTIN et al., 2021). It is calculated individually over an average period of 2.5 to 3 months in which animals remain in individual barns or automated collective bays, by keeping a daily record of the amount of feed ingested and rejected, and the average daily weight gain of each animal.

This trait allows identification and selection of more efficient animals (negative RFI) when compared with less efficient ones (positive RFI), since they ingest less feed than estimated for the same weight gain, thereby making it possible to select animals with lower consumption and lower maintenance requirements (KOCH et al., 1963; BASARAB et al., 2003; CARBERRY et al., 2012). Animals with low RFI (more efficient) emit about 28% less methane into the atmosphere (NKRUMAH et al., 2006). The most efficient animals produced less manure (N, P and K) than the most inefficient animals (BASARAB et al., 2003; HEGARTY et al., 2007; HAYES; LEWIN; GODDARD, 2013).

There is an individual variation in nutrient utilization efficiency among animals of the same breed, sex and age ingesting the same type of feed, which indicates genetic variations in nutrient utilization efficiency (RICHARDSON; HERD, 2004). In addition, RFI is a trait with moderate heritability (from 0,19 to 0,57) (DEL CLARO; MERCADANTE; VASCONCELOS SILVA, 2012; BERRY; CROWLEY, 2013;

SANTANA et al., 2014). For this reason, the traits related to feed efficiency are increasingly studied and gradually implemented in genetic improvement programs.

3.2. RNA coding, alternative splicing and RNA non-coding

Gene expression involves transcription, translation and the turnover of mRNAs and proteins (BUCCITELLI; SELBACH, 2020). The central dogma of molecular biology (CRICK, 1970) indicates that most eukaryotic genes (DNA fragments of the genome) are copied in the form of messenger RNA precursors (pre-mRNAs). Pre-mRNAs contain coding sequences (exons) separated by non-coding sequences (introns), which are removed by means of the process of splicing exons (Splicing) to give rise to exons connected to each other forming mature messenger RNAs (mRNAs). Finally, the mRNA is translated into protein through ribosomes). It is known that cells contain substantial amounts of intronic RNA that originates from unprocessed (primary) gene transcripts called non-coding RNAs. It seems that gene expression is mainly regulated by transcription factors (proteins that can bind specific DNA sequences to control transcription) (LAMBERT et al., 2018) at the transcriptional level and miRNAs (small regulatory non-coding RNAs) at the post-transcriptional level (MITSIS et al., 2020).

Splicing is regulated by a complex machinery called spliceosome that consists of five small nuclear ribonucleoproteins (snRNP) (U1, U2, U4 / U6, U5) (STALEY; GUTHRIE, 1998; WILL; LÜHRMANN, 2011) and more than 200 auxiliary proteins that recognize and assemble exons through consensus sequences that mark the exon-intron junctions, recognized by the spliceosome (MATLIN; MOORE, 2007; WILL; LÜHRMANN, 2011).

Splicing was discovered in 1970, and soon afterward it was observed that there are alternative splicing patterns within a single pre-mRNA molecule that can produce different functional mRNAs (BERGET; MOORE; SHARP, 1977), which means that different proteins can be obtained from one gene, through this process known as alternative Splicing (KALSOTRA; COOPER, 2011). However, our knowledge of RNA splicing is quite recent (ROY; GILBERT, 2006). In humans, up to 95% of genes are alternately spliced (JIANG; CHEN, 2021), creating diversity and allowing a gene to generate multiple protein isoforms with different functions. It can express itself in a specific way in a tissue, at a specific stage of development, therefore coding important domains (CASTLE et al., 2008; TAPIAL et al., 2017).

Alternative splicing genes regulate the localization of proteins, their enzymatic properties and their interaction with ligands. In most cases, changes caused by individual splicing isoforms are small. Due to its widespread usage and molecular

versatility, alternative splicing emerges as a central element in gene regulation that interferes with almost every biological function analyzed (KELEMEN et al., 2013), such as cell development and differentiation (JIANG; CHEN, 2021). In addition, it is known that about 15%-30% of mutations associated with hereditary diseases have been related to failures in the regulation of alternative splicing (PARK et al., 2018; ELLINGFORD et al., 2019).

Alternative splicing explains that cattle have fewer genes (approximately 23,000) compared to the number of proteins (AEBERSOLD et al., 2018; LAU et al., 2019). The different isoforms of a gene show dramatic differences in their biological properties and functions (NILSEN; GRAVELEY, 2010; KALSOTRA; COOPER, 2011). However, our knowledge about function of isoforms and identification of the most functionally important protein isoforms/variants in cattle is still limited (RODRIGUEZ et al., 2012; LI et al., 2017). There is still a profound lack of knowledge about the functions of biologically important isoforms, especially in the Nelore breed.

Regulation of alternative splicing

Constitutive splicing is the process in which mRNA is spliced identically, producing the same isoforms, while alternative splicing generates different isoforms by using different sets of exons (NAKKA et al., 2018). Figure 1 shows the five main subtypes of alternative splicing.

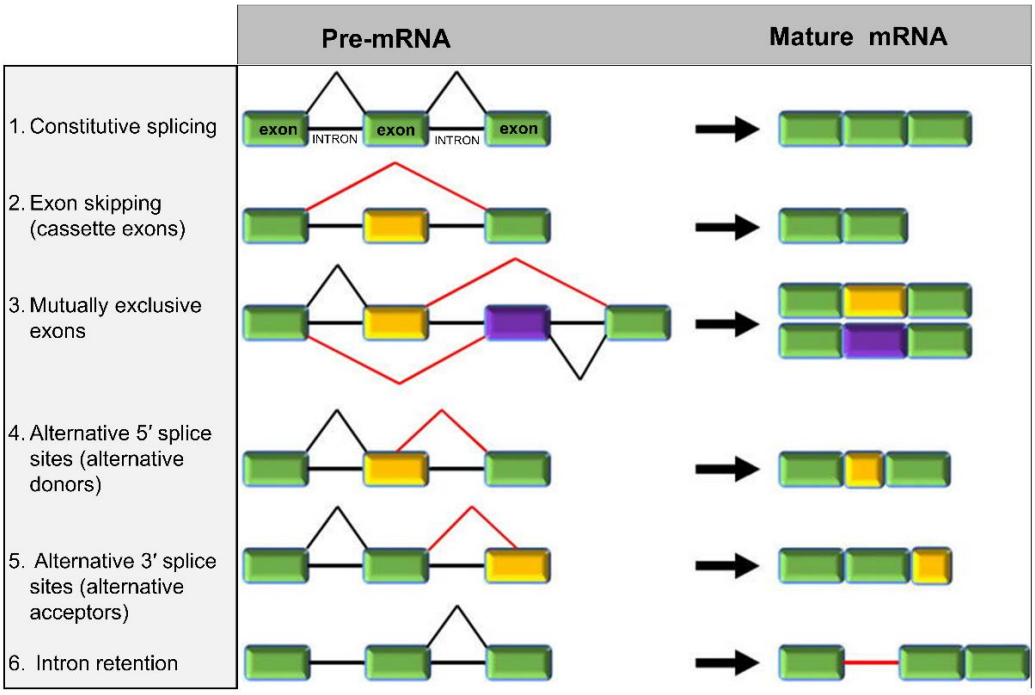


Figure 1. Constitutive splice and the five main types of alternative splicing (image adapted from JIANG; CHEN, 2021).

Following the representation in Figure 1: 1) Constitutive splicing; 2) Exon omission (cassette exons) is the most common in mammalian cells, resulting in the complete omission of one or more exons (SUGNET et al., 2004); 3) Mutually exclusive exons are uncommon, in which two or more splicing events are no longer independent. They run or deactivate in a coordinated manner (SAMMETH, 2009); 4) Alternative 5′ junction sites (alternative donors): the use of an alternative 5′ donor site, which changes the 3′ boundary of the upstream exon (JIANG; CHEN, 2021); 5) Alternative 3' junction sites (alternative acceptors): compared to alternative 5' junction sites, the use of an alternative 3' junction site causes the change of the 5' boundary of the exon downstream (JIANG; CHEN, 2021); 6) Intron retention, one or more introns remain unspliced in the mRNA (VANICHKINA et al., 2018). It is the rarest in mammals (SAMMETH; FOISSAC; GUIGÓ, 2008; NAKKA et al., 2018) and is often misinterpreted as a splicing error (JIANG; CHEN, 2021).

We can use different approaches to define exon junctions from RNA-seq data: isoform-based or exon-based (CHEN, 2013; JIANG; CHEN, 2021) and whether the genome is used as a reference or based on *de novo* assemblies (MARTIN; WANG, 2011). Isoform-based analysis approaches quantify full-length transcription isoforms and then estimate the relationship between the isoforms they included and the isoforms excluding an alternative exon of interest by calculating the splicing level. Exon-based approaches quantify individual exon inclusion ratios. It is more appropriate if the target is not a whole isoform.

The computer programs that can analyze alternative splices usually consist of three main steps: detection, statistical comparison and prediction of effects (HALPERIN et al., 2021). Currently, determination of isoform structures and quantification is still hampered by technological limitations (JIANG; CHEN, 2021) and the discovery of new transcripts using RNA-seq short reads is still a challenge (CONESA et al., 2016).

Non-coding RNA (ncRNAs)

Only 2% of genes encode proteins in the human genome (DJEBALI et al., 2012; LI et al., 2014; HON et al., 2017) the rest of the RNAs that do not translate into amino acids are functional ncRNAs acting as important mediators of gene regulation in biological processes (RÖNNAU et al., 2014). Some ncRNAs can encode small peptides (length less than 100 aa) regulating biological processes, although it is uncommon (HUBÉ; FRANCASTEL, 2018) (Figure 3.A). In addition, the complexity of an organism is highly associated with the abundance of ncRNA genes (RUBIN et al., 2000; MATTICK, 2001; KAPUSTA; FESCHOTTE, 2014).

.

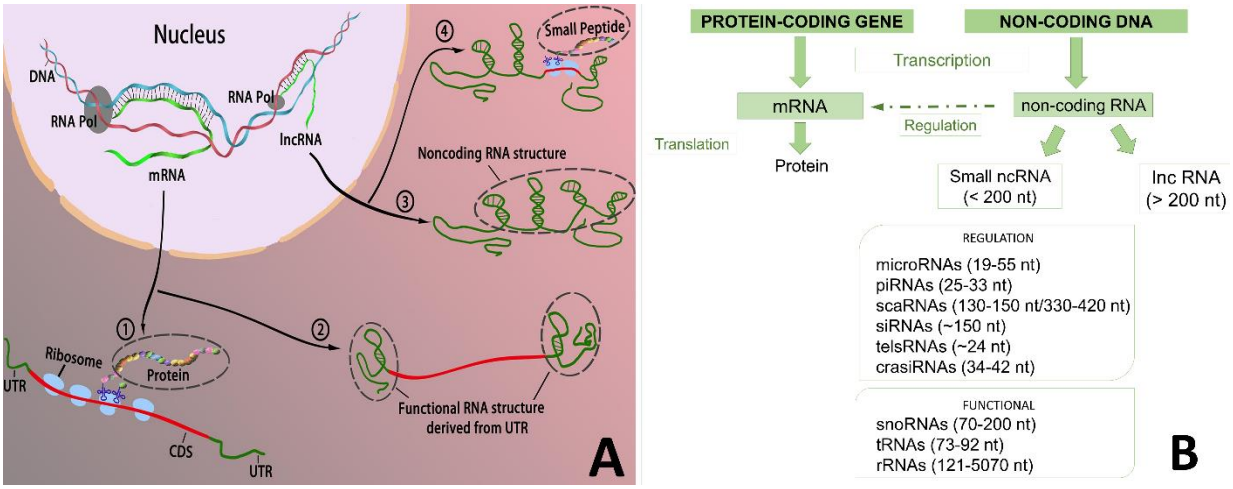


Figure 3. A. Structure of long coding and non-coding RNAs. ① Coding RNAs refer to the mRNA encoding the protein. ② Some coding mRNAs can function without translating into proteins by forming a secondary RNA structure derived primarily from untranslated regions (UTR). ③ Non-coding RNAs act as cell regulators without coding proteins. ④ Some lncRNAs (long non-coding RNAs) can bind to ribosomes and encode peptides to modulate cellular activities. **B.** Structure of coding and non-coding RNAs. Figure adapted from (RÖNNAU et al., 2014; LI; LIU, 2019; BHAT et al., 2020).

All these ncRNA classes (Figure 3.B) can perform numerous cellular functions, including structural, catalytic and regulatory functions. They are studied intensively because they direct various signaling pathways (RÖNNAU et al., 2014). For example, lncRNAs modulate enzymatic activities (WANG; CHANG, 2011), cell differentiation, growth (HOBUSS; BÄR; THUM, 2019), adipose function (SUN; LIN, 2019) and play an important role in immunology (CHEN; SATPATHY; CHANG, 2017). In addition, ncRNAs, including microRNAs and small interfering RNAs, can act as regulators in alternative splicing processes (LUCO; MISTELI, 2011).

In Figure 3B, the ncRNAs can be subdivided according to their size and function. A total of 2,578 microRNAs were identified in humans (BOLTON et al., 2014) and the number continues to rise. They regulate more than 50% of all

protein-coding genes in mammalian cells (SCHWARZENBACH et al., 2014). Genes can be attacked by multiple miRNAs and each miRNA is able to target hundreds of mRNAs directly or indirectly, which are involved in many physiological processes (CROCE, 2009). miRNA studies for RFI trait are still rare in Nelore cattle (DE OLIVEIRA et al., 2018; CARVALHO et al., 2019).

Another type of ncRNA is the snoRNA (small nucleolar RNAs). They are known to guide the modifications and processing of nucleotides (CECH; STEITZ, 2014; SLOAN et al., 2016). There are 400 species of snoRNA in the human genome (MARTENS-UZUNOVA; OLVEDY; JENSTER, 2013) and some miRNAs originate from snoRNAs (RÖNNAU et al., 2014). SnoRNAs are post-transcriptional (SUN; HAO; PRASANTH, 2018) or transcriptional regulators interfering with the maturation and modification of other ncRNAs (CECH; STEITZ, 2014; BAZIN et al., 2017).

Some studies have identified lncRNA in various tissues of cattle (HUANG; LONG; KHATIB, 2012; QU; ADELSON, 2012; WEIKARD; HADLICH; KUEHN, 2013; BILLEREY et al., 2014; GAO et al., 2019; WANG; KOGANTI; YAO, 2020). The detection changes from 304 to 23,735 lncRNA depending on the tissue (KOSINSKA-SELBI; MIELCZAREK; SZYDA, 2020) and also the approach. The number of functional lncRNAs has not yet been accurately determined, however they have been increasingly described as having important cellular functions (STATELLO et al., 2020). They are also mediators of interactions with proteins (and/or other molecules), although the nature and dynamics of such interactions have yet to be elucidated (KIRK et al., 2018).

Due to the fundamental role of ncRNAs in the regulation of gene expression, and consequently its possible impact on phenotypes of interest for animal production, it is an important topic of study. In cattle, the number of studies are slowly increasing, but the identification of ncRNAs is limited and even more so in the Nelore breed.

3.3. RNA-seq workflow.

The study of RNA-seq data generally involves trimming, alignment, counting and normalization of sequencing data and differential expression analysis (CORCHETE et al., 2020). Trimming is used to increase the rate of mapped reads by removing the adaptor sequences and removing poor-quality nucleotides with a wisely chosen reading length, to avoid unpredictable changes in gene expression (WILLIAMS et al., 2016) and transcriptome assembly (MACMANES, 2014). The beneficial effects of this process were evaluated for assembly based on genomic

references (DEL FABBRO et al., 2013; CHEN et al., 2014) and a *de novo* assembly approach (MBANDI et al., 2014). Recently, methods have been reported without alignment with reference genome, pre-index reference transcripts, divide sequence readings into k-mers, and then quick matches as "pseudo" alignments have been performed (JIANG; CHEN, 2021).

Once the reads have been mapped, they must be assigned to a gene or transcript (called a count or quantification). This is followed by a normalization procedure to eliminate potential sequencing bias. There are several methods and studies comparing methodologies with various algorithms used for identifying differently expressed genes with different results (MOULOS; HATZIS, 2015; SEYEDNASROLLAH; LAIHO; ELO, 2015; CORCHETE et al., 2020; GERACI; SAHA; BIANCHINI, 2020).

The biggest challenge in RNA-seq analysis is combining the different steps sequentially into a complete workflow and users must choose from many possible approaches, parameters and methodological options. There is no consensus on the most appropriate methodology that should be used for the analysis of RNA-seq (CORCHETE et al., 2020).

3.4. Transcriptome analyses in cattle through RNA-seq

The RNA-seq is a tool of no more than a decade, currently considered the most powerful, robust and adaptable technique for analyzing genes that are expressed in each tissue or cell type depending on the conditions in which they are found (STARK; GRZELAK; HADFIELD, 2019; CORCHETE et al., 2020; GERACI; SAHA; BIANCHINI, 2020).

RNA-seq is a more accurate and informative method compared to other techniques for transcriptome analysis, such as microarrays (XUAN et al., 2013; FINOTELLO; DI CAMILLO, 2015; HAN et al., 2015; SERNA et al., 2016). It makes it possible to study novel transcripts with a higher resolution, a better detection range and lower technical variability (PERKINS et al., 2014; SU et al., 2014; Y et al., 2020). It has been well described that RNA-seq also offers a high degree of agreement with validations using qRT-PCR (quantitative polymerase chain reaction; real time PCR) (SU et al., 2014).

Several studies have used RNA-seq to investigate gene expression in Nelore cattle which has made it possible to identify DEGs in muscle tissue associated with tenderness (FONSECA et al., 2017), fatty acid composition (OLIVIERI et al., 2021), loin area and subcutaneous fat (SILVA-VIGNATO et al., 2017), intramuscular fat (CESAR et al., 2015) and residual feed intake (RFI) (TIZIOTO et al., 2016) as well

as in the liver (TIZIOTO et al., 2015) and in multiple tissues (CHEN et al., 2021; DU et al., 2021).

Another application of RNA-seq is the detection of alternative splicing events. In cattle, these events have been associated with mastitis (WANG et al., 2016) embryonic development (HUANG; KHATIB, 2010) carcass traits (HE; LIU, 2013) and meat quality (SILVA et al., 2020; MALANE et al., 2020; FANG et al., 2021; SUN et al., 2021). This technology is also capable of detecting ncRNA associated with RFI in Nelore cattle (ALEXANDRE et al., 2020). For Nelore cattle, knowledge about the molecular mechanisms regulated by alternative splicing events and ncRNA, especially in feed efficiency traits, is still limited.

4. List of abbreviations

Food and Agriculture Organization of the United Nations (FAO), Sustainable Development Goal 2 (SDG2), Sustainable Development Goals (SDGs), United Nations (UN), greenhouse gas (GHG), Residual Food Intake (RFI), RNA sequencing (RNA-seq), Intergovernmental Panel on Climate Change (IPCC), Sustainable Development Goal 13 (SDG 13), World Health Organization (WHO), Kilograms (Kg), Ribonucleic acid (RNA), Nitrogen (N), Phosphorus (P), Potassium (K), Average metabolic weight (BMR),Rate of weight gain (AWG, kg/day), Expected progeny reference (DEP), Quantitative polymerase chain reaction real time (qRT-PCR), Differential expression gene (DEGs), Gene Set Enrichment Analysis (GSEA), Messenger RNA (mRNA), ARN noncoding (NCRNAs), Long ARN noncoding (lncRNA), MicroRNAs (miRNAs), Piwi-interacting RNAs (piRNAs), Ribosomal RNAs (rRNAs), Small Cajal-specific body RNAs (scaRNAs), Small-interfering RNAs (siRNAs), Small nuclear RNAs (snRNAs). Small nucleolar RNAs (snoRNAs), Transfer RNAs (tRNAs).

5. References

ABIEC. Perfil da pecuária no Brasil. **Beef REPORT**, p. 49, 2019. Disponível em:<http://www.abiec.com.br/controle/uploads/arquivos/sumario2019portugues.pdf>.

AEBERSOLD, R. et al. How many human proteoforms are there? **Nature Chemical Biology**, v. 14, n. 3, p. 206–214, fev. 2018.

AL-HUSSEINI, W. et al. Characterization and Profiling of Liver microRNAs by RNA-sequencing in Cattle Divergently Selected for Residual Feed Intake. **Asian-Australasian journal of animal sciences**, v. 29, n. 10, p. 1371–1382, out. 2016.

ALEXANDRE, P. A. et al. Exploring the Regulatory Potential of Long Non-Coding RNA in Feed Efficiency of Indicine Cattle. **Genes**, v. 11, n. 9, p. 1–22, set. 2020.

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, jun. 2003.

BHAT, A.A. et al. Role of non-coding RNA networks in leukemia progression,

metastasis and drug resistance. **Mol Cancer**, v. 19, n. 57. Mar. 2020.

BAZIN, J. et al. Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 114, n. 46, p. E10018–E10027, nov. 2017.

BERGET, S. M.; MOORE, C.; SHARP, P. A. Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. **Proceedings of the National Academy of Sciences**, v. 74, n. 8, p. 3171–3175, ago. 1977.

BERRY, D. P.; CROWLEY, J. J. CELL BIOLOGY SYMPOSIUM: Genetics of feed efficiency in dairy and beef cattle. **Journal of Animal Science**, v. 91, n. 4, p. 1594–1613, abr. 2013.

BILLEREY, C. et al. Identification of large intergenic non-coding RNAs in bovine muscle using next-generation transcriptomic sequencing. **BMC Genomics**, v. 15, n. 1, p. 499, jun. 2014.

BOLTON, E. M. et al. Noncoding RNAs in prostate cancer: the long and the short of it. **Clinical cancer research : an official journal of the American Association for Cancer Research**, v. 20, n. 1, p. 35–43, jan. 2014.

BRANCO, R. et al. Consumo alimentar residual de machos nelore selecionados para peso pós-desmame. **46° Reunião Anual da Sociedade Brasileira de Zootecnia**, p. 3, jul. 2009.

BUCCITELLI, C., SELBACH, M. mRNAs, proteins and the emerging principles of gene expression control. **Nature Reviews Genetics, v.** 21, n. 10, p. 630-644, oct. 2020.

CARBERRY, C. A. et al. Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. **Applied and Environmental Microbiology**, v. 78, n. 14, p. 4949–4958, jul. 2012.

CARVALHO, E. B. et al. Differentially expressed mRNAs, proteins and miRNAs associated to energy metabolism in skeletal muscle of beef cattle identified for low and high residual feed intake. **BMC genomics**, v. 20, n. 1, jun. 2019.

CASTLE, J. C. et al. Expression of 24,426 human alternative splicing events and predicted cis regulation in 48 tissues and cell lines. **Nature Genetics**, v. 40, n. 12, p. 1416–1425, nov. 2008.

CEACERO, T. M. et al. Phenotypic and Genetic Correlations of Feed Efficiency Traits with Growth and Carcass Traits in Nelore Cattle Selected for Postweaning Weight. **PLOS ONE**, v. 11, n. 8, p. 66, ago. 2016.

CECH, T. R.; STEITZ, J. A. The noncoding RNA revolution-trashing old rules to forge new ones. **Cell**, v. 157, n. 1, p. 77–94, mar. 2014.

CESAR, A. S. M. et al. Putative regulatory factors associated with intramuscular fat content. **PloS one**, v. 10, n. 6, jun. 2015.

CHEN, C. et al. Software for pre-processing Illumina next-generation sequencing short read sequences. **Source Code for Biology and Medicine**, v. 9, n. 1, p. 8–8, maio 2014.

CHEN, L. Statistical and Computational Methods for High-Throughput Sequencing Data Analysis of Alternative Splicing. **Statistics in biosciences**, v. 5, n. 1, p. 138–155, maio 2013.

CHEN, W. et al. Identification of Predictor Genes for Feed Efficiency in Beef Cattle

by Applying Machine Learning Methods to Multi-Tissue Transcriptome Data. **Frontiers in genetics**, v. 12, fev. 2021.

CHEN, Y. G.; SATPATHY, A. T.; CHANG, H. Y. Gene regulation in the immune system by long noncoding RNAs. **Nature immunology**, v. 18, n. 9, p. 962–972, ago. 2017.

CONESA, A. et al. A survey of best practices for RNA-seq data analysis**.Genome Biology,** v 17, n. 13, jan. 2016.

CORCHETE, L. A. et al. Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. **Scientific Reports** , v. 10, n. 1, p. 1–15, nov. 2020.

CRICK, F. Central Dogma of Molecular Biology. **Nature** , v. 227, n. 5258, p. 561–563, nov.1970.

CROCE, C. M. Causes and consequences of microRNA dysregulation in cancer. **Nature reviews. Genetics**, v. 10, n. 10, p. 704–714, out. 2009.

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. 1, dez. 2018.

DEL CLARO, A. C.; MERCADANTE, M. E. Z.; VASCONCELOS SILVA, J. A. Meta-análise de parâmetros genéticos relacionados ao consumo alimentar residual e a suas características componentes em bovinos. **Pesquisa Agropecuaria Brasileira**, v. 47, n. 2, p. 302–310, fev. 2012.

DEL FABBRO, C. et al. An Extensive Evaluation of Read Trimming Effects on Illumina NGS Data Analysis. **PLoS ONE**, v. 8, n. 12, dez. 2013.

DJEBALI, S. et al. Landscape of transcription in human cells. **Nature**, v. 489, n. 7414, p. 101–108, set. 2012.

DOBIN, A. et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics (Oxford, England)**, v. 29, n. 1, p. 15–21, jan. 2013.

DU, L. et al. Transcriptome profiling analysis of muscle tissue reveals potential candidate genes affecting water holding capacity in Chinese Simmental beef cattle. **Scientific Reports ,** v. 11, n. 1, p. 1–14, jun. 2021.

EDENHOFER, O. et al. Climate Change 2014 Mitigation of Climate Change Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Edited by. 2014. Disponível em: <www.ipcc.ch>.

ELLINGFORD, J. M. et al. Functional and in-silico interrogation of rare genomic variants impacting RNA splicing for the diagnosis of genomic disorders. **bioRxiv**, sep. 2019.

FANG, S. et al. NONCODEV5: a comprehensive annotation database for long non-coding RNAs. **Nucleic acids research**, v. 46, n. D1, p. D308–D314, jan. 2018.

FANG, X. et al. Comparative Genome-Wide Alternative Splicing Analysis of Longissimus Dorsi Muscles Between Japanese Black (Wagyu) and Chinese Red Steppes Cattle. **Frontiers in veterinary science**, v. 8, abr. 2021.

FAO. **How to Feed the World in 2050**. Disponível em: <http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf>.

FERRELL, C. L.; JENKINS, T. G. Cow Type and the Nutritional Environment:

Nutritional Aspects. **Journal of Animal Science**, v. 61, n. 3, p. 725–741, set. 1985.

FIDELIS, H. A. et al. Residual feed intake, carcass traits and meat quality in Nelore cattle. **Meat science**, v. 128, p. 34–39, jun. 2017.

FINOTELLO, F.; DI CAMILLO, B. Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. **Briefings in Functional Genomics**, v. 14, n. 2, p. 130–142, mar. 2015.

FONSECA, L. F. S. et al. Differences in global gene expression in muscle tissue of Nelore cattle with divergent meat tenderness. **BMC Genomics**, v. 18, n. 1, p. 945, dez. 2017.

FONTOURA, A. B. P. et al. Associations between feed efficiency, sexual maturity and fertility-related measures in young beef bulls. **animal**, v. 10, n. 1, p. 96–105, jul. 2016.

GAO, D. et al. A survey of statistical software for analysing RNA-seq data. **Human Genomics**, v. 5, n. 1, p. 56–60, out. 2010.

GAO, Y. et al. Analysis of Long Non-Coding RNA and mRNA Expression Profiling in Immature and Mature Bovine (Bos taurus) Testes. **Frontiers in Genetics**, v. 10, n. JUL, p. 646, jul. 2019.

GERACI, F.; SAHA, I.; BIANCHINI, M. Editorial: RNA-Seq Analysis: Methods, Applications and Challenges. **Frontiers in Genetics**, v. 11, p. 220, mar. 2020.

GÓMEZ-GALLEGO, C. et al. Human Breast Milk NMR Metabolomic Profile across Specific Geographical Locations and Its Association with the Milk Microbiota. **Nutrients**, v. 10, n. 10, p. 1355, 21 set. 2018.

GRION, A. L. et al. Selection for feed efficiency traits and correlated genetic responses in feed intake and weight gain of Nelore cattle. **Journal of Animal Science**, v. 92, n. 3, p. 955–965, jul. 2014.

HALPERIN, R. F. et al. Improved methods for RNAseq-based alternative splicing analysis. **Scientific Reports**, v. 11, n. 1, p. 1–15, maio 2021.

HAN, Y. et al. Advanced Applications of RNA Sequencing and Challenges. **Bioinformatics and Biology Insights**, v. 9, n. Suppl 1, p. 29–46, nov. 2015.

HAYES, B. J.; LEWIN, H. A.; GODDARD, M. E. The future of livestock breeding: Genomic selection for efficiency, reduced emissions intensity, and adaptation**.Trends in Genetics**,v. 29, n 4, p 206-214, abr. 2013.

HE, H.; LIU, X. Characterization of Transcriptional Complexity during Longissimus Muscle Development in Bovines Using High-Throughput Sequencing. **PLoS ONE**, v. 8, n. 6, jun. 2013.

HEGARTY, R. S. et al. Cattle selected for lower residual feed intake have reduced daily methane production. **Journal of Animal Science**, v. 85, n. 6, p. 1479–1486, jun. 2007.

HOBUSS, L.; BÄR, C.; THUM, T. Long non-coding RNAs: At the heart of cardiac dysfunction? **Frontiers in Physiology**, v. 29, n. 10, p. 30, jan, 2019.

HOLDEN, N. M. et al. Review of the sustainability of food systems and transition using the Internet of Food. **npj Science of Food 2018 2:1**, v. 2, n. 1, p. 1–7, 9 out. 2018.

HON, C. C. et al. An atlas of human long non-coding RNAs with accurate 5’ ends. **Nature**, v. 543, n. 7644, p. 199–204, 9 mar. 2017.

HRDLICKOVA, R.; TOLOUE, M.; TIAN, B. RNA-Seq methods for transcriptome analysis. **WIREs RNA.**v. 8,n. 1, jan. 2017.

HUANG, W.; KHATIB, H. Comparison of transcriptomic landscapes of bovine embryos using RNA-Seq. **BMC Genomics**, v. 11, n. 1, p. 1–10, dez. 2010.

HUANG, W.; LONG, N.; KHATIB, H. Genome-wide identification and initial characterization of bovine long non-coding RNAs from EST data. **Animal Genetics**, v. 43, n. 6, p. 674–682, dez. 2012.

HUBÉ, F.; FRANCASTEL, C. Coding and non-coding RNAs, the frontier has never been so blurred. **Frontiers in Genetics**, v. 9, n. APR, abr. 2018.

JIANG, W.; CHEN, L. Alternative splicing: Human disease and quantitative analysis from high-throughput sequencing. **Computational and Structural Biotechnology Journal**, v. 19, p. 183, jan. 2021.

KALSOTRA, A.; COOPER, T. A. Functional consequences of developmentally regulated alternative splicing. **Nature Reviews. Genetics**, v. 12, n. 10, p. 715, out. 2011.

KAPUSTA, A.; FESCHOTTE, C. Volatile evolution of long noncoding RNA repertoires: mechanisms and biological implications. **Trends in Genetics : TIG**, v. 30, n. 10, p. 439–452, set. 2014.

KELEMEN, O. Function of alternative splicing. Gene, v. 514, n.1, p. 1-30, fev. 2013.

KIM, D. et al. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. **Genome Biology**, v. 14, n. 4, abr. 2013.

KIRK, J. M. et al. Functional classification of long non-coding RNAs by k-mer content. **Nature genetics**, v. 50, n. 10, p. 1474–1482, out. 2018.

KOCH, R. M. et al. Efficiency of Feed Use in Beef Cattle. **Journal of Animal Science**, v. 22, n. 2, p. 486–494, maio 1963.

KOSINSKA-SELBI, B.; MIELCZAREK, M.; SZYDA, J. Review: Long non-coding RNA in livestock. **Animal**, v. 14, n. 10, p. 2003–2013, jan. 2020.

LAMBERT S.A. et al. The Human Transcription Factors. **Cell Reports**. v. 4, n. 75, p. 598-599, ago. 2018.

LANGFELDER, P.; MISCHEL, P. S.; HORVATH, S. When Is Hub Gene Selection Better than Standard Meta-Analysis? **PLoS ONE**, v. 8, n. 4, abr. 2013.

LANNA, D. P.; DE ALMEIDA, R. Residual feed intake: um novo critério de seleção?. **V Simpósio da Sociedade Brasileira de Melhoramento Animal** .jul, 2003.

LAU, E. et al. Splice-Junction-Based Mapping of Alternative Isoforms in the Human Proteome. **Cell Reports**, v. 29, n. 11, p. 3751- 3765, dez. 2019.

LI, G. et al. Long noncoding RNA plays a key role in metastasis and prognosis of hepatocellular carcinoma. **BioMed Research International**, v. 2014, jun. 2014.

LI, H. et al. Annotation of Alternatively Spliced Proteins and Transcripts with Protein-Folding Algorithms and Isoform-Level Functional Networks. **Methods in Molecular Biology**, v. 1558, p. 415–436, jul. 2017a.

LI, J.; LIU, C. Coding or noncoding, the converging concepts of RNAs. **Frontiers in Genetics**, v. 10, n. MAY, p. 496, apr. 2019.

LUCO, R. F.; MISTELI, T. More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. **Current opinion in genetics & development**, v. 21, n. 4, p. 366, ago. 2011.

MACMANES, M. D. On the optimal trimming of high-throughput mRNA sequence data. **Frontiers in genetics**, v. 5, n. JAN, jan. 2014.

MALANE M.M., et al. Structural variants affecting mRNAs isoforms splice sites associated with marbling in Nellore cattle. **Journal of Animal Science**, v. 98, n. 4, p. 24–25, Nov. 2020.

MARTENS-UZUNOVA, E. S.; OLVEDY, M.; JENSTER, G. Beyond microRNA - Novel RNAs derived from small non-coding RNA and their implication in cancer. **Cancer Letters**, v. 340, n. 2, p. 201–211, nov. 2013.

MARTIN, J. A.; WANG, Z. Next-generation transcriptome assembly. **Nature Reviews Genetics 2011 12:10**, v. 12, n. 10, p. 671–682, set. 2011.

MARTIN, P. Climate change, complexity, agriculture and challenged governance. In: **Research Handbook on Climate Change and Agricultural Law**.p. 74–102. fev. 2013.

MARTIN, P. et al. Invited review: Disentangling residual feed intake—Insights and approaches to make it more fit for purpose in the modern context. **Journal of Dairy Science**, v. 104, n. 6, p. 6329–6342, jun. 2021.

MATLIN, A. J.; CLARK, F.; SMITH, C. W. J. Understanding alternative splicing: towards a cellular code. **Nature Reviews Molecular Cell Biology 2005 6:5**, v. 6, n. 5, p. 386–398, maio 2005.

MATLIN, A. J.; MOORE, M. J. Spliceosome assembly and composition. **Advances in experimental medicine and biology**, v. 623, p. 14–35, jun. 2007.

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

MBANDI, S. K. et al. A glance at quality score: Implication for de novo transcriptome reconstruction of Illumina reads. **Frontiers in Genetics**, v. 5, n. FEB, p. 17, feb. 2014.

MEHRBAN, H. et al. Genetic parameters and correlations of related feed efficiency, growth, and carcass traits in Hanwoo beef cattle. **Animal Bioscience**, v. 34, n. 5, p. 824, maio. 2021.

MENEZES, A. C. B. et al. Feeding behavior, water intake, and energy and protein requirements of young Nelore bulls with different residual feed intakes. **Journal of animal science**, v. 98, n. 9, set. 2020.

MITSIS et al. Transcription factors and evolution: An integral part of gene expression (Review). **World Academy of Sciences Journal**, v. 2, p. 3-8, jan. 2020.

MOULOS, P.; HATZIS, P. Systematic integration of RNA-Seq statistical algorithms for accurate detection of differential gene expression patterns. **Nucleic acids research**, v. 43, n. 4, fev. 2015.

MUKIIBI, R. et al. Liver transcriptome profiling of beef steers with divergent growth rate, feed intake, or metabolic body weight phenotypes1. **Journal of animal science**, v. 97, n. 11, p. 4386–4404, nov. 2019.

MUKIIBI, R. et al. Bovine hepatic miRNAome profiling and differential miRNA expression analyses between beef steers with divergent feed efficiency

phenotypes. **Scientific reports**, v. 10, n. 1, dez. 2020.

NAKKA, K. et al. Diversification of the muscle proteome through alternative splicing. **Skeletal Muscle**, v. 8, n. 1, p. 8, mar. 2018.

NAKKA, K. et al. Diversification of the muscle proteome through alternative splicing. **Skeletal Muscle**, v. 8, n. 1, p. 8, mar. 2018.

NASCIMENTO, C. F. et al. Residual feed intake and blood variables in young Nelore cattle. **Journal of animal science**, v. 93, n. 3, p. 1318–1326, mar. 2015

NILSEN, T. W.; GRAVELEY, B. R. Expansion of the eukaryotic proteome by alternative splicing. **Nature**, v. 463, n. 7280, p. 457–463, jan. 2010.

NKRUMAH, J. D. et al. Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. **J Anim Sci**, v. 84, n. 1, p. 145–153, jan. 2006.

OLIVIERI, B. F. et al. Differentially expressed genes identified through RNA-seq with extreme values of principal components for beef fatty acid in Nelore cattle. **Journal of animal breeding and genetics**, v. 138, n. 1, p. 80–90, jan. 2021.

PARK, E. et al. The Expanding Landscape of Alternative Splicing Variation in Human Populations. **The American Journal of Human Genetics**, v. 102, n. 1, p. 11–26, jan. 2018.

PERKINS, J. R. et al. A comparison of RNA-seq and exon arrays for whole genome transcription profiling of the L5 spinal nerve transection model of neuropathic pain in the rat. **Molecular Pain**, v. 10, n. 1, p. 7, jan. 2014.

POORE, J.; NEMECEK, T. Reducing food's environmental impacts through producers and consumers. **Science**, v. 360, n. 6392, p. 987–992, jul. 2018a.

PORTER, J. R. et al. Food security and food production systems. **Climate Change 2014 Impacts, Adaptation and Vulnerability: Part A: Global and Sectoral Aspects**, p. 485–534, jul. 2014.

QU, Z.; ADELSON, D. L. Bovine ncRNAs are abundant, primarily intergenic, conserved and associated with regulatory genes. **PLoS ONE**, v. 7, n. 8, ago. 2012.

RICHARDSON, E. C.; HERD, R. M. Biological basis for variation in residual feed intake in beef cattle. 2. Synthesis of results following divergent selection. **Australian Journal of Experimental agriculture.**, v. 44, n. 4–5, p. 431–440, jan. 2004.

RODRIGUEZ, J. M. et al. APPRIS: annotation of principal and alternative splice isoforms. **Nucleic Acids Research**, v. 41, n. Database issue, p. D110-7, nov. 2012.

ROJAS-DOWNING, M. M. et al. Climate change and livestock: Impacts, adaptation, and mitigation. **Climate Risk Management**, v. 16, p. 145–163, apr. 2017.

RÖNNAU, C. G. H. et al. Noncoding RNAs as Novel Biomarkers in Prostate Cancer. **BioMed Research International**, v. 2014, jun. 2014.

ROSEGRANT, M. W. et al. Looking into the future for agriculture and AKST. **Agriculture at a Crossroads, global report. Washington, DC, USA: Island Press**. pp.307-376. jun. 2009.

ROY, B.; M. HAUPT, L.; R. GRIFFITHS, L. Review: Alternative Splicing (AS) of Genes As An Approach for Generating Protein Complexity. **Current Genomics**, v. 14, n. 3, p. 182–194, maio 2013.

ROY, S. W.; GILBERT, W. The evolution of spliceosomal introns: patterns, puzzles and progress. **Nature reviews. Genetics**, v. 7, n. 3, p. 211–221, mar. 2006.

RUBIN, G. M. et al. Comparative genomics of the eukaryotes. **Science (New York, N.Y.)**, v. 287, n. 5461, p. 2204–2215, mar. 2000.

SACKS, D. et al. Multisociety Consensus Quality Improvement Revised Consensus Statement for Endovascular Therapy of Acute Ischemic Stroke. **International Journal of Stroke**, v. 13, n. 6, p. 612–632, ago. 2018.

SAMMETH, M. Complete alternative splicing events are bubbles in splicing graphs. **Journal of Computational Biology**, v. 16, n. 8, p. 1117–1140, ago. 2009.

SAMMETH, M.; FOISSAC, S.; GUIGÓ, R. A General Definition and Nomenclature for Al-ternative Splicing Events. **PLoS Computational Biology**, v. 4, n. 8, nov. 2008.

SANTANA, M. H. A. et al. Genome-wide association study for feedlot average daily gain in Nelore cattle (*Bos indicus*). **Journal of Animal Breeding and Genetics**, v. 131, n. 3, p. 210–216, jun. 2014.

SCHWARZENBACH, H. et al. Clinical relevance of circulating cell-free microRNAs in cancer. **Nature Reviews Clinical Oncology**, v. 11, n. 3, p. 145–156, mar. 2014.

SILVA, D.B.S., et al. Spliced genes in muscle from Nelore Cattle and their association with carcass and meat quality. **Scientific Reports,** v. 10, n. 14701, sep. 2020.

SERNA, M. et al. Ovarian Transcriptomic Analysis Reveals Differential Expression Genes Associated with Cell Death Process after Selection for Ovulation Rate in Rabbits. **Animals**, v. 10, n. 10, p.1924. oct 2020.

SEYEDNASROLLAH, F.; LAIHO, A.; ELO, L. L. Comparison of software packages for detecting differential expression in RNA-seq studies. **Briefings in Bioinformatics**, v. 16, n. 1, p. 59–70, jan. 2015.

SILVA-VIGNATO, B. et al. Comparative muscle transcriptome associated with carcass traits of Nelore cattle. **BMC Genomics**, v. 18, n. 1, p. 506, dez. 2017.

SILVA DE OLIVEIRA, J. et al. Fisiologia, manejo e alimentação de bezerros de corte. **arq. Ciênc. Vet. Zool. Unipar**. v. 10, n. 1, 2007.

SLOAN, K. E. et al. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. **RNA Biology**, v. 14, n. 9, p. 1138–1152, dez. 2016.

STALEY, J. P.; GUTHRIE, C. Mechanical Devices of the Spliceosome: Motors, Clocks, Springs, and Things. **Cell**, v. 92, n. 3, p. 315–326, fev. 1998.

STARK, R.; GRZELAK, M.; HADFIELD, J. RNA sequencing: the teenage years. **Nature Reviews Genetics 2019 20:11**, v. 20, n. 11, p. 631–656, jul. 2019.

STATELLO, L. et al. Gene regulation by long non-coding RNAs and its biological functions. **Nature Reviews Molecular Cell Biology 2020 22:2**, v. 22, n. 2, p. 96–118, dez. 2020.

SU, Z. et al. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. **Nature Biotechnology 2014 32:9**, v. 32, n. 9, p. 903–914, ago. 2014.

SUGNET, C. W. et al. Transcriptome and genome conservation of alternative splicing events in humans and mice. **Pacific Symposium on Biocomputing.**

Pacific Symposium on Biocomputing, p. 66–77, jul. 2004.

SUN, H. Z. et al. Gene co-expression and alternative splicing analysis of key metabolic tissues to unravel the regulatory signatures of fatty acid composition in cattle. **RNA biology**, v. 18, n. 6, p. 854–862, nov. 2021.

SUN, L.; LIN, J. D. Function and Mechanism of Long Noncoding RNAs in Adipocyte Biology. **Diabetes**, v. 68, n. 5, p. 887–896, feb. 2019.

SUN, Q.; HAO, Q.; PRASANTH, K. V. Nuclear Long Noncoding RNAs: Key Regulators of Gene Expression. **Trends in Genetics : TIG**, v. 34, n. 2, p. 142–157, fev. 2018.

TAPIAL, J. et al. An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. **Genome research**, v. 27, n. 10, p. 1759–1768, out. 2017.

TIZIOTO, P. C. et al. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC Genomics**, v. 16, n. 1, p. 1–14, jul. 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 9. Dec. 2016.

TRAPNELL, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. **Nature Biotechnology**, v. 28, n. 5, p. 511–515, maio. 2010a.

VANICHKINA, D. P. et al. Title: Challenges in defining the role of intron retention in normal biology and disease. **Semin Cell Dev Biol**. n. 75, p. 0-49. jul. 2017.

WANG, J.; KOGANTI, P. P.; YAO, J. Systematic identification of long intergenic non-coding RNAs expressed in bovine oocytes. **Reproductive Biology and Endocrinology**, v. 18, n. 1, p. 1–9, fev. 2020.

WANG, K. C.; CHANG, H. Y. Molecular mechanisms of long noncoding RNAs. **Molecular cell**, v. 43, n. 6, p. 904–914, set. 2011.

WANG, X. G. et al. Deciphering Transcriptome and Complex Alternative Splicing Transcripts in Mammary Gland Tissues from Cows Naturally Infected with Staphylococcus aureus Mastitis. **PLoS ONE**, v. 11, n. 7, p. e0159719, jan. 2016.

WEIKARD, R.; HADLICH, F.; KUEHN, C. Identification of novel transcripts and noncoding RNAs in bovine skin by deep next generation sequencing. **BMC Genomics**, v. 14, n. 1, nov. 2013.

WILL, C. L.; LÜHRMANN, R. Spliceosome structure and function. **Cold Spring Harbor perspectives in biology**, v. 3, n. 7, p. 1–2, jul. 2011.

WILLIAMS, C. R. et al. Trimming of sequence reads alters RNA-Seq gene expression estimates. **BMC bioinformatics**, v. 17, n. 1, fev. 2016.

UN. World 2019 Revision of World Population Prospects. **United Nations**. Disponível em: <https://population.un.org/wpp/>.

UN. Goal 2: Zero Hunger - United Nations Sustainable Development. Disponível em: <https://www.un.org/sustainabledevelopment/hunger/>.

WRIGHT, I. A. et al. Integrating crops and livestock in subtropical agricultural systems**. Journal of the Science of Food and Agriculture**, mar. 2012. .

XU, X. et al. Global greenhouse gas emissions from animal-based foods are twice those of plant-based foods. **Nature Food 2021 2:9**, v. 2, n. 9, p. 724–732, set. 2021.

XUAN, J. et al. Next-generation sequencing in the clinic: promises and challenges. **Cancer letters**, v. 340, n. 2, p. 284–295, nov. 2013.

Y, W. et al. Analysis of the physical meat quality in partridge (Alectoris chukar) and its relationship with intramuscular fat. **Poultry science**, v. 99, n. 2, p. 1225–1231, fev. 2020.

YANG, C. et al. The impact of RNA-seq aligners on gene expression estimation. **Conference on Bioinformatics, Computational Biology and Biomedicine,** v. 2015, p. 462–471, set. 2015.

CHAPTER 2 - Transcriptome profiling of liver in Nelore Cattle phenotypically divergent for RFI in two genetic groups

ABSTRACT – The identification and selection of genetically superior animals for Residual Feed Intake (RFI) resulting in improving food efficiency could enhance productivity and simultaneously minimize environmental impact. The aim of this study was to use RNA-seq data to identify differentially expressed genes (DEGs), specific biomarkers and enriched biological processes associated with RFI of liver in Nelore cattle in two genetic groups. The animals used in this study belong to two genetic groups. In the first group (G1), the animals (n=24) were from a stabilized Nelore herd submitted to RFI evaluation: 12 low RFI (LRFI) and 12 high RFI (HRFI). The samples from their liver tissues were frozen at -80 °C for subsequent extraction of total RNA. The libraries (TruSeq RNA Library) were constructed according to the manufacturer's protocol and sequenced using HiSeq 2500 System (Illumina). For the second group (G2), 20 samples of liver tissue of Nelore cattle divergent for RFI were selected. The RNA-seq was performed using a HiSeq 2000 System (Illumina®). The raw data of G2 were chosen from the ENA repository (EMBL-EBI). The same pipeline for RNA-seq for both genetic groups was performed: the low quality was trimmed using Trimmomatic software; HitSat2 software was used to map paired end reads; the scrip own was used to assemble the aligned reads for each sample and to estimate the number of counts per million (CPM). EdgeR software normalized reads and analysis differences were used. GSEA software and Pathway Studio (Elsevier) were used for differential and enrichment analysis, respectively. 1,811 DEGs were found for G1 and 2,054 for G2 (p-value 0.05). Common biological pathways, such as energy metabolism, protein turnover, redox homeostasis and immune response, were observed. We detected 88 common genes in both genetic groups, of which 33 were involved in immune response and blocking oxidative stress. In addition, seven (*B2M, ADSS, SNX2, TUBA4A, ARHGAP18, MECR, ABCF3*) possible gene biomarkers were found (in the 88 common ones), considering AUC > 0.70. The *B2M* gene was overexpressed in LRFI. This gene regulates the lipid metabolism protein turnover and inhibits cell death. We also found non-coding RNAs in both groups. MIR25 was up-regulated and SNORD16 was down-regulated in RLFI for G1. For G2, up-regulated RNase_MRP and SCARNA10 were found. We highlight MIR25 being able to act by blocking cytotoxicity and oxidative stress and RMRP as a blocker of mitochondrial damage.

Keywords: Bovine, biomarkers, feed efficiency, liver tissue, RNA-Seq.

1. Background

The world population is concerned about the environment and fulfilling Sustainable Development Goals (SDGs), such as SDG 13 (Climate Action) (UN, 2021); moreover, growing pressure from the media is putting the focus, in recent times, on livestock. Thirty-seven percent of greenhouse gas (GHG) emissions are partly bought about by global food supply even though the world population, particularly the high-income countries, has been recommended to reduce global meat consumption (MARTIN, 2017; POORE; NEMECEK, 2018), despite the recommendation of the World Health Organization (WHO) to consume 500 grams per week (SACKS et al., 2018). Brazil has the advantage of mostly having an extensive production system, which is more sustainable. It produces and exports most of the meat consumed by European countries, but it could benefit from increasing its market, as long as it seeks sustainable production (Meat|OECD-FAO, 2020). Furthermore, Brazil is the second highest meat consumer per capita (42.12 kg / year) (ABIEC, 2019), 80% of which is beef consumption. In the context of climate change and a more restrictive environmental legislation, bovine meat production is under close observation. The European market is particularly demanding wit1h regard to increasing sustainability.

Food intake on farms is approximately 60% of the total cost of a livestock farm, therefore selection of more efficient animals would reduce production costs and increase the profitability of the agricultural system (SILVA DE OLIVEIRA et al., 2007; DEL CLARO; MERCADANTE; VASCONCELOS SILVA, 2012). The identification and selection of genetically superior animals for Residual Feed Intake (RFI) (more efficient animals ingest less feed than estimated for the same weight gain compared with less efficient animals) would reduce feed costs, thereby increasing profits and minimizing environmental impact (YANG et al., 2022; BASARAB et al., 2003; FORESTRY, 2006).

There are various studies on the RFI trait in bovines from different breeds, genders and management systems (FONSECA et al 2015; TIZIOTO et al., 2015; CARDOSO et al., 2018; CHEN et al., 2021; DU et al., 2021; MARTIN et al., 2021), but little experimental information has been published that is thorough enough to unravel the biological regulation trait. The expression of the potential for feed efficiency is complex and depends on the interaction of numerous biochemical pathways through a multitude of tissues. The liver constitutes the main location of energy and substance metabolism (YANG et al., 2022). It can maintain glucose homeostasis in the blood and participates in protein biosynthesis as well as

immune and detoxification functions (TREFTS; GANNON; WASSERMAN, 2017). According to FERRELL and JENKINS (1985) in beef cattle, 70% of the total energy is used for maintenance, which illustrates the need for understanding biological mechanisms to optimize the regulation of maintenance.

The non-coding RNAs have been shown to regulate gene expression (AGOSTINO et al., 2022). Non-coding RNAs (ncRNAs) are generated from the larger part of the genome that does not encode proteins but produces transcripts that regulate gene expression and protein function (WINKLE et al., 2021). ncRNAs also have the potential to be superior to established protein-based biomarkers (RÖNNAU et al., 2014), however the use of these is still limited by the wide range of concentrations and difficult detection when there is low abundance (VELONAS et al., 2013).

In bovines, ncRNAs were reported as regulator of adipose tissue deposition (LEI et al., 2021) and fertility traits (GAO et al., 2019; WANG; KOGANTI; YAO, 2020). In Nelore cattle, ncRNAs were associated with intramuscular fat trait (DE OLIVEIRA et al., 2018). For RFI trait, there is studies to indicine Cattle (ALEXANDRE et al., 2020) Angus cattle (AL-HUSSEINI et al., 2016) and another *Bos taurus* breeding (NOLTE et al., 2020). Furthermore, studies to detect noncoding for Nelore cattle are still limited.

Studying all transcriptomes in liver tissue associated with the RFI trait allows us to observe the influence and complexity of gene network interactions. Furthermore, it allows us to understand the molecular mechanisms involved in more efficient animals. The aim of this study was to use RNA-seq data to identify DEGs, specific biomarkers, non-coding RNAs detection and enriched biological processes associated with RFI of liver in Nelore cattle in two genetic groups.

2. Methods

2.2. Sampling method

In this study, RNA-Seq data from two genetics groups (G1 and G2) of Nelore herds were evaluated for RFI. The animals constitutes two genetic lineages from two Brazilian researches institutes and they underwent a selection process as described below:

2.2.1. Genetic group 1 (G1)

The animals belonged to a stabilized Nelore herd, with 160 matrices and 8 bulls, from the Institute of Animal Science (Brazil). Since 1978, animals have been

selected within the herd for yearling weight (about 8% of males and 60% of females) based on individual yearling weight performance. Animals of the entire contemporary group (n = 60) had been submitted to weight gain test for a period of up 4 months (May 4 and October 19, 2010).

The animals were weighed without prior fasting three times per week on consecutive days estimated the RFI model proposed by KOCH et al., (1963). After the weight gain test, 24 animals were selected based on RFI values in order to compose a representative sample of more (low RFI, LRFI) and less efficient (high RFI, HRFI) animals. These animals were slaughtered and samples from liver tissue were collected immediately and stored in an RNA holder (BioAgency, São Paulo, Brazil) at -80 C until RNA extraction. Liver tissue samples (50-80 mg) were collected immediately after slaughter and stored in a 15 mL Falcon tube containing 5 mL of RNA holder (BioAgency, São Paulo, SP, Brazil). The samples were frozen at -80 °C for subsequent total RNA extraction. For more details, see FONSECA et al., (2015).

Total RNA was extracted using the RNeasy Lipid Tissue Mini Kit (Qiagen, CA, USA) according to the manufacturer's protocol. The purity of the extracted RNA was determined by evaluating absorbance using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Santa Clara, CA, USA, 2007). The quality of the RNA samples was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA, 2009) and the RNA concentration and genomic DNA contamination were measured using a Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA, 2010).

The mRNA libraries for sequencing were prepared using the TruSeq RNA Sample Preparation Kit® (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. Libraries were pooled to enable multiplexed sequencing and, on average, were generated 25 million reads per sample. RNA sequencing (RNA- Seq) was performed using a HiSeq 2500 System (Illumina®) that generated 100 bp paired-end reads.

2.2.2. Genetic group 2 (G2)

These steers comprised half-sib families produced by the artificial insemination of commercial and purebred Nelore dams, derived from 18 sires representing the main breeding lineages commercialized in Brazil. The 83 calves used in this expression study were allocated to feedlots in Embrapa Southeast Research. Measures of daily feed intake were collected for at least 70 days and body weight

was measured every 14 days. Liver samples were available for only 83 of the animals, which were ranked for RFI to select 20 animals genetically divergent for RFI.

The raw data of G2 were chosen from ENA repository (EMBL-EBI) under accession PRJEB7696. Total RNA from liver tissue samples of 20 Nelore animals divergent for RFI were selected. The RNA-Seq was performed using the HiSeq 2000 System (Illumina®). Paired-end reads (2 × 100 bp) were produced. All the information about the methodology used can be found in TIZIOTO et al. (2015).

2.3. Transcriptome analyses

Analyses were performed separately for each genetic group, using the same pipeline. All samples from G1 and G2 were compared using two groups; Low Residual feed Intake, LRFI (efficient) and High Residual Feed Intake , HRFI (inefficient) groups.

The quality of RNA-Seq reads (quality indices, GC content, N content, length distributions, duplication, overrepresented sequences and K-mer content) was verified using Fastqc software (v.0.11.4) (WINGETT; ANDREWS, 2018). The low quality reads (adapter sequence and reads containing poly-N) from raw data using Trimmomatic v.0.36 (BOLGER; LOHSE; USADEL, 2014) with following parameters: java -jar /usr/local/bin/trimmomatic-0.36.jar PE -phred33 <IN_R1.fastq.gz> <IN_R2.fastq.gz> <OUT_R1.fastq.gz> <OUT_R1_UN.fastq.gz> <OUT_R2.fastq.gz> <OUT_R2_UN.fastq.gz> LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50

All downstream analysis was based on the trimmed data with high quality reads. HISAT2 v.2.0.5 (KIM; LANGMEAD; SALZBERG, 2015) was used to align the paired-end trimmed reads to the bovine reference genome (ARS-UCD1.2; Bos taurus) deposited in National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/). This reads counting were carried out using the own script and Ensembl (Index of /pub/release-105/gtf/bos_taurus/ (ensembl.org)) gene annotations (scrip in Additional File 1).

EdgeR packages available from the Bioconductor project was used to estimation of normalized CPM reads (counts per million) and identify differentially expressed genes (DEGs) between RFI groups (OSHLACK; ROBINSON; YOUNG, 2010; LAW et al. 2016). We considered a gene to be differentially expressed when the p-value ≤0.05.

2.3.1. Non-coding RNAs analysis

An alignment of the paired-end trimmed reads were performed with the Bos taurus bovine reference genome ncRNAs file downloaded from the Ensembl database:https://www.ensembl.org/Bos_taurus/Info/Index through HISAT2 v.2.0.5 (KIM; LANGMEAD; SALZBERG, 2015).

This reads counting were carried out using the own script and Ensembl (Index of /pub/release-105/gtf/bos_taurus/ (ensembl.org)) gene annotations. Only reads showing a MAPQ equal to 0 were filtered out in order to remove noise related to multiple mapping sites. Counts for each non-coding RNA detected was obtained by using an in-house script, keeping only those non-coding RNAs with a mean count value higher or equal to 25 mean reads per sample.

The program was used to estimation of normalized CPMs reads (counts per million) and identify differentially expressed RNA non-coding between RFI groups was used the EdgeR packages available from the Bioconductor project (OSHLACK; ROBINSON; YOUNG, 2010; SMYTH et al., 2016). We considered a gene to be differentially expressed when the p-value ≤0.05.

2.3.2. Gene set enrichment analysis

To investigate biological processes and differentially expressed molecular functions (enrichment analysis) we used functional annotation gene set enrichment analysis, GSEA software (SUBRAMANIAN et al., 2005) using MSigDB database v7.5. Pathway Studio software v.10 (Elsevier Inc., Rockville, MD, USA) was used to identify the main biological processes and pathways with database ResNet v11.

2.3.3. Prediction of footprint gene profile

To obtain RFI predictors or biomarkers receiver operating characteristic curves (ROC) were generated. For analysis used genefilters package (Bioconductor) with the used function "rowpAUCs". The function has applied on the normalized values (counts) in two genetic goups (independent). Statistical correction (FDR) is not applied. Rowwise calculation of Receiver Operating Characteristic (ROC) curves and the corresponding partial area under the curve (pAUC) for a given data matrix or ExpressionSet. Cutpoints (theta are calculated before the first, in between and after the last data value. By default, both classification rules x > theta and x < theta are tested and the (partial) area under the curve of the better one of the two i returned. This is only valid for symmetric cases, where the classification is independent of the magnitude of x (e.g., both over- and under-expression of

different genes in the same class).

An AUC (Area under the ROC Curve) of 0.5 represents a test with no discriminating ability (ie, no better than chance), while an AUC of 1.0 represents a test with perfect discrimination (HOO; CANDLISH; TEARE, 2017). AUC of the proposed test was estimated to be >0.70.

Results

Alignment statistics

The alignment parameters pipeline were the same to genetic group 1 and genetic group 2. The mean number of paired end reads per sample after filtering was approximately 19.4 million reads for G1 and for G2 were 20 million reads. Of the trimmed reads were around 86% for G1 and around 88% G2 mapped with bovine reference genome ARS-UCD1.2 Bos Taurus (Additional Table 1).

The box plot containing CPMs for each group, the plot of principal component analysis (PCA), the correlation distance and Euclidian distance showed the formation of different groups (highest and lowest RFI), indicating differences in the expression of genes between the LRFI and HRFI groups (see Additional file: Figure S2). The distribution of quartiles, on box plot, was consistent between groups, indicating high quality of the data. In addition, the medians were similar in the two groups, indicating that the level of sequencing coverage permitted the identification of low-expressed genes.

Differential expression analysis of RNA seq data

Nelore animals from two genetic groups (G1 and G2), each divided into two groups, divergent for RFI, were compared. Results from the analysis of differential expression between LRFI and HRFI ($p \leq 0.05$) for G1 showed 1,811 DEGs (Table 2, Figures 1 and 2), 1,007 being underexpressed (55.6%) and 804 overexpressed (44.4%). Genetic group 2 showed 2,054 DEGs of which 1,140 were underexpressed (55.5%) and 914 were overexpressed (44.5%).

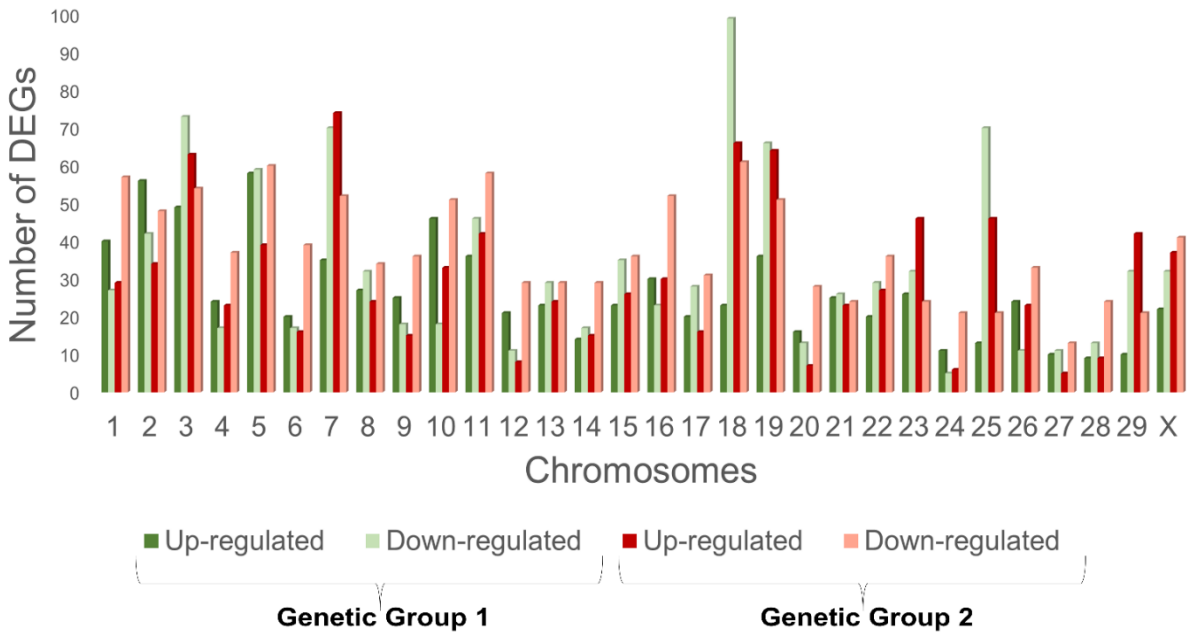


Figure 1. Distribution of the differentially expressed genes across the chromosomes. in the G1 and G2.

The 88 common DEGs obtained for both genetics groups with the same biological orientation (Fold-change) (Figure 3 and Table 2).

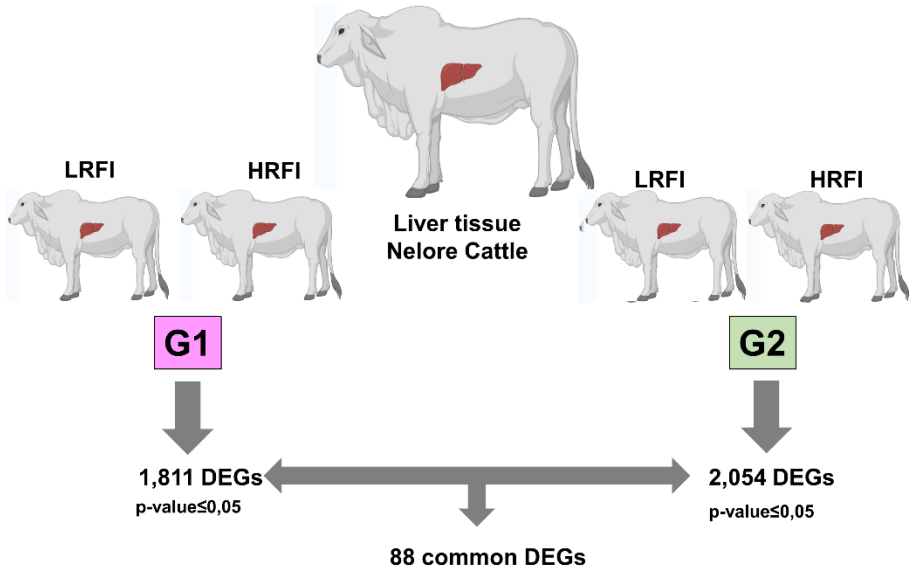


Figure 2. DEGs genes in G1 and G2 and common genes with the same biological orientation.

Table 2. Eighty-eight common DEGs in G1 and G2 for LRFI (p≤0.05).

			Genetic group 1			Genetic group 2		
Gene Name	**ID**	**Gene Symbol**	**Fold change (log2)***	**p-value**	**padj**	**Fold change (log2)***	**p-value**	**padj**
Ankyrin Repeat y SOCS Box Protein 8	ENSBTAG00000000289	*ASB8*	-0.1495	0.0253	0.571	-0.1683	0.0056	0.308
ATP Binding Cassette Subfamily F Member 3	ENSBTAG00000006920	*ABCF3*	-0.1319	0.0412	0.571	-0.1681	0.0056	0.308

Vascular cell adhesion molecule 1-like		*LOC534578*	0.8064	0.0011	0.571	0.4097	0.0253	0.445
Integral Membrane Protein 2B	ENSBTAG00000003109	*ITM2B*	0.1688	0.0025	0.571	0.1242	0.0308	0.462
Contactin 4	*ENSBTAG00000033225*	*CNTN4*	0.9359	0.0066	0.571	0.6749	0.0472	0.51
Pleckstrin Homology Domain Containing B2	ENSBTAG00000000941	*PLEKHB2*	0.3896	0.0078	0.571	0.1984	0.0323	0.468
Zinc finger protein 665	ENSBTAG00000000943	*ZNF286A*	-0.2862	0.0085	0.571	-0.1821	0.0373	0.483
Transmembrane And Coiled-Coil Domain Family 1	ENSBTAG00000006614	*TMCC1*	-0.2	0.0093	0.571	-0.2596	0.0009	0.22
Adenylosuccinate Synthase 2	ENSBTAG00000185100	*ADSS*	0.276	0.0097	0.571	0.1704	0.0267	0.451
Methionine Adenosyltransferase 2B	ENSBTAG00000012350	*MAT2B*	0.2139	0.0097	0.571	0.1131	0.0435	0.5
Major histocompatibility complex, class II, DR alpha Cyclin B3		*BoLA-DRB3*	0.4489	0.0098	0.571	0.3182	0.0262	0.45
Syntaxin 7	ENSBTAG00000017139	*STX7*	0.2613	0.0106	0.571	0.1475	0.0479	0.51
TNF Alpha Induced Protein 8	ENSBTAG00000014551	*TNFAIP8*	0.3635	0.0107	0.571	0.2624	0.009	0.35
F-Box Protein 15	ENSBTAG00000009686	*FBXO15*	-0.4387	0.0108	0.571	-0.4727	0.0367	0.48
LRR Binding FLII Interacting Protein 1	ENSBTAG00000005354	*LRRFIP1*	0.327	0.0108	0.571	0.2132	0.0062	0.312
Toll Like Receptor 5	ENSBTAG00000000477	*TLR5*	0.5309	0.0112	0.571	0.8541	0.0159	0.394
Joining Chain Of Multimeric IgA And IgM	ENSBTAG00000018531	*JCHAIN*	0.5275	0.0124	0.571	0.3372	0.0146	0.387
Uncharacterized		*LOC515089*	-0.2801	0.0124	0.571	-0.3338	0.0054	0.306
Cyclin B3	ENSBTAG00000046286	*CCNB3*	-0.3393	0.0131	0.571	-0.437	0.0166	0.396
Annexin A1	ENSBTAG00000015978	*ANXA1*	0.3743	0.0136	0.571	0.2777	0.0171	0.401
Inositol Polyphosphate-5-Phosphatase K	ENSBTAG00000011479	*INPP5K*	-0.1223	0.0138	0.571	-0.1309	0.0405	0.494
Adhesion G Protein-Coupled Receptor A3	ENSBTAG00000004653	*ADGRA3*	-0.1803	0.0139	0.571	-0.1609	0.0074	0.333
Zinc Finger CCHC-Type Containing 7	ENSBTAG00000025859	*ZCCHC7*	-0.2154	0.0143	0.571	-0.2452	0.0201	0.425
Lymphocyte Cytosolic Protein 1	ENSBTAG00000007079	*LCP1*	0.5266	0.0145	0.571	0.4347	0.0078	0.338
Major Histocompatibility Complex, Class II, DR Alpha	ENSBTAG00000010645	*BoLA-DRA*	0.4832	0.0157	0.571	0.2185	0.0499	0.512
CD74 Molecule	ENSBTAG00000015228	*CD74*	0.5054	0.0158	0.571	0.2425	0.0213	0.432
Transmembrane Protein 50A	ENSBTAG00000001296	*TMEM50A*	0.1392	0.016	0.571	0.1724	0.0416	0.497
Mannose Receptor C Type 2	ENSBTAG00000015739	*MRC2*	0.5687	0.0169	0.571	0.3	0.0424	0.5
Uncharacterized		*LOC788175*	0.5953	0.0178	0.571	0.3008	0.027	0.451
Adrenoceptor Alpha 1A	ENSBTAG00000031632	*ADRA1A*	-0.2719	0.0193	0.571	-0.2767	0.0002	0.158
Stathmin 1	ENSBTAG00000013761	*STMN1*	0.4345	0.0194	0.571	0.2428	0.0382	0.487
Helicase, Lymphoid Specific	ENSBTAG00000005979	*HELLS*	0.4887	0.0205	0.571	0.4734	0.0255	0.446
IFI30 Lysosomal Thiol Reductase	ENSBTAG00000018349	*IFI30*	0.4426	0.0206	0.571	0.2357	0.0311	0.463
Mitochondrial Trans-2-Enoyl-CoA Reductase Helicase, lymphoid-specific	ENSBTAG00000017253	*MECR*	-0.2213	0.0217	0.571	-0.2052	0.0095	0.353
DENN Domain Containing 10		*FAM45A*	0.2852	0.0223	0.571	0.1354	0.027	0.451
Sorting Nexin 2	ENSBTAG00000025358	*SNX2*	0.214	0.0231	0.571	0.1328	0.0299	0.46
Sorting nexin 6	ENSBTAG00000012873	*SNX6*	0.1672	0.0234	0.571	0.1112	0.0456	0.508
Rho Guanine Nucleotide Exchange Factor 5	ENSBTAG00000001815	*ARHGEF5*	-0.2641	0.024	0.571	-0.2617	0.04	0.493
Ribosomal Protein S6 Kinase A1	ENSBTAG00000014447	*RPS6KA1*	0.252	0.025	0.571	0.252	0.025	0.571
Protein Regulator Of Cytokinesis 1	ENSBTAG00000018643	*PRC1*	0.4199	0.0252	0.571	0.1908	0.023	0.434

Lumican	ENSBTAG00000001745	*LUM*	0.4358	0.0263	0.571	0.3917	0.039	0.488
Damage Specific DNA Binding Protein 2	ENSBTAG00000020999	*DDB2*	0.1129	0.0264	0.571	0.2711	0.0081	0.342
FYN Binding Protein 1	ENSBTAG00000001492	*FYB1*	0.3626	0.0266	0.571	0.192	0.024	0.437
Granulin Precursor	ENSBTAG00000018823	*GRN*	0.1808	0.0271	0.571	0.1586	0.0036	0.282
Pyruvate Kinase M1/2	ENSBTAG00000001601	*PKM*	0.3742	0.0274	0.571	0.1707	0.0364	0.479
Tetratricopeptide Repeat Domain 19	ENSBTAG00000013270	*TTC19*	-0.1468	0.0278	0.571	-0.1329	0.0498	0.512
ARV1 Homolog, Fatty Acid Homeostasis Modulator	ENSBTAG00000021822	*ARV1*	-0.1859	0.0288	0.571	-0.175	0.0308	0.462
BTB Domain Containing 6	ENSBTAG00000012751	*BTBD6*	-0.169	0.0291	0.571	-0.0768	0.0388	0.488
Vimentin	ENSBTAG00000018463	*VIM*	0.3079	0.0291	0.571	0.3593	0.0085	0.342
Uncharacterized		*LOC783680*	0.2802	0.0297	0.571	0.1726	0.0247	0.441
NPC Intracellular Cholesterol Transporter 2		*NPC2*	0.2977	0.0297	0.571	0.3	0.0424	0.5
Kelch Like Family Member 36	ENSBTAG00000008385	*KLHL36*	-0.1304	0.0301	0.571	-0.1356	0.0355	0.474
Lymphocyte Cytosolic Protein 2	ENSBTAG00000009381	*LCP2*	0.3456	0.0302	0.571	0.2466	0.0003	0.161
Rho GTPase Activating Protein 18	ENSBTAG00000015381	*ARHGAP18*	0.3437	0.0304	0.571	0.2553	0.0292	0.458
Calpain 2	ENSBTAG00000012778	*CAPN2*	0.2296	0.031	0.571	0.1569	0.0408	0.495
Thymus, Brain And Testes Associated	ENSBTAG00000021180	*TBATA*	-0.4652	0.0319	0.571	-0.5368	0.0426	0.5
Coactosin Like F-Actin Binding Protein 1	ENSBTAG00000016315	*COTL1*	0.2718	0.0321	0.571	0.1888	0.0162	0.395
LETM1 Domain Containing 1	ENSBTAG00000010924	*LETMD1*	-0.1713	0.0335	0.571	-0.1957	0.0442	0.504
Toll Like Receptor 2	ENSBTAG00000008008	*TLR2*	0.3937	0.0337	0.571	0.3998	0.0008	0.204
Cytochrome B-245 Beta Chain	ENSBTAG00000019953	*CYBB*	0.5508	0.0338	0.571	0.4176	0.0248	0.443
Amyloid Beta Precursor Protein Binding Family B Member 1 Interacting Protein	ENSBTAG00000012526	*APBB1IP*	0.3322	0.034	0.571	0.2228	0.0061	0.31
Uncharacterized		*LOC112446800*	-0.5062	0.0343	0.571	-0.5978	0.0305	0.461
P21 (RAC1) Activated Kinase 1	ENSBTAG00000010191	*PAK1*	0.318	0.0344	0.571	0.2334	0.0356	0.474
Thymosin Beta 4 X-Linked	ENSBTAG00000038488	*TMSB4X*	0.3005	0.035	0.571	0.3199	0.0032	0.277
Moesin	ENSBTAG00000003418	*MSN*	0.3817	0.0356	0.571	0.1346	0.0447	0.506
Nuclear Receptor Subfamily 4 Group A Member 2	ENSBTAG00000003650	*NR4A2*	0.4061	0.0358	0.571	0.5458	0.0214	0.433
Integrin Subunit Beta 2	ENSBTAG00000017060	*ITGB2*	0.4358	0.0359	0.571	0.2014	0.0401	0.493
Myosin IF	ENSBTAG00000007661	*MYO1F*	0.3206	0.0366	0.571	0.1629	0.0369	0.482
GDNF Inducible Zinc Finger Protein 1	ENSBTAG00000008908	*GZF1*	-0.1394	0.0369	0.571	-0.2676	0.0003	0.161
LDL Receptor Related Protein 5	ENSBTAG00000005903	*LRP5*	-0.1151	0.038	0.571	-0.0944	0.0411	0.495
Serine Palmitoyltransferase Small Subunit B	ENSBTAG00000049223	*SPTSSB*	0.4723	0.04	0.571	0.5892	0.0184	0.413
CD63 Molecule	ENSBTAG00000011931	*CD63*	0.0771	0.0402	0.571	0.0907	0.0178	0.408
Myosin Binding Protein H	ENSBTAG00000011465	*MYBPH*	0.9146	0.0402	0.571	0.6306	0.0169	0.399
Marginal Zone B And B1 Cell Specific Protein	ENSBTAG00000038337	*MZB1*	0.4825	0.0407	0.571	0.57	0.0191	0.418
Lymphocyte Antigen 9	ENSBTAG00000002947	*LY9*	0.3472	0.0409	0.571	0.3123	0.0053	0.306
SET Domain And Mariner Transposase Fusion Gene	ENSBTAG00000018935	*SETMAR*	-0.2426	0.0409	0.571	-0.2787	0.017	0.399

Cathepsin S	ENSBTAG00000017135	*CTSS*	0.4946	0.041	0.571	0.3211	0.0209	0.428
Ecotropic Viral Integration Site 2B	ENSBTAG00000009353	*EVI2B*	0.2939	0.0427	0.571	0.1647	0.0454	0.508
Solute Carrier Family 16 Member 9	ENSBTAG00000019792	*SLC16A9*	0.5142	0.0431	0.571	0.5586	0.0058	0.308
Parvin Gamma	ENSBTAG00000008057	*PARVG*	0.3136	0.0454	0.571	0.2878	0.0205	0.427
TATA-Box Binding Protein Associated Factor 5 Like	ENSBTAG00000135801	*TAF5L*	-0.1487	0.0454	0.571	-0.1746	0.0458	0.508
Beta-2-microglobulin	ENSBTAG00000012330	*B2M*	0.2578	0.046	0.571	0.2057	0.0115	0.365
Tubulin alpha 4a	ENSBTAG00000030974	*TUBA4A*	-0.2453	0.0461	0.571	-0.2353	0.0178	0.408
FGR Proto-Oncogene, Src Family Tyrosine Kinase	ENSBTAG00000011784	*FGR*	0.3737	0.0462	0.571	0.202	0.0284	0.456
Actin beta	ENSBTAG00000026199	*ACTB*	0.2713	0.0466	0.571	0.1693	0.0346	0.472
Rho GTPase Activating Protein 30	ENSBTAG00000017875	*ARHGAP30*	0.299	0.0469	0.571	0.1761	0.0333	0.469
Adaptor Related Protein Complex 4 Subunit Beta 1	ENSBTAG00000003614	*AP4B1*	-0.2135	0.0481	0.571	-0.2169	0.0301	0.46
Death Associated Protein Kinase 1	ENSBTAG00000000738	*DAPK1*	0.2284	0.0482	0.571	0.2626	0.0192	0.418

*The Fold-change estimates (relative expression) refer to the LRFI group.

Biomarkers

To find the most important biomarkers that describe our trait, we performed ROC curves using AUC >0.70. We found seven genes: *B2M, ADSS, SNX2, TUBA4A, ARHGAP18, MECR, ABCF3* that were on our list of the 88 DEGs (Table 2). The changes in expression and significance can be observed in Table 2. These biomarkers were associated with important biological processes, such as cell proliferation, insulin release, apoptosis, regulated cell death, lipid metabolism, fatty acid synthesis, lipid homeostasis, protein targeting, protein metabolic process, protein folding, glucose import, carbohydrate metabolism, reactive oxygen species generation, immune response, cell cycle, cell renewal and cell death (Figure 3).

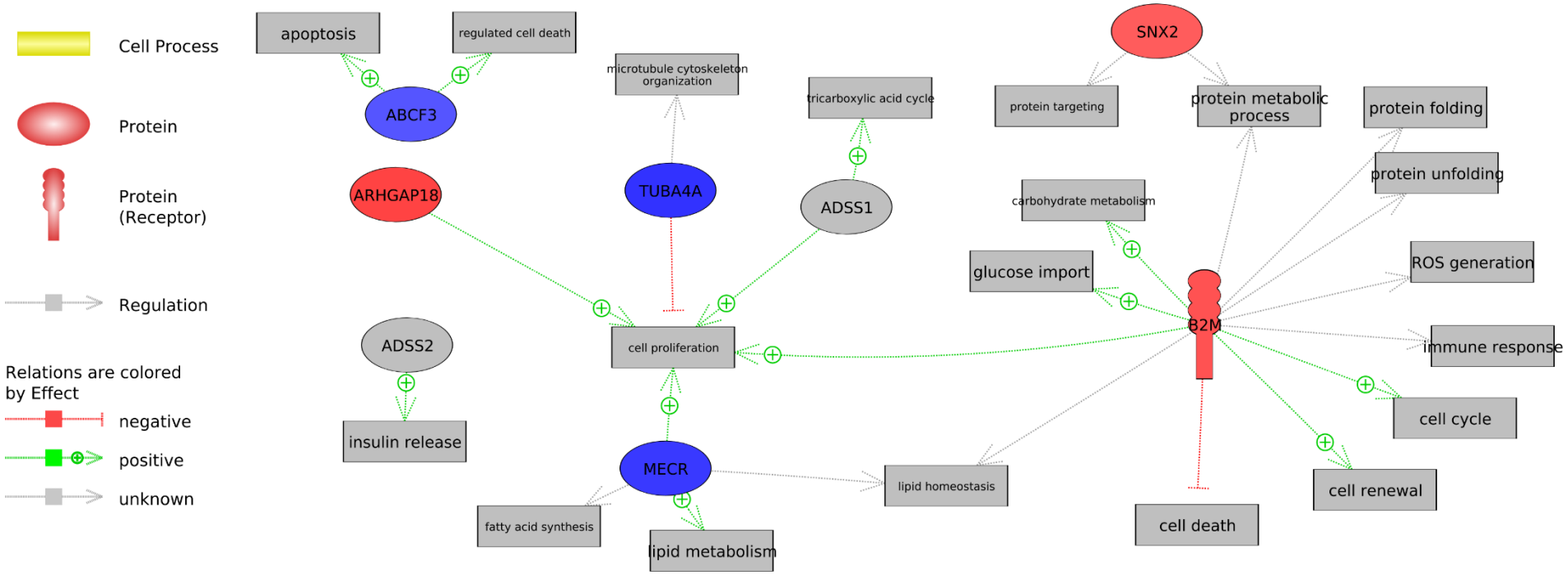


Figure 3. Biological processes associated with biomarkers related to the RFI trait. The genes highlighted in red are up-regulated in LRFI, the blue circles present the down-regulated genes in the LRFI group compared with the HRFI; the green arrows indicate a positive regulation effect, while the red arrows denote a negative regulation effect.

Gene set enrichment analysis

We followed two strategies to address the biological analysis, one using all transcripts with significant fold change for each genetic group and the other filtering the DEGs common in the two genetic groups (Table 2).

Biological analysis of the whole transcriptome data

The gene enrichment analysis showed 1,275 biological processes in G1 (1,163 with up-regulation and 112 with down-regulation), 124 molecular functions (95 down-regulated and 29 down-regulated), and 200 cellular components (148 down-regulated and 52 down-regulated) when comparing LRFI and HRFI. For G2, 375 biological processes were detected (214 with up-regulation and 161 with down-regulation), 65 molecular functions (38 with up-regulation and 27 with down-regulation) and 90 cellular components (38 with down-regulation and 52 with down-regulation) when comparing LRFI and HRFI.

Out of all the biological processes we highlight, the ones related to energy metabolism, protein turnover, redox homeostasis and immune system all common to G1 (see Additional Table 2) and G2 (see Additional Table 3). These biological processes were described previously in the analysis of biomarkers (see Figure 4).

The following describes each general biological process found and studied for each genetic group:

Energy Metabolism

Significant biological processes were identified for G1 related to negative energy balance in more efficient cattle like a negative regulation of energy derivation by oxidation of organic compounds (GO:0015980) through body use to produce energy and the adenosine triphosphate (ATP) metabolic process (GO:0046034) (Additional Table 2). Furthermore, in the liver synthesis of lipids, lipoproteins or phospholipids which are then used in the rest of the tissues of the organism. However, in the LRFI group an inhibited lipoprotein biosynthetic process was found (GO:0042158) as was the lipoprotein metabolic process (GO:0042157), oligosaccharide-lipid intermediate biosynthetic process (GO:0006490), and there may be less lipid accumulation in the more efficient group (LRFI). On the other hand, in the LRFI group we saw regulation of phospholipase activity (GO:0010517), regulation of lipid kinase activity (GO:0043550) and positive regulation of lipid kinase activity (GO:0090218) (Additional Table 2).

This tendency was also observed for G2 where we found, in LRFI animals,

inactive regulation of generation of precursor metabolites and energy (GO:0043467) even though we obtained a fatty acid derivative biosynthetic process activated (GO:1901570), lipopolysaccharide-mediated signaling pathway (GO:0031663) (Additional Table 3).

The liver regulates glucose in blood, converting glucose to glycogen or triacylglycerides or converting glycogen to glucose to elevate its level. When there is a lack of glucose and glycogen, the liver can convert amino acids and lipids to glucose. In the more efficient group (LRFI), an inactive glycogen biosynthetic process (GO:0005978), a glucan biosynthetic process (GO:0009250) and regulation of the glycolytic process (GO:0006110), the cellular response to glucose starvation (GO:0042149) can be observed (Additional Table 3).

Protein Turnover

The renewal of proteins (protein turnover) is related to both the synthesis of new proteins and the degradation of already existing proteins that provide essential components to synthesize other proteins.

The LRFI group of the G1 displayed tyrosine phosphorylation of STAT (GO:0007260) protein processes, regulation of protein serine/threonine kinase activity (GO:0071900), protein autophosphorylation (GO:0046777), activation of protein kinase activity (GO:0032147), positive regulation of protein-containing complex assembly (GO:0031334), positive regulation of cellular protein localization (GO:1903829), negative regulation of protein phosphorylation (GO:0001933), regulation of protein binding (GO:0043393) (Additional Table 2).

The LRFI group of G2 exhibited processes such as positive regulation of peptidase activity (GO:0010952), SRP-dependent cotranslational protein targeting to membrane (GO:0006614) in conjunction with essential proteins to carry out the translation in the ribosome and we found inactivated amino-acid betaine metabolic process (GO:0006577) in the LRFI (Additional Table 3).

Redox homeostasis

Various biological processes stand out in both genetic groups that are related to an increase in the oxidative processes in less efficient bovine.

Several biological processes related to the mitochondrial function were observed in G1: mitochondrial respiratory chain complex I assembly (GO:0032981), mitochondrial respiratory chain complex assembly (GO:0033108), mitochondrial ATP synthesis coupled with proton transport (GO:0042776), respiratory chain complex IV assembly (GO:0008535),

respiratory electron transport chain (GO:0022904) in the LRFI group. In addition, we found that the Wnt signaling pathway (GO:0016055) was activated in the LRFI group. The Wnt controls multiple processes of development, cell proliferation and apoptosis (GUPTA et al., 2014) (Additional Table 2).

We highlight several antioxidant processes in G2 in the LRFI group, such as mitochondrial electron transport, NADH to ubiquinone (GO:0006120), ATP synthesis coupled electron transport (GO:0042773), the glutathione derivative metabolic process (GO:1901685), the glutathione derivative biosynthetic process (GO:1901687), cellular detoxification (GO:1990748) and cellular response to toxic substances (GO:0097237). It is of utmost importance to maintain enough liver antioxidants given the large number of oxidation compounds in the liver that can trigger inflammation. In the LRFI group, there is a positive regulation of cysteine-type endopeptidase activity (GO:2001056), which is involved in the apoptotic process of the superoxide metabolic process (Additional Table 3).

Immune System

Additionally, hepatocytes are responsible for the synthesis of various enzymes and proteins with relevant functions in the immune system. Their main tasks are to ensure the peripheral immunotolerance of the organism with the help of hematopoietic cells and transforming growth factor- β. The liver participates in determining the shape of the immune response (JAKAB, 2015). We observed the activation of an immune response (GO:0002253) and an adaptive immune response (GO:0002250) in G1 and G2 (Additional Table 2 and Additional Table 3).

Biological analysis of differentially expressed genes

We present the biological analysis aimed at detecting the signaling pathways involving the 88 genes detected between both genetic groups (Table 2) showing relevant biological processes (Table 3). Results from the biological analysis showed processes that had already been observed in the analysis of the entire transcriptome, such as immune response (p-value $3.4x10^{-2}$) and actin cytoskeleton reorganization (p-value $1.7\ x10^{-2}$) (Table 3).

Table 3. The most representative biological processes based on enrichment p-value≤0.05.

Go term	Biological processes	Genes	p-value
GO:0071803	positive regulation of podosome assembly	*ARGEF5,LCP1,MSN*	8,0 x10^{-4}
GO:0045087	innate immune response	*FGR,ANXA1,B2M, CYBB,JCHAIN,LYA, TLR2,TLR5*	2,4x10^{-3}
GO:0051493	regulation of cytoskeleton organization	*ARGEF5,CAPN2,STMN1*	3,0x10^{-3}
GO:0008360	regulation of cell shape	*FGR,ARHGAP18,ANXA1, ITGB2,MSN*	3,1x10^{-3}
GO:0001666	response to hypoxia	*CAPN2,NR4A2,PAK1, PKM,TLR2*	6,4x10^{-3}
GO:0035556	intracellular signal transduction	*ARHGEP5,ADRAIA,ASB8, DAPK1,LCP1,ARS6KA1,STMN1*	7,6x10^{-3}
GO:0032366	intracellular sterol transport	*ARV1,NPC2*	8,6x10^{-3}
GO:0007409	axonogenesis	*CNTN4.LUM,PAK,STMN1*	8,6x10^{-3}
GO:0007165	signal transduction	*CD74,ARHGAP30,ADRAIA,APBBIKP, AMXA1,DAPK1,GRN,MRC2,NR4A2, RPS6KA1,STMN1,TLR2*	1,0x10^{-2}
GO:0006915	apoptotic process	*TNFAIP8,ADRAIA,DAPK1,ITGB2, MZB1,PAK1,RPRGKA1,TLR2*	1,1x10^{-2}
GO:0001895	retina homeostasis	*ACTB,B2M,JCHAIN*	1,3 x10^{-2}
GO:0031532	actin cytoskeleton reorganization	*ANXA1,PAK1,PARVG*	1,7 x10^{-2}
GO:0006886	intracellular protein transport	*CD74,AP4B1,SNX2, SNX6,STX7*	1,9 x10^{-2}
GO:0034123	positive regulation of toll-like receptor signaling pathway	*TLR2,TLR5*	2,1 x10^{-2}
GO:0006897	endocytosis	*LRP5,MRC2,SNX2,SNX6*	2,2 x10^{-2}
GO:0098609	cell-cell adhesion	*ABI3,LRRFIPI,ARHGAP18, PKM,SNX2*	2,9 x10^{-2}
GO:0006955	immune response	*CD74,B2M,CTSS,JCHAIN, LCP2,TLR2*	3,4 x10^{-2}
GO:0022617	extracellular matrix disassembly	*CAPN2,CTSS,LCP1*	4,2 x10^{-2}
GO:0050707	regulation of cytokine secretion	*TLR2,TLR5*	4,6 x10^{-2}

In addition, it delved into the genes that are integrated into the general processes highlighted in the analysis of all transcriptomes: energy homeostasis and lipid metabolism, the protein metabolic process, oxidative stress and the immune system (Figures 4, 5 and 6). Previously, in Figure 4, we show 33 genes involved in the immune response and oxidative stress, almost 38% of total 88 DEGs. We have potentially identified genes involved in antioxidant mechanisms that play a key role in hepatic metabolic adaptation by inhibiting oxidative stress, such as T*MSB4X, COTL1, RPS6KA1, VIM, STMN1, TNFAIP3, IFI30, ANXA1, TLR5* and others by activating the immune system, such as *MYO1F, ITM2B, CYBB, TLR2* and *LUM*.

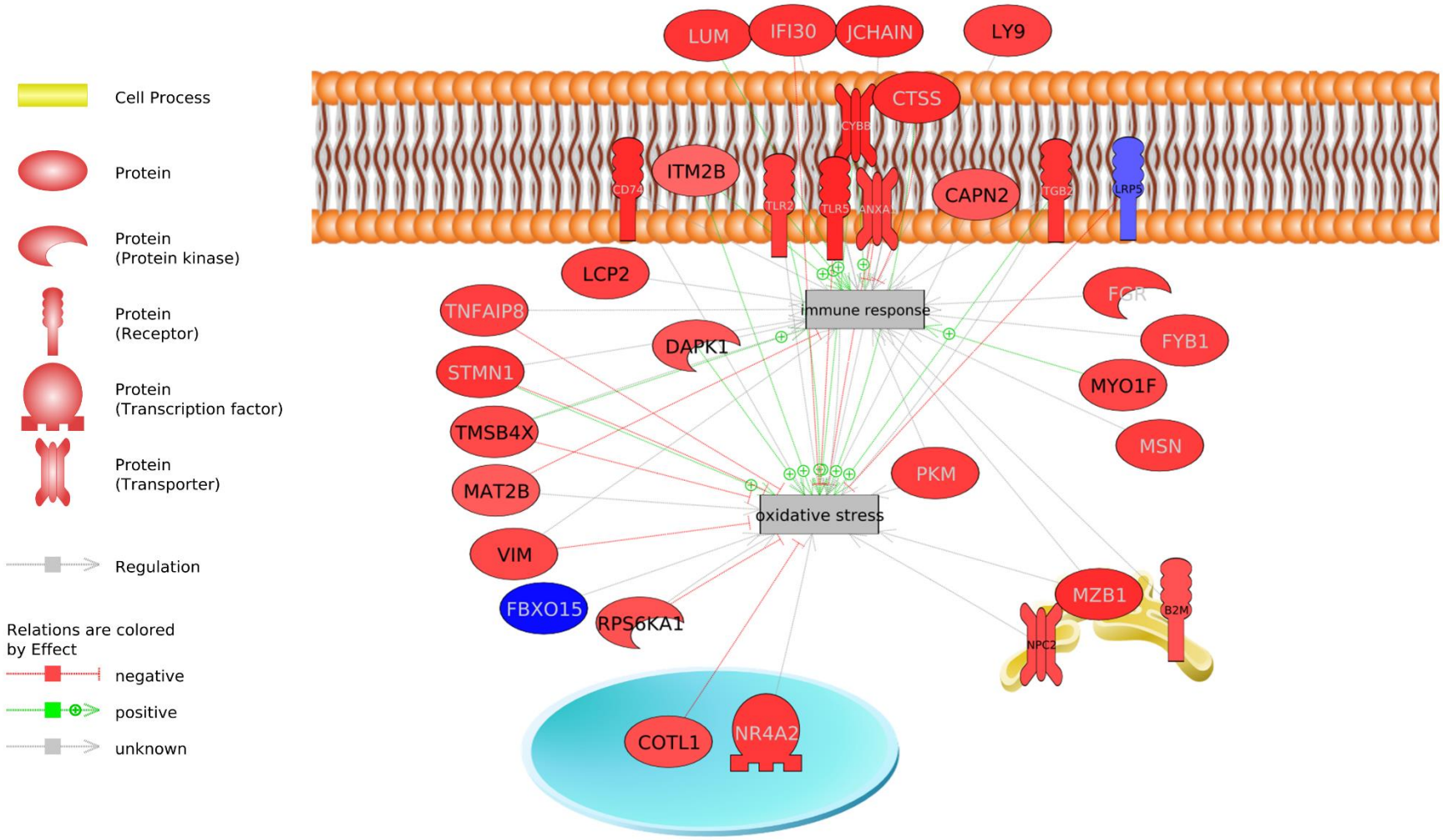


Figure 4. DEGs involved in immune response and oxidative stress identified using Pathway Studio software. Highlighted in red are genes overexpressed in LRFI and in blue are underexpressed genes in the LRFI group. Green arrows indicate a positive regulation effect, while red arrows has a negative regulation effect. and gray arrows has a unknown regulation effect.

Figure 5 shows filtered genes with a known regulatory effect which positively regulate the immune response (*LUM, CYBB, ITM2B, TLR2, MYO1F* and *TMSB4X*) and those that inhibit oxidative stress (*IFI30, ANXA1, TLR5, TMSB4X, COTL1, RPS6KA1, VIM, STMN1* and *TNFAIP8*).

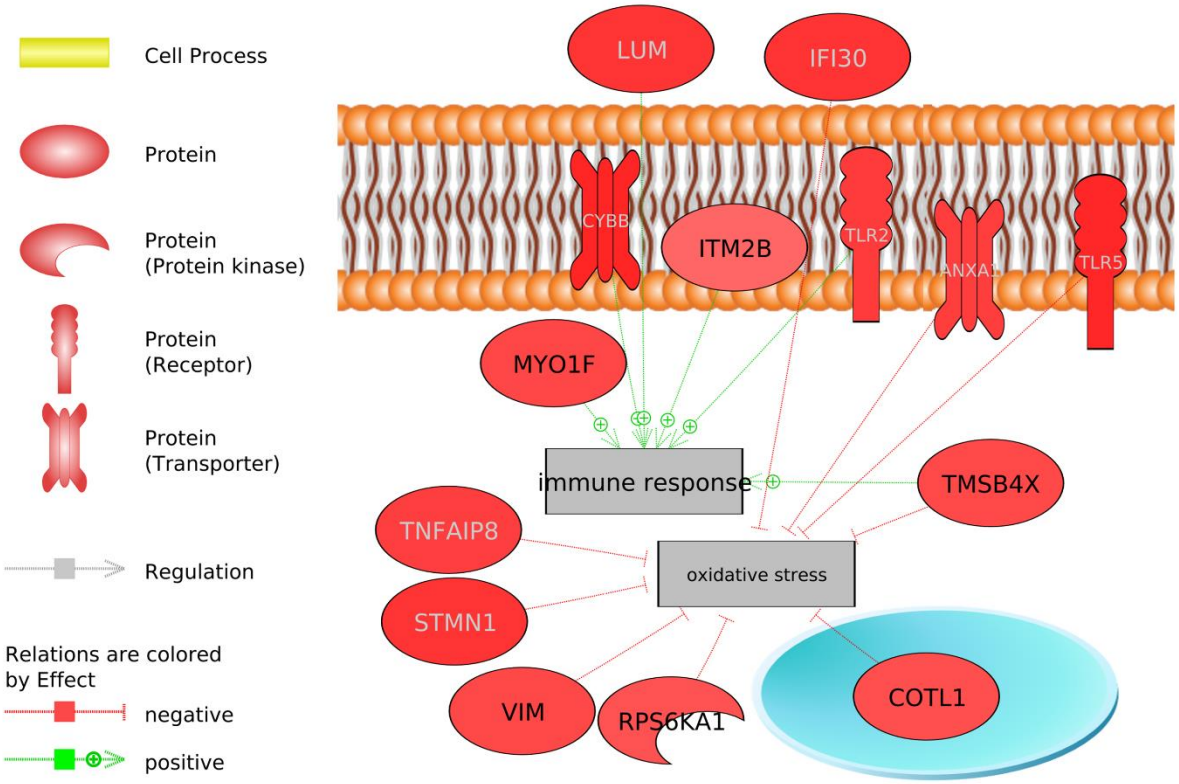


Figure 5. DEGs involved in immune response and oxidative stress with effect known using Pathway Studio software. Highlighted in red are genes overexpressed in LRFI and in blue are underexpressed genes in LRFI group. Green arrows has a known positive regulation effect and red arrows has a known negative regulation effect.

Figure 6 shows common DEGs in both genetic groups associated with the protein metabolic process (e.g., *B2M, CAPN2, SNX2*), with energy homeostasis (e.g., *NR4A2*) and lipid metabolism (e.g., *NPC2, ANXA1, STX7, TLR2, VIM, LRP5, ARV1* and *MECR*). *PKM* and *RPS6KA1* genes are involved in energy homeostasis and lipid metabolism and the *CTSS* gene is common in three biological processes.

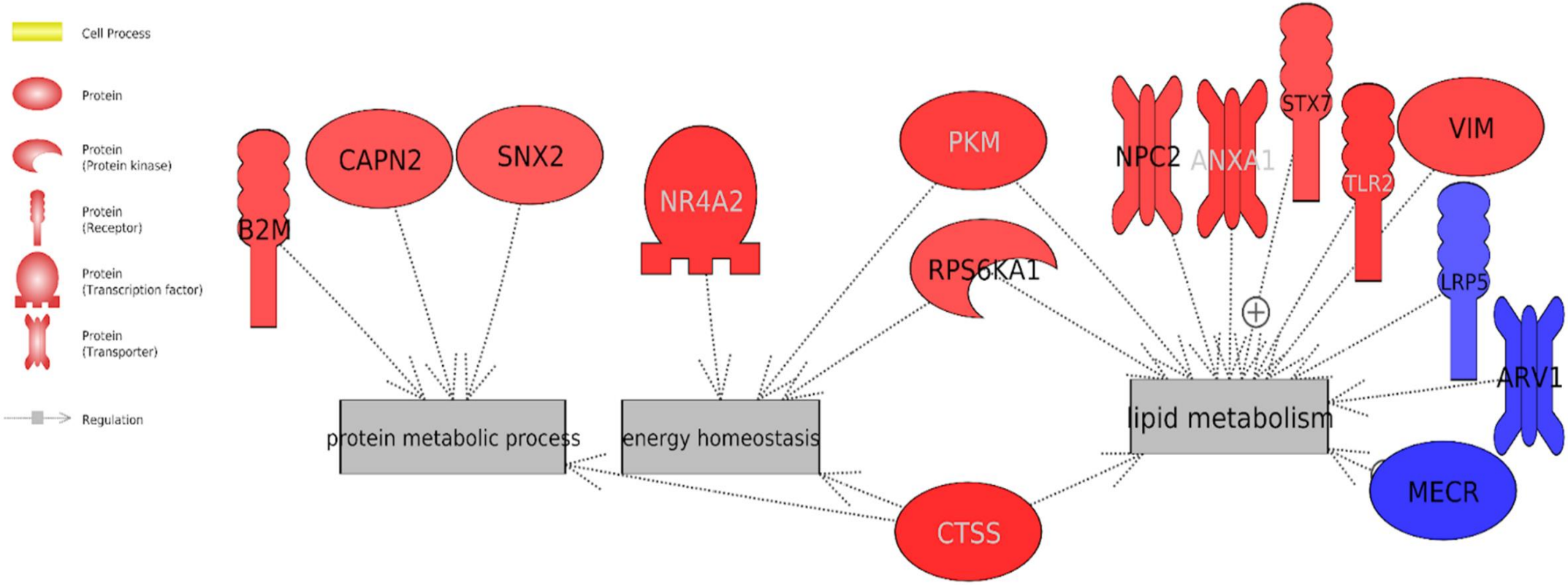


Figure 6. DEGs involved in protein metabolism, energy homeostasis and lipid metabolism using Pathway Studio software. Highlighted in red are genes overexpressed in LRFI and in blue are underexpressed genes in LRFI group. Gray arrows has a unknown regulation effect.

Analysis of non-coding RNAs

Transcriptome analysis was performed to investigate ncRNAs. For G1 were 226 non-coding RNAs and 49 non-coding RNAs for G2 (supplementary table 4). Applying the filter (25 mean count/samples) there were 14 for G1 and 4 for G2 (supplementary Table 5). Differential analysis, without filter, found two significant non-coding RNAs in LRFI group for G1: small nucleolar RNA, C/D box 16 (SNORD16) and microRNA 25 (MIR25). For G2, we found two significant non-coding RNAs in LRFI group; small Cajal body-specific RNA 10 (SCARNA10) and RMRP (RNase_MRP) (Table 2).

Table 2. The differential analysis of non-coding RNAs in G1 and G2.

Genetic group 1			
ID	**name**	**Fold change (log2)***	**p-value**
SNORD16	Small nucleolar RNA, C/D box 16	-0.356	0.01276
MIR25	MicroRNA 25	0.511	0.04741
Genetic group 2			
ID	**name**	**Fold change (log2)***	**p-value**
SCARNA10	Small Cajal body-specific RNA 10	0.636	0.01855
RNase_MRP	RNA component of mitochondrial RNA processing endoribonuclease	0.664	0.02635

*The fold-change estimates (relative expression) refer to the LRFI group.

Functional analysis of the non-coding

We showed cellular processes for all differentially expressed non-coding RNAs found in the G1 and G2 in LRFI (Figure 7). For G1, the MIR25 acts by blocking stress oxidative, cell death, toxicity and acts by activating cell survival, angiogenesis and cell proliferation, among others. The SNORD16 acts by blocking cell invasion, although it is also activates apoptosis. For G2, RMRP was associated with important cellular processes such as inactivation of mitochondrial damage, apoptosis, cellular aging and activation of cell growth and cell proliferation. The SCARNA10 acts by activating apoptosis.

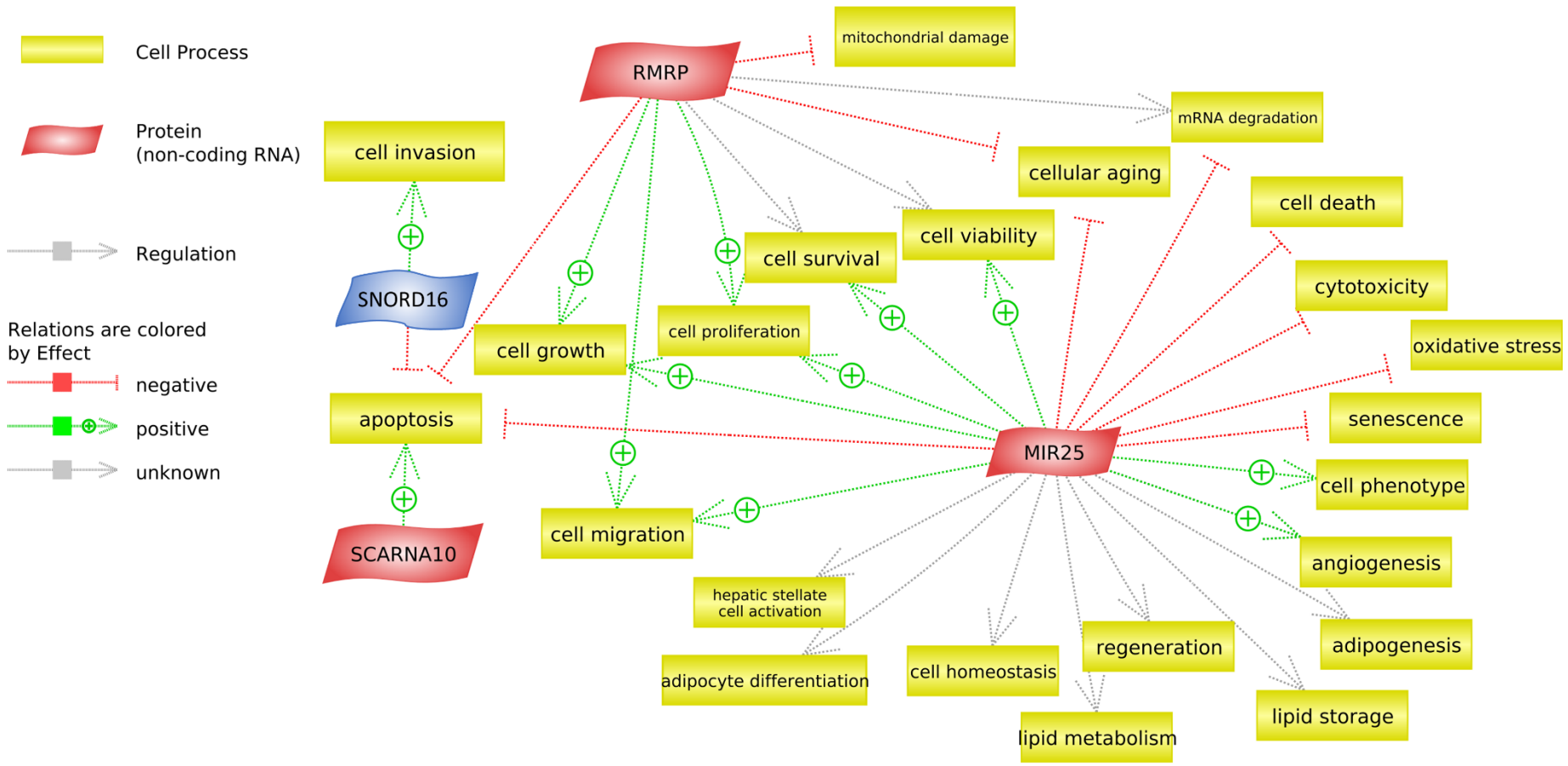


Figure 7. The main cellular processes for non-coding differentially expressed associated to RFI trait using Pathway Studio software. The non-coding RNA highlighted in red are up-regulated in LRFI, the blue circles present the down-regulated non-coding RNA in LRFI group compared with HRFI; green arrows indicate a positive regulation effect, while red arrows has a negative regulation effect.

Discussion

The results of this study may help to elucidate the biological processes involved in the liver of Nelore cattle. Steers with a low RFI had greater feed efficiency as they use energy to increase protein synthesis and turnover to the detriment of glucose and lipid utilization. In addition, a decrease in oxidative stress was observed in the LRFI groups, which confers greater liver protection against infections or pathologies and a better immune response. Moreover, an increase in antioxidant capacity or power, a greater cell proliferation capacity and protection against cell death were observed in the efficient group. Important biological processes were observed that explain the differences between groups with divergent RFI phenotypes.

Energy Metabolism

We observed the participation of several biological processes related to mitochondrial and lipid metabolism in G1 and G2 (see Additional Tables 2 and 3). Hepatocytes have great metabolic capacity contributing to a variation in energy utilization. They generate 90% of the energy produced in the body, mainly in the mitochondria (HEGDE; VOGEL; FEANY, 2014). Furthermore, they are responsible for metabolizing lipids, proteins and carbohydrates into biologically useful molecules (GRUMMER, 1993; CONNOR et al., 2009). BOTTJE et al.(2002) were the first to demonstrate that energy metabolism could be a true determinant of variation in feed efficiency. It is estimated that two-thirds of RFI variation is due to differences in resting energy expenditure (HERD; ARTHUR, 2009). Moreover, LRFI cattle consume 20% less food for the same level of yield (KELLY et al., 2010; FITZSIMONS; KENNY; MCGEE, 2014). The variation in the total energy expenditure of cattle of the same breed and of similar management may be due to differences in physiological processes, such as lipid metabolism (JOHNSON; FERRELL; JENKINS, 2003).

Both G1 and G2 presented down-regulation of the activation of MAPKK pathway activity (see Additional Table 2 and Table 3) in the LRFI group. The MAPK pathway regulates food consumption and energy expenditure (ASSIFI et al., 2005; KAHN et al., 2005). Inactivation in our more efficient group could result in inhibition of gluconeogenesis (LECLERC; KAHN; DOIRON, 1998) and lipogenesis (VIOLLET et al., 2006) which could favor fewer lipids being stored or formed in the liver, because constant activation of the MAKPP pathway causes fat accumulation (CARRASCO-CHAUMEL et al., 2005).

The animals in the HRFI group could consequently have an increased fat accumulation in hepatic cells, which can lead to the development of fatty liver (GOCERI et al., 2016), thereby exposing the animals to several metabolic stressors (DRACKLEY et al., 2016). In addition, we have observed a negative trend for lipid synthesis in LRFI animals which may be caused by processes that increase the flow and hepatic uptake of fatty acids and/or alter its metabolism (synthesis, oxidation or esterification) (LAGOMARSINO, 2012) (Figure 6). Our results are consistent with different studies of the liver transcriptome in bovine that dictate the great importance of lipid regulation for the RFI trait (ALEXANDRE et al., 2015; TIZIOTO et al., 2016), as well as those of a study on pig liver (LKHAGVADORJ et al., 2010).

Protein turnover

We observed a trend towards protein synthesis and turnover in the LRFI groups (Figure 6, Additional Tables 2 and 3) as other studies have already reported (CHEN et al., 2011; SANTANA et al., 2016). This suggests that feed-efficient animals divert the energy or nutrients consumed towards protein synthesis. This feature may give feed-efficient animals a great metabolic advantage over less efficient animals. Our results show a modulation in actin cytoskeleton reorganization (Additional Table 2 and 3 and Table 3), thereby contributing to muscle regeneration that is usually accompanied by proliferation and fusion into myotubes (ZANOU; GAILLY, 2013). It has been observed that a higher protein turnover results in a decrease in oxidative stress and a better efficiency in the use of energy in steers. (CASAL et al., 2020).

Redox homeostasis

Oxidative stress is caused by an imbalance between the production of reactive oxygen species (ROS) and the antioxidant capacity of the cell. Malfunction of these mechanisms can result in excess ROS production or a decrease in ROS reduction capacity, which in turn can cause lipid peroxidation and apoptosis (SU et al., 2019). In addition, ROS generated by mitochondria play an important role in the regulation of other proapoptotic proteins (OTT et al., 2007). Mitochondria are intrinsically associated with oxidative stress and cell signaling responsible for maintaining homeostasis and survival (HEKIMI; WANG; NOË, 2016; BRAVO-SAGUA et al., 2017; ANNESLEY; FISHER, 2019; MOEHLE; SHEN; DILLIN, 2019).

The more efficient RFI group (LRFI) would be more protected against oxidative stress since many genes involved in its inhibition have been observed (Figure 5), and they would be less likely to have liver lesions as reported by a study on Nelore cattle (ALEXANDRE et al., 2015). In the same way, a better ability to modulate oxidative stress was observed in more efficient animals (TIZIOTO et al., 2015; YANG et al., 2020) and decreased respiratory capacity and increased ROS production in HRFI animals (BOTTJE et al., 2002; FERGUSON et al., 2005; KONG et al., 2016). Previous studies have suggested a link between mitochondrial function and feed efficiency in broilers (BOTTJE et al., 2002; IQBAL et al., 2005). However, few studies have reported this link in beef cattle, especially with skeletal muscle tissue (ACETOZE et al., 2015, FONSECA et al., 2015) and in liver tissue (ACETOZE et al., 2015; FONSECA et al., 2015; CASAL et al., 2020).

This result is corroborated by the finding of overexpression of common genes among the genetic groups involved in the reduction of oxidative stress and may act as antioxidants in LRFI animals: *TMSB4X, COTL1, RPS6KA1, VIM, STMN1, TNFAIP3, IFI30, ANXA1, TLR5* (see Figure 6).

Antioxidants are molecules that slow down or inhibit the oxidation of diverse species and oxidative forms capable of attacking the organism, among which glutathione peroxidase (GPx) and glutathione (GSH) can be mentioned (ZUELKE et al., 2003). Our results demonstrate that the LRFI group experienced an increase in antioxidants such as glutathione (Additional Table 3). The glutathione family has previously been implicated in food efficiency in several species (BARENDSE et al., 2007; CHEN et al., 2011; LINDHOLM-PERRY et al., 2016; KONG et al., 2017). Furthermore, the participation of the Wnt pathway in the LRFI group was identified (Additional Table 2); this pathway acts as an antioxidant, anti-inflammatory, antitumor and hepatoprotective agent (ASHRAFIZADEH et al., 2020) and controls multiple development processes, such as apoptosis. Its dysregulation has been implicated in various pathologies (SEGDITSAS; TOMLINSON, 2006; GUPTA et al., 2014).

Our study has shown potential biomarkers in the LRFI and HRFI groups (Figure 3 and 6), highlighting *B2M* as the biomarker involved in most of the biological processes. *B2M* is overexpressed in LRFI animals when compared to the inefficient group. It has been identified in more efficient pigs, relating it to lipid metabolism (HORODYSKA et al., 2016). Furthermore, it plays a key role in lipid homeostasis, regulating its downstream target genes, such as FAS, and

accumulation of fatty acids and lipid droplets (HUANG; ZHAU; CHUNG, 2010). *B2M* reduced secretion of lipid binding protein-2 (GOOD et al., 2021) and it also participates in protein turnover (TAURINES et al., 2011). This gene regulates processes such as inhibition of cell death (LEÓN-LETELIER; BONIFAZ; FUENTES-PANANÁ, 2019). Moreover, it seems that its overexpression would help to reduce infections (CHEVALIEZ et al., 2008), as it plays an important role in the regulation of the immune system by binding to T cells and other immune receptors (KOLLNBERGER, 2016). Additionally, It has been described as a biomarker of kidney filtration and increased cell turnover (PRIZMENT et al., 2016), as has another biomarker, SNX2 (sorting nexin 2) (Figure 3) (ZHOU; LUO, 2018).

Immune response

Genetic group 1 showed a signaling pathway receptor activation via STAT that could come from the activation transforming growth factor beta production and subsequently activating the ERK1 and ERK2 cascade, resulting in greater cell proliferation (WEI; LIU, 2002). Growth factor pathways like IGF-1 have already been described as biomarkers of food efficiency (MONTELLI et al., 2021) whose function is linked, among others, to the immune system and tissue repair (PRIETO et al., 2002). In addition, the ERK 1 and 2 pathway is related to immunity regulation, stress response, survival, and apoptosis (SHAUL; SEGER, 2007).

Cytokines also activate the ERK 1 and 2 pathway and phosphorylation of threonine appears to be necessary to induce the synthesis of these cytokines (KONTNY et al., 2000). We found the modulation of this route (peptidyl-threonine phosphorylation) as well as the production of cytokines (Additional Table 2 and Table 3) in the two genetic groups. Moreover, this is corroborated by the analysis of biological processes of the genes common to both genetic groups (Table 3 and Figures 4 and 5). Activation of NF-κB was observed in the LRFI of both genetic groups (Additional Tables 2 and 3), which may play an important role in immune responses and stress and retain its impact in processes such as apoptosis, proliferation, differentiation and development (OECKINGHAUS; GHOSH, 2009).

Non-coding RNA

The central dogma of molecular biology states that RNA functions revolve around protein translation. The messenger RNAs (mRNAs) as temporary copies of genetic information, ribosomal RNAs (rRNAs) as a main component of ribosome, or translators of codon sequence (tRNAs). It is now known that this process represents less than 2% of the genome, and insufficiently explains the

functionality of 98% of the RNA transcripts (SAW et al., 2020). The ENCODE (Encyclopedia of DNA Elements) project was launched in 2005 and its recent reports revealed that up to 80% of the human genome has the capacity to transcribe into ncRNAs (DJEBALI et al., 2012; DUNHAM et al., 2012).

Our results show two non-coding RNAs in RLFI for G1: up-regulated MIR25 and down- regulated SNORD16., and two other non-coding RNAs in RLFI for G2: up-regulated RNase_MRP and SCARNA10. However, we detected that there are many common biological processes that can be regulated by these non-coding (Figure 7).

MicroRNAs are short non-coding RNAs that act by regulating the expression of genes that drive numerous cellular processes. MIR25 is involved in biological processes such as DNA damage response, cell cycle regulation, cell proliferation, migration and differentiation. In addition, many MIR25 target molecules can be found among extracellular matrix components and membrane receptors (SÁRKÖZY; KAHÁN; CSONT, 2018). It also has a role in the modulation of oxidative stress by targeting NADPH oxidase 4 (VARGA et al., 2013). Its overexpression protected cardiomyocytes against oxidative damage by inactivating the mitochondrial apoptosis pathway (PAN et al., 2015; MAGENTA et al., 2016; YAO; SUN; LEI, 2018). MIR25 can decrease p53 levels thus impairing downstream p53 effects including senescence, apoptosis and cell cycle arrest (CHEN et al., 2021).

In addition, MIR25 acts as a negative feedback anti-fibrotic control during hepatic stellate cell activation by reducing the reactivity of hepatic stellate cells to TGF-β-induced collagen expression and modulating the cross-talk between Notch, Wnt and TGF-β signaling (GENZ et al., 2019). Therefore, this gene could regulate mitochondrial functions in animals with low RFI. Several other studies were also associated mitochondrial function with RFI in cattle, included Nelore (FONSECA et al., 2015; CASAL et al., 2020; FERNÁNDEZ et al., 2020).

The other non-coding RNA were SNORD16, down-regulated in LRFI. The SnoRNAs are 60 to 300 nucleotides in length, are predominately found in the nucleolus, function mainly as guide RNAs for the posttranscriptional modification of ribosomal RNAs (rRNAs), The SNORD16 is categorized according to the modification type as the box H/ACA snoRNAs are associated with rRNA pseudouridylation (HE et al., 2020). Overexpression of snoRNAs plays a critical role in the onset and development of diseases, especially cancer (MIRISOLA et al., 2011; LÓPEZ-CORRAL et al., 2012). This non-coding may be helping the

immune system response to avoid diseases or help in the face of a dysregulation of the immune system.

SCARNA10 is a lncRNA transcript detected in G2. Hepatic fibrosis is characterized by the accumulation of excessive amounts of extracellular matrix components in the liver. This gene could have an importance in liver fibrosis (ZHANG et al., 2019c). However, our model and may be somehow regulating the extracellular matrix in LRFI group.

The last detected non-coding was RNase_MRP l in G2, RNA component of mitochondrial componente de ARN de la endoribonuclease (RMRP). RMRP is a recognized non-coding transcript (HUSSEN et al., 2021). RMRP inhibits lipopolysaccharide-induced apoptosis of cardiomyocytes and mitochondrial damage by suppressing the post-transcriptional regulatory function of miR-1-5p on HSPA4. RMRP lncRNA prevents mitochondrial dysfunction and cardiomyocyte apoptosis via the miR-1-5p/hsp70 axis in a sepsis model mice (HAN et al., 2020). Thus, RMRP gene mutations lead to decreased cell growth by impairing ribosomal assembly and altering cyclin-dependent cell-cycle regulation. Overexpressing of RMRP in human fibroblasts markedly elevates the production of cleaved or processed 5.8S ribosomal RNA and consequently, accelerates cellular growth rates (TRAINOR; MERRILL, 2014). In addition , RMRP inhibition attenuates lipid accumulation (YIN et al., 2022). This gene could be involved in mitochondrial functions, ribosomal assembly as well as helping the lipid accumulation in LRFI group.

In summary, we only found scientific evidence of the MIR25 in bovines, relating it to fertility trait (KROPP; SALIH; KHATIB, 2014; POHLER et al., 2017). This gene could be used as a biomarker for early pregnancy diagnosis in cattle (ONO et al., 2022).

Conclusion

Eighty-eight common DEGs were identified in the two genetic groups. The biomarkers relevant to our trait in liver tissue are identified as *B2M*, *ADSS, SNX2, TUBA4A, ARHGAP18, MECR, ABCF3*, the overexpressed *B2M* gene participating in most cellular processes that differentiate groups with greater or lesser food efficiency. The biological pathways associated with RFI of liver in Nelore cattle in the two genetic groups were energy metabolism, protein turnover, redox homeostasis and immune response. The common transcripts, biomarkers and metabolic pathways found in the two genetic groups make this

unprecedented work even more important, since the results are valid for different herds raised and slaughtered in different ways. We found up-regulated MIR25 and down-regulated SNORD16 non-coding RNAs in RLFI for G1. For G2, up-regulated RNase_MRP and SCARNA10. We highlight MIR25 being able to act by blocking cytotoxicity and oxidative stress and RMRP as a blocker of mitochondrial damage. The results reinforce the biological importance of these known processes but also reveal new insights into the complexity of the liver tissue transcriptome of Nelore cattle.

List of abbreviations

DEGs: Differentially expressed genes, Sustainable Development Goals (SDGs), Residual Feed Intake (RFI), ribonucleic acid (RNA), Low Residual feed Intake (LRFI), High Residual Feed Intake (HRFI), Fragments per kilobase of transcribed by millions of mapped fragments (FPKM), Counts per million (CPMs), Transcripts per million mapped reads (TPM), False discovery rates (FDR), Gene Set Enrichment Analysis (GSEA), Receiver operating characteristic (ROC), Area under the curve (AUC), Principal component analysis (PCA), Logarithm Of Ratio (LOR), Adenosine triphosphate (ATP), Nicotinamide adenine dinucleotide (NADH), Mitogen-activated protein kinase (MAPKK), Reactive oxygen species (ROS), Glutathione peroxidase (GPx), Fas Cell Surface Death Receptor (FAS), Signal Transducer and Activator of Transcription (STAT), Serine/threonine kinase 1(ERK1), Serine/threonine kinase 2 (ERK1), Insulin Like Growth Factor 1 (IGF-1), Nuclear factor kappa-light-chain-enhancer of activated B cells (NF-Kb).

Declarations

Ethics approval and consent to participate

All experimental procedures were approved by Ethics Committee of the School of Agricultural and Veterinarian Sciences, São Paulo State University (UNESP), Jaboticabal, SP, Brazil (protocol number 18.340/16).

Funding

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and São Paulo Research Foundation – FAPESP (grant #2009/16118-5, #2017/10630-2 e #2018/20026-8).

Acknowledgements

The authors thank the Qualitas Nelore breeding program company for providing the tissue samples and database used in this study.

References

ACETOZE, G. et al. Relationship between Liver Mitochondrial Respiration and Proton Leak in Low and High RFI Steers from Two Lineages of RFI Angus Bulls. **International Scholarly Research Notices**, v. 2015, p. 194014–194014, abr. 2015.

AGOSTINO, S. Di et al. Secreted Non-Coding RNAs: Functional Impact on the Tumor Microenvironment and Clinical Relevance in Triple-Negative Breast Cancer. **Non-Coding RNA** , v. 8, n. 1, p. 5, 11 jan. 2022.

AL-HUSSEINI, W. et al. Characterization and Profiling of Liver microRNAs by RNA-sequencing in Cattle Divergently Selected for Residual Feed Intake. **Asian-Australasian journal of animal sciences**, v. 29, n. 10, p. 1371–1382, 1 out. 2016.

ALEXANDRE, P. A. et al. Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. **BMC genomics**, v. 16, n. 1, dez. 2015.

ALEXANDRE, P. A. et al. Exploring the Regulatory Potential of Long Non-Coding RNA in Feed Efficiency of Indicine Cattle. **Genes**, v. 11, n. 9, p. 1–22, 1 set. 2020.

ANDERS, S.; PYL, P. T.; HUBER, W. HTSeq—a Python framework to work with high-throughput sequencing data. **Bioinformatics**, v. 31, n. 2, p. 166, jan. 2015.

ANNESLEY, S. J.; FISHER, P. R. Mitochondria in Health and Disease. **Cells,** v. 8, n. 7, p. 680, jul. 2019.

ASHRAFIZADEH, M. et al. Resveratrol targeting the Wnt signaling pathway: A focus on therapeutic activities. **Journal of cellular physiology**, v. 235, n. 5, p. 4135–4145, may. 2020.

ASSIFI, M. M. et al. AMP-activated protein kinase and the coordination of hepatic fatty acid metabolism of starved/carbohydrate-refed rats. **Am J Physiol Endocrinol Meta,** v. 289, n. 5 52-5, nov. 2005.

BARENDSE, W. et al. A validated whole-genome association study of efficient food conversion in cattle. **Genetics**, v. 176, n. 3, p. 1893–905, jul. 2007.

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, jun. 2003.

BENJAMINI, Y.; HOCHBERG, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. **Source: Journal of the Royal Statistical Society.** v. 57, n. 1, p. 289–300, jun. 1995.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, ago. 2014.

BOTTJE, W. et al. Association of mitochondrial function with feed efficiency within a single genetic line of male broilers. **Poultry Science**, v. 81, n. 4, p. 546–555, abr. 2002.

BRAVO-SAGUA, R. et al. Calcium Transport and Signaling in Mitochondria. **Comprehensive Physiology**, v. 7, n. 2, p. 623–634, abr. 2017.

BRAY, N. L. et al. Near-optimal probabilistic RNA-seq quantification. **Nature**

Biotechnology, v. 34, n. 5, p. 525–527, abr. 2016.

CARDOSO, D. F. et al. Genome-wide scan reveals population stratification and footprints of recent selection in Nelore cattle. **Genetics Selection Evolution**, v. 50, n. 1, p. 22, dez. 2018.

CARRASCO-CHAUMEL, E. et al. Adenosine monophosphate-activated protein kinase and nitric oxide in rat steatotic liver transplantation. **Journal of Hepatology**, v. 43, n. 6, p. 997–1006, dez. 2005.

CASAL, A. et al. Differential hepatic oxidative status in steers with divergent residual feed intake phenotype. **Animal : an international journal of animal bioscience**, v. 14, n. 1, p. 78–85, jan. 2020.

CHEN, Y. et al. Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. **Animal genetics**, v. 42, n. 5, p. 475–490, out. 2011.

CHEN, S. et al. Non-coding RNAs, guardians of the p53 galaxy. **Seminars in cancer biology**, v. 75, p. 72–83, 1 out. 2021.

CHEVALIEZ, S. et al. Role of class I human leukocyte antigen molecules in early steps of echovirus infection of rhabdomyosarcoma cells. **Virology**, v. 381, n. 2, p. 203–214, nov. 2008.

CONNOR, E. E. et al. Enhanced mitochondrial complex gene function and reduced liver size may mediate improved feed efficiency of beef cattle during compensatory growth. **Functional & Integrative Genomics**, v. 10, n. 1, p. 39–51, set. 2009.

DILLIES, M. A. et al. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. **Briefings in bioinformatics**, v. 14, n. 6, p. 671–683, nov. 2013.

DJEBALI, S. et al. Landscape of transcription in human cells. **Nature**, v. 489, n. 7414, p. 101–108, 6 set. 2012.

DRACKLEY, J. K. et al. Physiological and pathological adaptations in dairy cows that may increase susceptibility to periparturient diseases and disorders. **Italian Journal of Animal Science**, v. 4, n. 4, p. 323–344, jul. 2016.

DUNHAM, I. et al. An integrated encyclopedia of DNA elements in the human genome. **Nature**, v. 489, n. 7414, p. 57–74, 6 set. 2012.

EVERAERT, C. et al. Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. **Scientific Reports**, v. 7, n. 1, p. 1–11, may. 2017.

FERGUSON, M. et al. Age-associated decline in mitochondrial respiration and electron transport in Drosophila melanogaster. **Biochemical Journal**, v. 390, n. 2, p. 501–511, set. 2005.

FERNANDEZ, E.E. et al. Mitochondrial abundance and function in muscle from beef steers with divergent residual feed intakes. **Animal**, v.14, n. 3, p.560-565, fev. 2020

FITZSIMONS, C.; KENNY, D. A.; MCGEE, M. Visceral organ weights, digestion and

carcass characteristics of beef bulls differing in residual feed intake offered a high concentrate diet. **Animal**, v. 8, n. 6, p. 949–959, 2014.

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, fev. 2015.

FORESTRY, A. Residual Feed Intake (Net Feed Efficiency) in Beef Cattle. Agri-facts**. Open Government Licence - Alberta**. jul. 2006.

GAO, Y. et al. Analysis of Long Non-Coding RNA and mRNA Expression Profiling in Immature and Mature Bovine (*Bos taurus*) Testes. **Frontiers in Genetics**, v. 10, n. JUL, p. 646–646, 5 jul. 2019.

GENZ, B. et al. Overexpression of miRNA-25-3p inhibits Notch1 signaling and TGF-β-induced collagen expression in hepatic stellate cells. **Scientific reports**, v. 9, n. 1, 1 dez. 2019.

GOCERI, E. et al. Quantification of liver fat: A comprehensive review. **Computers in biology and medicine**, v. 71, p. 174–189, abr. 2016.

GOOD, S. C. et al. Global proteotoxicity caused by human β2 microglobulin variants impairs the unfolded protein response in c. elegans. **International Journal of Molecular Sciences**, v. 22, n. 19, out. 2021.

GRUMMER, R. R. Etiology of lipid-related metabolic disorders in periparturient dairy cows. **Journal of dairy science**, v. 76, n. 12, p. 3882–3896, jun. 1993.

HAN, Y. et al. lncRNA RMRP Prevents Mitochondrial Dysfunction and Cardiomyocyte Apoptosis via the miR-1-5p/hsp70 Axis in LPS-Induced Sepsis Mice. **Inflammation**, v. 43, n. 2, p. 605–618, 1 abr. 2020.

HE, J. Y. et al. Small Nucleolar RNA, C/D Box 16 (SNORD16) Acts as a Potential Prognostic Biomarker in Colon Cancer. **Dose-Response**, v. 18, n. 2, 1 abr. 2020.

HEGDE, V. R.; VOGEL, R.; FEANY, M. B. Glia are critical for the neuropathology of complex I deficiency in Drosophila. **Human Molecular Genetics**, v. 23, n. 17, p. 4686–4692, abr. 2014.

HEKIMI, S.; WANG, Y.; NOË, A. Mitochondrial ROS and the effectors of the intrinsic apoptotic pathway in aging cells: The discerning killers! **Frontiers in Genetics**, v. 7, n. SEP, p. 161, set. 2016.

HERD, R. M.; ARTHUR, P. F. Physiological basis for residual feed intake. **Journal of animal science**, v. 87, n. 14 Suppl, jun. 2009.

HOO, Z. H.; CANDLISH, J.; TEARE, D. What is an ROC curve? **Emergency medicine journal**, v. 34, n. 6, p. 357–359, jun. 2017.

HORODYSKA, J. et al. P3024 Transcriptome analysis of longissimus thoracis et lumborum from pigs divergent in residual feed intake. **Journal of Animal Science**, v. 94, n. suppl_4, p. 63–64, set. 2016.

HUANG, W. C.; ZHAU, H. E.; CHUNG, L. W. K. Androgen receptor survival signaling is blocked by anti-β2- microglobulin monoclonal antibody via a MAPK/lipogenic

pathway in human prostate cancer cells. **Journal of Biological Chemistry**, v. 285, n. 11, p. 7947–7956, mar. 2010.

HUSSEN, B. M. et al. Long Non-coding RNA RMRP in the Pathogenesis of Human Disorders. **Frontiers in Cell and Developmental Biology**, v. 9, p. 1130, 30 abr. 2021.

IQBAL, M. et al. Compromised liver mitochondrial function and complex activity in low feed efficient broilers are associated with higher oxidative stress and differential protein expression. **Poultry Science**, v. 84, n. 6, p. 933–941, jun. 2005.

JOHNSON, D. E.; FERRELL, C. L.; JENKINS, T. G. The history of energetic efficiency research: Where have we been and where are we going? **Journal of Animal Science**, v. 81, n. 13_suppl_1, p. E27–E38, jan. 2003.

KAHN, B. B. et al. AMP-activated protein kinase: Ancient energy gauge provides clues to modern understanding of metabolism. **Cell Metabolism**, v. 1, n. 1, p. 15–25, jan. 2005.

KELLY, A. K. et al. Effect of divergence in residual feed intake on feeding behavior, blood metabolic variables, and body composition traits in growing beef heifers. **Journal of animal science**, v. 88, n. 1, p. 109–123, jan. 2010.

KIM, D. et al. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. **Genome Biology**, v. 14, n. 4, abr. 2013.

KIM, D.; LANGMEAD, B.; SALZBERG, S. L. HISAT: A fast spliced aligner with low memory requirements. **Nature Methods**, v. 12, n. 4, p. 357–360, 31 mar. 2015.

KOCH, R. M. et al. Efficiency of Feed Use in Beef Cattle. **Journal of Animal Science**, v. 22, n. 2, p. 486–494, may. 1963.

KOLLNBERGER, S. The Role of HLA-Class I Heavy-Chain Interactions with Killer-Cell Immunoglobulin-Like Receptors in Immune Regulation. **Critical reviews in immunology**, v. 36, n. 3, p. 269–282, jul. 2016.

KONG, B. W. et al. Proteomics of Breast Muscle Tissue Associated with the Phenotypic Expression of Feed Efficiency within a Pedigree Male Broiler Line: I. Highlight on Mitochondria. **PLOS ONE**, v. 11, n. 5, p. 5679, may. 2016.

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, p. 1–19, jan. 2017.

KONTNY, E. et al. Rottlerin, a PKC isozyme-selective inhibitor, affects signaling events and cytokine production in human monocytes. **Journal of Leukocyte Biology**, v. 67, n. 2, p. 249–258, fev. 2000.

KROPP, J.; SALIH, S. M.; KHATIB, H. Expression of microRNAs in bovine and human pre-implantation embryo culture media. **Frontiers in genetics**, v. 5, n. APR, 2014.

LAGOMARSINO, TO. Liver regeneration in non-alcoholic fatty liver disease. **Medwave,**v. 12, n. 11, p. e5559–e5559, dez. 2012.

LAMARRE, S. et al. Optimization of an RNA-seq differential gene expression analysis depending on biological replicate number and library size. **Frontiers in Plant Science**, v. 9, p. 108, fev. 2018.

LAW C.W et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. **F1000Res**, v. 5. jun. 2016.

LECLERC, I.; KAHN, A.; DOIRON, B. The 5′-AMP-activated protein kinase inhibits the transcriptional stimulation by glucose in liver cells, acting through the glucose response complex. **FEBS Letters**, v. 431, n. 2, p. 180–184, jul. 1998.

LEI, Z. et al. ncRNAs regulate bovine adipose tissue deposition. **Molecular and cellular biochemistry**, v. 476, n. 7, p. 2837–2845, 1 jul. 2021.

LEÓN-LETELIER, R. A.; BONIFAZ, L. C.; FUENTES-PANANÁ, E. M. OMIC signatures to understand cancer immunosurveillance and immunoediting: Melanoma and immune cells interplay in immunotherapy. **Journal of leukocyte biology**, v. 105, n. 5, p. 915–933, may. 2019.

LINDHOLM-PERRY, A. K. et al. Differential gene expression in the duodenum, jejunum and ileum among crossbred beef steers with divergent gain and feed intake phenotypes. **Animal genetics**, v. 47, n. 4, p. 408–427, ago. 2016.

LKHAGVADORJ, S. et al. Gene expression profiling of the short-term adaptive response to acute caloric restriction in liver and adipose tissues of pigs differing in feed efficiency. **American journal of physiology**, v. 298, n. 2, fev. 2010.

LÓPEZ-CORRAL, L. et al. Genomic analysis of high-risk smoldering multiple myeloma. **Haematologica**, v. 97, n. 9, p. 1439–1443, 2012.

MAGENTA, A. et al. Oxidative stress, microRNAs and cytosolic calcium homeostasis. **Cell calcium**, v. 60, n. 3, p. 207–217, 1 set. 2016.

MARTIN, P. et al. Invited review: Disentangling residual feed intake—Insights and approaches to make it more fit for purpose in the modern context. **Journal of Dairy Science**, v. 104, n. 6, p. 6329–6342, jun. 2021.

Meat | OECD-FAO Agricultural Outlook 2020-2029 | OECD iLibrary. Disponível em: <https://www.oecd-ilibrary.org/sites/29248f46en/index.html?itemId=/content/component/29248f46-en>.

MIRISOLA, V. et al. A prognostic multigene classifier for squamous cell carcinomas of the larynx. **Cancer letters**, v. 307, n. 1, p. 37–46, 1 ago. 2011.

MOEHLE, E. A.; SHEN, K.; DILLIN, A. Mitochondrial proteostasis in the context of cellular and organismal health and aging. **Journal of Biological Chemistry**, v. 294, n. 14, p. 5396–5407, abr. 2019.

MONTELLI, N. L. L. L. et al. Associations of feed efficiency with circulating IGF-1 and leptin, carcass traits and meat quality of lambs. **Meat Science**, v. 173, p. 108379, mar. 2021.

NOLTE, W. et al. Identification and Annotation of Potential Function of Regulatory

Antisense Long Non-Coding RNAs Related to Feed Efficiency in Bos taurus Bulls. **International journal of molecular sciences**, v. 21, n. 9, 1 maio 2020.

OECKINGHAUS, A.; GHOSH, S. The NF-κB Family of Transcription Factors and Its Regulation. **Cold Spring Harbor Perspectives in Biology**, v. 1, n. 4, jul. 2009.

ONO, K. et al. Analysis of circulating microRNA during early gestation in Japanese black cattle. **Domestic animal endocrinology**, v. 79, 1 abr. 2022.

OSHLACK, A.; ROBINSON, M. D.; YOUNG, M. D. From RNA-seq reads to differential expression results**. Genome Biology**, dez. 2010.

OTT, M. et al. Mitochondria, oxidative stress and cell death. **Apoptosis 2007 12:5**, v. 12, n. 5, p. 913–922, fev. 2007.

PAN, L. et al. MiR-25 protects cardiomyocytes against oxidative damage by targeting the mitochondrial calcium uniporter. **International journal of molecular sciences**, v. 16, n. 3, p. 5420–5433, 10 mar. 2015.

POHLER, K. G. et al. Circulating microRNA as candidates for early embryonic viability in cattle. **Molecular reproduction and development**, v. 84, n. 8, p. 731–743, 1 ago. 2017.

PRIETO, M. et al. TGF-beta synthesis and mechanism of action. **Nephrology.** p. 135–143, jun. 2002.

PRIZMENT, A. E. et al. Circulating Beta-2 Microglobulin and Risk of Cancer: The Atherosclerosis Risk in Communities Study (ARIC). **Cancer Epidemiol Biomarkers Prev,** v. 25, n. 4, p.657-64, nov. 2016.

RÖNNAU, C. G. H. et al. Noncoding RNAs as Novel Biomarkers in Prostate Cancer. **BioMed Research International**, v. 2014, 2014.

SANTANA, M. H. A. et al. Systems genetics and genome-wide association approaches for analysis of feed intake, feed efficiency, and performance in beef cattle. **Genetics and Molecular Research**, v. 15, n. 4, out 2016.

SÁRKÖZY, M.; KAHÁN, Z.; CSONT, T. A myriad of roles of miR-25 in health and disease. **Oncotarget**, v. 9, n. 30, p. 21580, 20 abr. 2018.

SAW, P. E. et al. Non-coding RNAs: the new central dogma of cancer biology. **Science China Life Sciences 2020 64:1**, v. 64, n. 1, p. 22–50, 11 set. 2020.

SEGDITSAS, S.; TOMLINSON, I. Colorectal cancer and genetic alterations in the Wnt pathway. **Oncogene 2006 25:57**, v. 25, n. 57, p. 7531–7537, dez. 2006.

SEYEDNASROLLAH, F.; LAIHO, A.; ELO, L. L. Comparison of software packages for detecting differential expression in RNA-seq studies. **Briefings in Bioinformatics**, v. 16, n. 1, p. 59–70, jan. 2015.

SHAUL, Y. D.; SEGER, R. The MEK/ERK cascade: From signaling specificity to diverse functions. **Molecular Cell Research**, v. 1773, n. 8, p. 1213–1226, ago. 2007.

STARK, R.; GRZELAK, M.; HADFIELD, J. RNA sequencing: the teenage years. **Nature Reviews Genetics** , v. 20, n. 11, p. 631–656, jul. 2019.

SU, L. J. et al. Reactive Oxygen Species-Induced Lipid Peroxidation in Apoptosis, Autophagy, and Ferroptosis. **Oxidative Medicine and Cellular Longevity**, v. 2019, jun. 2019.

SUBRAMANIAN, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. **Proceedings of the National Academy of Sciences**, v. 102, n. 43, p. 15545–15550, out. 2005.

TAURINES, R. et al. Review: Proteomic research in psychiatry. **Journal of Psychopharmacology**, v. 25, n. 2, p. 151–196, fev. 2011.

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

TIZIOTO, P. C. et al. Gene expression differences in Longissimus muscle of Nelore steers genetically divergent for residual feed intake. **Scientific Reports 2016 6:1**, v. 6, n. 1, p. 1–12, dez. 2016.

TRAINOR, P. A.; MERRILL, A. E. Ribosome biogenesis in skeletal development and the pathogenesis of skeletal disorders. **Biochimica et Biophysica Acta - Molecular Basis of Disease**, v. 1842, n. 6, p. 769–778, 2014..

TRAPNELL, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. **Nature Protocols**, v. 7, n. 3, p. 562–578, mar. 2012.

TRAPNELL, C. et al. Differential analysis of gene regulation at transcript resolution with RNA-seq. **Nature Biotechnology**, v. 31, n. 1, p. 46–53, jan. 2013.

TREFTS, E.; GANNON, M.; WASSERMAN, D. H. The liver. **Current Biology**, v. 27, n. 21, p. R1147–R1151, nov. 2017.

VARGA, Z. V. et al. MicroRNA-25-dependent up-regulation of NADPH oxidase 4 (NOX4) mediates hypercholesterolemia-induced oxidative/nitrative stress and subsequent dysfunction in the heart. **Journal of Molecular and Cellular Cardiology**, v. 62, p. 111–121, 1 set. 2013.

VELONAS, V. M. et al. Current status of biomarkers for prostate cancer. **International Journal of Molecular Sciences**, v. 14, n. 6, p. 11034–11060, 2013.

VIOLLET, B. et al. Activation of AMP-activated protein kinase in the liver: a new strategy for the management of metabolic hepatic disorders. **The Journal of Physiology**, v. 574, n. 1, p. 41–53, jul. 2006.

UN. United Nations Sustainable Development – 17 Goals to Transform Our World. **UN Department of Economic and Social Affairs,** Disponível em: <https://www.un.org/sustainabledevelopment/>.

WAGNER, G. P.; KIN, K.; LYNCH, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. **Theory in**

biosciences, v. 131, n. 4, p. 281–285, dez. 2012.

WANG, J.; KOGANTI, P. P.; YAO, J. Systematic identification of long intergenic non-coding RNAs expressed in bovine oocytes. **Reproductive Biology and Endocrinology**, v. 18, n. 1, p. 1–9, 21 fev. 2020.

WEI, Z.; LIU, H. T. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. **Cell Research**, v. 12, n. 1, p. 9–18, jul. 2002

WINGETT, S. W.; ANDREWS, S. FastQ Screen: A tool for multi-genome mapping and quality control. **F1000Research**, v. 7, p. 1338, set. 2018.

WINKLE, M. et al. Noncoding RNA therapeutics — challenges and potential solutions. **Nature Reviews Drug Discovery 2021 20:8**, v. 20, n. 8, p. 629–651, 18 jun. 2021.

YANG, L. et al. Association of residual feed intake with growth performance, carcass traits, meat quality, and blood variables in native chickens. **Journal of animal science**, v. 98, n. 7, p. 121,apr. (2020).

YAO, Y.; SUN, F.; LEI, M. miR-25 inhibits sepsis-induced cardiomyocyte apoptosis by targetting PTEN. **Bioscience reports**, v. 38, n. 2, 13 abr. 2018.

YIN, J. et al. RMRP inhibition prevents NAFLD progression in rats via regulating miR-206/PTPN1 axis. **Mammalian genome : official journal of the International Mammalian Genome Society**, 10 fev. 2022.

ZANOU, N.; GAILLY, P. Skeletal muscle hypertrophy and regeneration: interplay between the myogenic regulatory factors (MRFs) and insulin-like growth factors (IGFs) pathways. **Cellular and molecular life sciences**, v. 70, n. 21, p. 4117–4130, nov. 2013.

ZHANG, K. et al. SCARNA10, a nuclear-retained long non-coding RNA, promotes liver fibrosis and serves as a potential biomarker. **Theranostics**, v. 9, n. 12, p. 3622, 2019c.

ZHENG, W.; CHUNG, L. M.; ZHAO, H. Bias detection and correction in RNA-Sequencing data. **BMC bioinformatics**, v. 12, jul. 2011.

ZHOU, J. J.; LUO, J. The PIN-FORMED Auxin Efflux Carriers in Plants. **International journal of molecular sciences**, v. 19, n. 9, set. 2018.

ZUELKE, K. A. et al. Glutathione (GSH) concentrations vary with the cell cycle in maturing hamster oocytes, zygotes, and pre-implantation stage embryos. **Molecular Reproduction and Development**, v. 64, n. 1, p. 106–112, jan. 2003.

Additional Files

Additional File: Table S1, Table S2, Table S3

Additional File: Figure S1, Figure S2.A-B

ADDITIONAL FILES

Figure S1. The own script used for reads count.

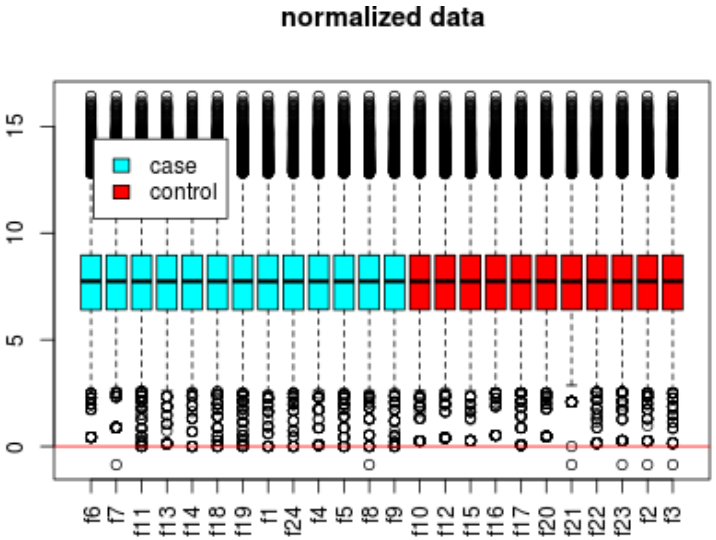
1. Download from Biomart database the following fields Chromosome, start, end, gene (mart_export.txt)
2. Information was splitted by chromosome (this action will reduce the computation times when annotating)
3. cut -f1,2,3,4,5 OUT.sam > OUT.txt (Read_name, flag, chromosome, position, MAPQ)
3. Python script for annotation

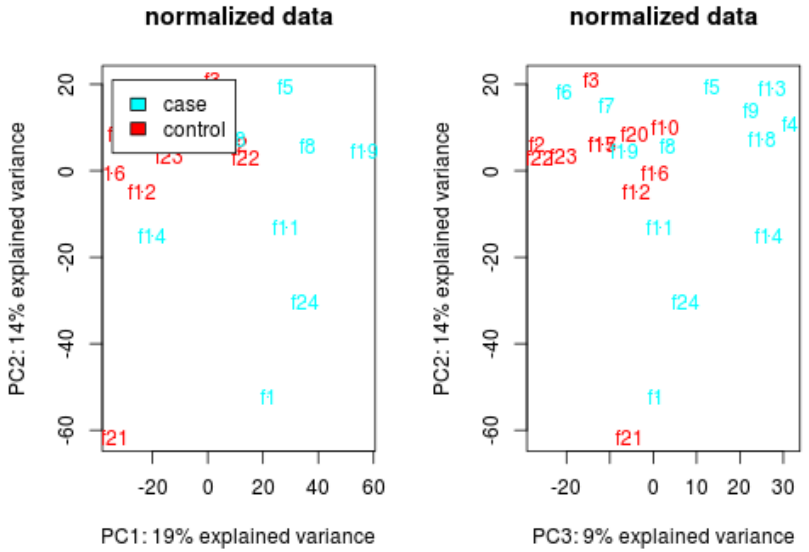
```
import os
from job import*
gene_list = []
folders = os.listdir(MAP_FOLDER_RESULTS)
for folder in folders:
    inFile = MAP_FOLDER_RESULTS + '/' + folder + '/OUT.txt'
    infile = open(MAP_FOLDER_RESULTS + '/' + folder + '/OUT.txt', 'r')
    lineas = sum(1 for line in open(inFile))
    for i in range (0,lineas,1):
        content = infile.readline()
        col = content.strip().split('\t')
        if int(col[4]) >= 5:
            chr = col[2]
            pos = int(col[3])
            infile_chr = open( 'INDEX_FOLDER/' + chr + '.txt', 'r')
            content_chr = infile_chr.readlines()
            for line in content_chr:
                col = line.strip().split('\t')
                start = int(col[3])
                end = int(col[4])
                if pos >= start and pos <= end:
                    info = col[8].strip().split(';')
                    gene = info[2].strip().split(' ' )
                    gene = gene[1].replace("","")
                    gene_list.append(gene)
            infile_chr.close()
uniq = list(set(gene_list))
outfile = open(MAP_FOLDER_RESULTS +  '/' + folder + '/conteos.txt', 'w')
for gene in uniq:
    n = gene_list.count(gene)
    outfile.write(gene + '\t' + str(n) + '\n')
outfile.close()
```

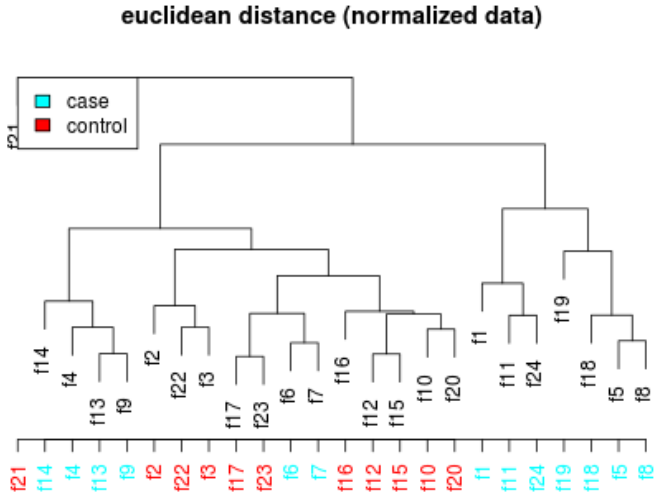
Table S1. Descriptive statistics for alignment parameter using HITSAT2**.**

GENETIC GROUP 1			
SEQUENCE ID	Raw reads	Clean Reads (Mb)	Overall alignment rate (%)
1	23,225,642	19,014,864	81.87013302
2	20,051,086	17,212,955	85.84549984
3	18,988,306	15,889,039	83.67802267
4	20,205,500	17,574,112	86.97687263
5	18,829,397	16,264,153	86.37638794
6	20,146,026	17,381,814	86.27912026
7	18,249,944	16,186,137	88.69143379
8	19,573,711	16,880,873	86.24257812
9	17,903,628	15,736,760	87.89704522
10	19,769,426	17,087,279	86.43285344
11	17,534,365	15,482,048	88.29545866
12	21,066,221	15,814,760	75.07165144
13	22,230,830	18,936,800	85.18260452
14	20,087,642	17,272,061	85.98351663
15	17,484,201	15,082,220	86.26199161
16	20,735,282	18,109,819	87.33818522
17	18,991,037	15,699,968	82.6704092
18	16,449,275	14,561,444	88.5233179
19	20,211,361	17,472,382	86.44831983
20	20,560,712	17,973,022	87.41439499
21	17,944,043	15,322,488	85.39038833
22	17,679,765	15,634,492	88.43156003
23	17,659,388	15,043,701	85.18812203
24	19,262,155	16,550,940	85.92465381
MEAN	19,368,289	16,591,005	85.66066527
GENETIC GROUP 2			
SEQUENCE ID	Reads raw	Reads cleanined (Mb)	Overall alignment rate (%)
1	8,205,798	7,254,724	88.40973175
2	12,342,938	10,774,148	87.2899791
3	11,174,277	9,899,320	88.59025063
4	9,411,232	8,328,025	88.49027417
5	10,210,766	9,017,716	88.31576397
6	10,441,025	9,227,088	88.37339246
7	10,691,977	9,361,547	87.55674465
8	8,200,461	7,328,352	89.36512228
9	10,407,413	9,191,569	88.31751945
10	9,680,276	8,322,409	85.97284829
11	10,977,085	9,747,828	88.80160808
12	9,609,337	8,581,017	89.29874142
13	9,711,313	8,588,163	88.43462259
14	8,758,872	7,734,901	88.30932796
15	10,250,216	8,932,195	87.14152951
16	9,346,752	8,236,241	88.1187497
17	12,612,367	11,168,260	88.55007153
18	8,100,136	7,143,194	88.18609959
19	9,168,261	8,114,508	88.50651176
20	9,867,034	8,820,733	89.39599276
MEAN	19,916,754	17,577,194	88.25330751

Figure S2. A. Global statistics and quality control for genetic group 1 data. (1) Box plot of CPMs distributions for individual conditions, (2) PCA plot , (3) Euclidean distance and (4) Correlation distance.







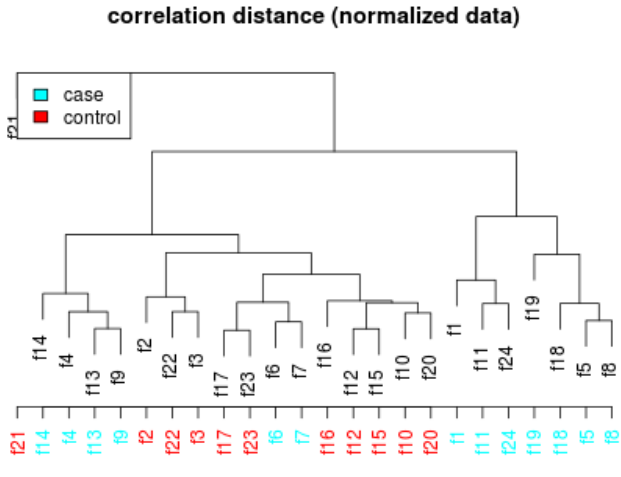
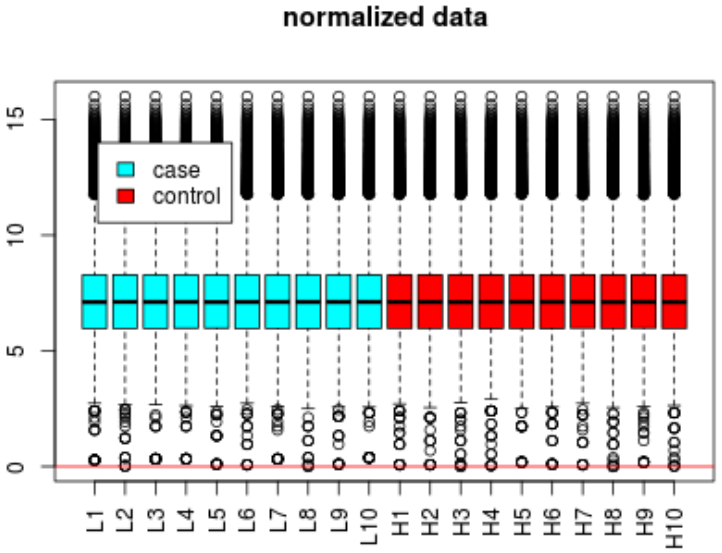
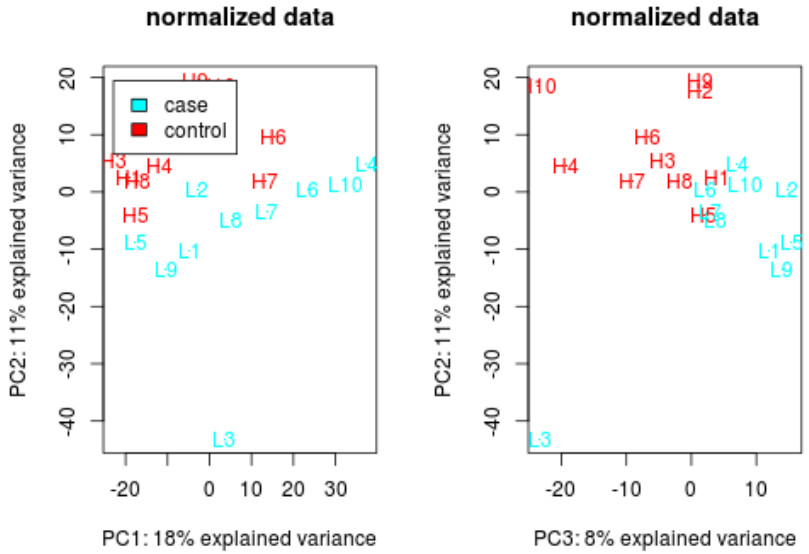
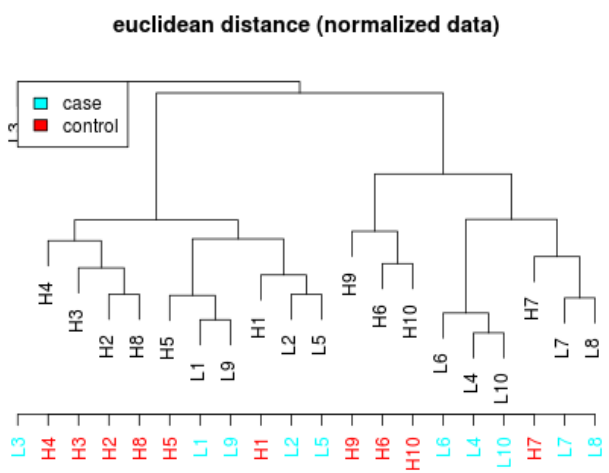


Figure S2. B. Global statistics and quality control for genetic group 2 data. (1) Box plot of CPMs distributions for individual conditions, (2) PCA plot , (3) Euclidean distance and (4) Correlation distance.







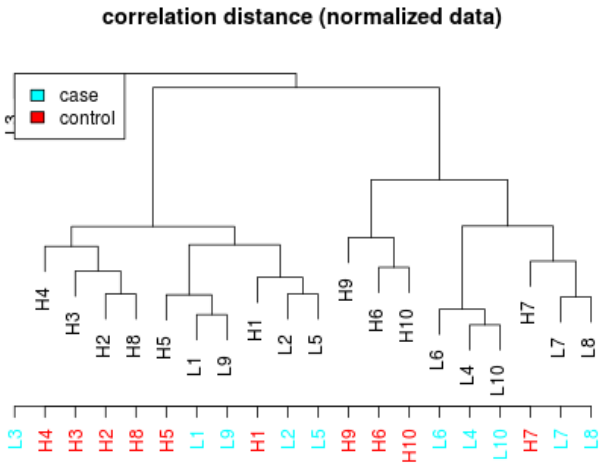


Table S2. The significant biological processes related to Energy. Protein turnover, and redox and immune system, in Genetic group 1 in the most efficient group.

Go Term	Number of genes	LOR*	p-value	padj	Biological process
GO:0006490	19	-0.866	2.3×10^{-04}	1.3×10^{-02}	oligosaccharide-lipid intermediate biosynthetic process
GO:0042158	93	-0.497	2.4×10^{-06}	2.6×10^{-04}	lipoprotein biosynthetic process
GO:0042157	121	-0.390	2.1×10^{-05}	1.3×10^{-04}	lipoprotein metabolic process
GO:0015980	271	-0.377	1.0×10^{-09}	2.5×10^{-07}	energy derivation by oxidation of organic compounds
GO:0006488	18	-0.928	1.4×10^{-04}	8.7×10^{-03}	dolichol-linked oligosaccharide biosynthetic process
GO:0010517	68	0.384	1.6×10^{-03}	6.7×10^{-02}	regulation of phospholipase activity
GO:0043550	63	0.429	7.5×10^{-04}	3.6×10^{-02}	regulation of lipid kinase activity
GO:0090218	37	0.560	7.5×10^{-04}	3.6×10^{-02}	positive regulation of lipid kinase activity
GO:0032981	58	-1.082	2.0×10^{-22}	3.6×10^{-19}	mitochondrial respiratory chain complex I assembly
GO:0033108	87	-1.039	3.5×10^{-28}	1.8×10^{-24}	mitochondrial respiratory chain complex assembly
GO:0042776	19	-0.938	7.0×10^{-06}	6.6×10^{-04}	mitochondrial ATP synthesis coupled proton transport
GO:0008535	21	-0.895	9.4×10^{-06}	8.5×10^{-04}	respiratory chain complex IV assembly
GO:0042773	81	-0.833	1.2×10^{-12}	5.1×10^{-10}	ATP synthesis coupled electron transport
GO:0022904	100	-0.772	8.9×10^{-14}	4.4×10^{-11}	respiratory electron transport chain
GO:0007260	75	0.380	1.1×10^{-03}	4.8×10^{-02}	tyrosine phosphorylation of STAT protein
GO:0071900	492	0.369	1.6×10^{-15}	9.7×10^{-13}	regulation of protein serine/threonine kinase activity
GO:0046777	236	0.474	4.5×10^{-13}	2.1×10^{-10}	protein autophosphorylation
GO:0032147	324	0.387	6.8×10^{-12}	2.6×10^{-09}	activation of protein kinase activity
GO:0031334	234	0.440	4.0×10^{-11}	1.4×10^{-08}	positive regulation of protein-containing complex assembly
GO:1903829	321	0.317	2.3×10^{-08}	4.2×10^{-06}	positive regulation of cellular protein localization
GO:0043393	209	0.265	1.4×10^{-04}	8.8×10^{-03}	regulation of protein binding
GO:0071634	35	0.721	1.2×10^{-05}	1.1×10^{-03}	regulation of transforming growth factor beta production
GO:0018107	114	0.385	5.4×10^{-05}	3.8×10^{-03}	peptidyl-threonine phosphorylation
GO:0007223	37	0.561	1.0×10^{-04}	2.8×10^{-02}	Wnt signaling pathway, calcium modulating pathway
GO:1901224	75	0.493	$6,4 \times 10^{-05}$	2.0×10^{-03}	positive regulation of NIK/NF-kappaB signaling
GO:0000186	55	-0.572	4.2×10^{-05}	9.8×10^{-03}	activation of MAPKK activity
GO:0070371	284	0.378	4.2×10^{-10}	1.1×10^{-07}	ERK1 and ERK2 cascade

GO:0002764	353	0.569	8.0×10^{-25}	2.6×10^{-21}	immune response-regulating signaling pathway
GO:0002253	402	0.471	7.6×10^{-25}	7.8×10^{-17}	activation of immune response
GO:0045785	386	0.568	1.9×10^{-27}	8.6×10^{-24}	positive regulation of cell adhesion
GO:0001819	356	0.534	2.1×10^{-22}	3.6×10^{-19}	positive regulation of cytokine production
GO:0045787	359	0.288	7.4×10^{-08}	1.2×10^{-05}	positive regulation of cell cycle
GO:0097696	139	0.325	1.5×10^{-04}	9.3×10^{-03}	receptor signaling pathway via STAT

*The LOR (Logarithm of Ratio) refer to LRFI group, this value categorizes each biological process in a classification in relation to the rest (total), allowing to identify biological processes, molecular mechanisms and significant cellular components and their regulation within the groups.

Table S3. Significant biological processes related to energy. turnover protein, and redox and immune system, in genetic group 2 in the most efficient group.

Go Term	Number of genes	LOR*	p-value	padj	Biological process
GO:1901570	88	0.397	2.2×10^{-04}	3.7×10^{-02}	fatty acid derivative biosynthetic process
GO:0031663	54	0.570	5.9×10^{-05}	1.3×10^{-02}	lipopolysaccharide-mediated signaling pathway
GO:0005978	49	-0.549	1.5×10^{-04}	2.7×10^{-02}	glycogen biosynthetic process
GO:0009250	49	-0.549	1.5×10^{-04}	2.7×10^{-02}	glucan biosynthetic process
GO:0006110	72	-0.524	1.5×10^{-05}	4.7×10^{-03}	regulation of glycolytic process
GO:0002718	63	0.474	3.3×10^{-04}	5.1×10^{-03}	regulation of cytokine production involved in immune response
GO:0042149	40	-0.697	8.7×10^{-06}	2.9×10^{-02}	cellular response to glucose starvation
GO:0006120	46	1.141	2.9×10^{-12}	6.8×10^{-09}	mitochondrial electron transport, NADH to ubiquinone
GO:0042773	81	-1.199	2.5×10^{-07}	1.7×10^{-03}	ATP synthesis coupled electron transport
GO:1901685	19	1.047	1.3×10^{-07}	1.0×10^{-04}	glutathione derivative metabolic process
GO:1901687	19	1.047	1.3×10^{-07}	1.0×10^{-04}	glutathione derivative biosynthetic process
GO:1990748	102	0.362	2.6×10^{-04}	4.2×10^{-02}	cellular detoxification
GO:0097237	110	0.357	1.9×10^{-04}	3.2×10^{-02}	cellular response to toxic substance
GO:0006614	94	1.139	6.3×10^{-18}	1.2×10^{-13}	SRP-dependent cotranslational protein targeting to membrane
GO:2001056	135	0.408	2.9×10^{-06}	1.3×10^{-03}	positive regulation of cysteine-type endopeptidase activity
GO:0010952	172	0.350	6.4×10^{-06}	2.4×10^{-03}	positive regulation of peptidase activity
GO:0006577	17	-0.736	1.2×10^{-03}	1.4×10^{-01}	amino-acid betaine metabolic process
GO:0052548	358	0.225	2.6×10^{-05}	6.6×10^{-03}	regulation of endopeptidase activity

GO:2001056	135	0.408	2.9×10^{-06}	1.3×10^{-03}	positive regulation of cysteine-type endopeptidase activity
GO:0043280	118	0.408	1.2×10^{-05}	3.8×10^{-03}	positive regulation of cysteine-type endopeptidase activity involved in apoptotic process
GO:0007249	253	0.222	4.7×10^{-04}	6.5×10^{-02}	I-kappaB kinase/NF-kappaB signaling
GO:0043123	176	0.276	2.9×10^{-04}	4.5×10^{-02}	positive regulation of I-kappaB kinase/NF-kappaB signaling
GO:0000186	55	-0.572	4.2×10^{-05}	9.8×10^{-03}	activation of MAPKK activity
GO:0032963	98	0.594	7.8×10^{-09}	8.8×10^{-06}	collagen metabolic process
GO:0030574	42	0.619	5.2×10^{-05}	1.1×10^{-03}	collagen catabolic process
GO:2000251	19	0.827	2.1×10^{-04}	3.5×10^{-02}	positive regulation of actin cytoskeleton reorganization
GO:0001516	24	0.794	1.6×10^{-04}	2.8×10^{-02}	prostaglandin biosynthetic process
GO:0018107	114	-0.364	1.4×10^{-04}	2.5×10^{-02}	peptidyl-threonine phosphorylation
GO:0052547	383	0.238	4.6×10^{-06}	1.9×10^{-03}	regulation of peptidase activity
GO:0002283	449	0.328	1.3×10^{-06}	2.6×10^{-08}	neutrophil activation involved in immune response
GO:0002446	460	0.329	6.4×10^{-12}	1.4×10^{-08}	neutrophil mediated immunity
GO:0002694	476	0.283	1.6×10^{-09}	2.1×10^{-06}	regulation of leukocyte activation
GO:0050867	306	0.349	2.3×10^{-09}	3.0×10^{-08}	positive regulation of cell activation
GO:0002250	352	0.279	2.8×10^{-07}	$1\text{-}9 \times 10^{-04}$	adaptive immune response
GO:0042110	431	0.226	3.8×10^{-06}	1.6×10^{-03}	T cell activation
GO:0033006	30	0.863	7.5×10^{-06}	2.6×10^{-03}	regulation of mast cell activation involved in immune response
GO:0002697	351	0.241	8.7×10^{-07}	2.9×10^{-03}	regulation of immune effector process

*The LOR (Logarithm Of Ratio) refer to LRFI group, this value positions each biological process in a classification in relation to the rest (total), being able to identify biological processes, molecular mechanisms and significant cellular components and their regulation within the groups.

Table S4. Non-coding RNA identified in liver tissue of Nelore bulls selected to be divergent for RFI in G1 applying the filter (25 mean count/samples).

ID	10F_q1	11F_q1	12F_q1	13F_q1	14F_q1	15F_q1	16F_q1	17F_q1	18F_q1	19F_q1	1F_q1	20F_q1	21F_q1	22F_q1	23F_q1	24F_q1	2F_q1	3F_q1	4F_q1	5F_q1	6F_q1	7F_q1	8F_q1	9F_q1
bta-mir-3533	24	27	52	125	38	44	33	29	33	94	117	36	35	54	34	111	80	39	23	75	46	42	50	40
U6	34	41	47	58	23	48	40	41	49	61	20	21	39	38	55	24	39	39	56	47	47	34	37	47
U1	37	36	29	53	19	39	40	45	27	50	13	7	27	50	36	41	38	49	31	69	26	40	21	30
U3	115	78	82	222	86	123	166	147	113	175	396	72	89	130	169	263	254	138	187	136	93	107	165	109
U2	49	109	64	90	71	92	77	132	61	59	136	100	65	48	71	151	149	74	106	58	35	94	159	124
7SK	278	342	219	729	347	439	410	338	333	525	314	667	351	432	418	869	404	533	558	377	293	873	594	286
5_8S_rRNA	2038	1755	1291	3470	1617	2190	1925	2566	1486	2507	1621	8242	1031	1947	2483	5597	2418	1729	2935	952	708	3625	2524	943
bta-mir-3064	17	43	43	38	30	49	47	80	28	50	15	2	22	34	40	31	61	26	35	48	36	35	40	25
EEF1AKMT1	552	621	530	749	557	676	543	666	553	577	525	382	475	571	551	474	643	543	676	566	629	637	466	532
5S_rRNA	468	528	423	3208	377	376	474	1098	346	769	669	1359	167	575	513	711	722	1516	1511	425	275	2245	1624	226
RNaseP_nuc	60	54	29	155	63	69	28	28	42	78	46	209	66	56	55	185	48	93	148	24	42	128	108	26
CDC42SE1	854	968	1342	1891	960	982	1163	1001	1152	1786	1915	717	791	1239	908	1789	1490	1089	933	1496	1328	876	1698	1124
SNORD14	16	38	21	13	28	26	39	41	16	31	4	0	24	16	27	23	46	13	39	39	44	36	13	23
CEBPG	79	130	150	169	116	178	165	122	160	186	185	151	112	185	129	215	159	150	129	183	148	141	198	152

Table S5. Non-coding RNA identified in liver tissue of Nelore bulls selected to be divergent for RFI in G2 applying the filter (25 mean count/samples).

ID	ERR669163	ERR669164	ERR669165	ERR669166	ERR669167	ERR669168	ERR669169	ERR669170	ERR669171	ERR669172	ERR669173	ERR669174	ERR669175	ERR669176	ERR669177	ERR669178	ERR669179	ERR669180	ERR669181	ERR669182
CEBPG	132	148	70	170	58	138	88	108	86	150	100	70	94	108	84	124	90	102	110	142
5_8S_rRNA	46	420	104	52	458	50	50	390	22	42	368	194	98	826	20	310	3338	28	80	138
CDC42SE1	378	624	296	480	374	494	370	426	404	368	574	370	424	544	368	568	396	490	490	390
EEF1AKM1	214	288	134	162	164	206	200	190	196	186	228	174	228	224	124	168	162	184	224	192
7SK	32	40	12	14	78	18	44	44	22	18	46	32	16	100	22	44	146	26	34	20

CHAPTER 3 - Gene co-expression network analysis and prediction of hub genes in liver of Nelore Cattle phenotypically divergent for RFI in two genetics groups

ABSTRACT – The network analysis is to identify complex transcriptional regulation and identification of central genes, being able to help in the selection of genetically superior animals for Residual Feed Intake (RFI) resulting in improving food efficiency could enhance productivity and simultaneously minimize environmental impact. The aim of this study was to use transcriptome RNA-seq data from Nelore cattle liver tissue to identify hub genes based on co-expression networks and network analysis with public databases using BioGRID. The normalized count data from 88 differentially expressed genes (DEGs) previously identified in both genetic groups associated with Residual Feed Intake (RFI) of liver in Nelore cattle were used for co-expression and interaction network analysis. In the first group (G1), the animals (n=24) were from a stabilized Nelore herd submitted to RFI evaluation: 12 low RFI (LRFI) and 12 high RFI (HRFI). The samples from their liver tissues were frozen at -80 °C for subsequent extraction of total RNA. The libraries (TruSeq RNA Library) were constructed according to the manufacturer's protocol and sequenced using HiSeq 2500 System (Illumina). For the second group (G2), 20 samples of liver tissue of Nelore cattle divergent for RFI were selected. The RNA-seq was performed using a HiSeq 2000 System (Illumina®). The raw data of G2 were chosen from the ENA repository (EMBL-EBI). To perform the biological enriched analyses, Pathway Studio (Elsevier) and DAVID software were used. The topological properties of the network were analyzed: those genes engaging in the most interactions with other genes were predicted to be the three top hub genes, BoLA-DRA, BoLA-DRB3 and CD74 which regulate the main significant biological processes found, such as immune response (GO:0006955) and MHC class II protein complex (GO:0042613). Up-regulated angiogenesis and down-regulated apoptosis processes were overlapped in both transcription factors, LRRFIP1 and NR4A2. In addition, seven candidate genes were highlighted using BioGRID network analysis: DDB2, ACTB, PAK1, VIM, PRC1, RPS6KA1 and PKM. These nodes have known effects in cell proliferation, energy metabolism, cytoskeletal organization and cell survival. The results reported in this study indicate candidate genes and molecular mechanisms involved in RFI in two genetic groups of Nelore cattle.

Keywords: bovine, co-expression, feed efficiency, hub genes, liver tissue, networks analysis, transcription factors.

1. Background

Gene expression plays the important role of converting information that is encoded in a gene into a functional product. Regulation of gene expression is of vital importance for all living species as it controls specific developmental stages and the response to prevailing environmental conditions (DE SOUZA et al., 2018). It also contributes to phenotypic diversity within and between species (WRAY et al., 2003). In the simplest scenario, the subset of gene transcripts in a tissue can be responsible for the differences between phenotypes. But Eukaryotic cellular functions are highly connected through networks of transcriptional regulators that regulate other transcriptional regulators (NEPH et al., 2012). However, the structure of core regulatory networks and their component sub-networks are largely undefined (SKINKYTE-JUSKIENE et al., 2018).

The primary regulators of gene expression are transcription factors (TFs), which are pivotal regulatory proteins that control gene expression in a context-dependent and tissue-specific manner (DE SOUZA et al., 2018; MITSIS et al., 2020). One method to understand gene function and gene–phenotype associations from genome-wide gene expression is to perform co-expression network analysis which is based on similarity in gene expression. The altered co-expression patterns of genes between two phenotypes represents significant potential to identify gene clusters associated with the phenotype of interest (GOV et al., 2017; BUSRA et al., 2019). They have become popular in recent years as large transcriptional data sets can be entered (SERIN et al., 2016).

The focus of network analysis is to identify complex transcriptional regulation and identification of central genes (hub genes, master regulators, nodes and key players). Hub genes are highly correlated with a great number of genes and could play an important role, acting as the main regulators in the co-expression network (FILTEAU et al., 2013; LIM et al., 2013; KOGELMAN et al., 2014) and can be used to build the interaction between genes and other cellular molecules or their involvement in pathways associated with Residual Feed Intake (RFI).

Residual feed intake is a methodology for assessing feed efficiency, regardless of metabolic live weight and daily weight gain, applied as a selection criterion for cattle with the aim of increasing the efficiency of individual feeding (BRANCO et al., 2009; GRION et al., 2014; CEACERO et al., 2016) without detriment to reproduction or deterioration of meat quality. The RFI trait is an

important, difficult and time-consuming economic measure of livestock (GERACI; SAHA; BIANCHINI, 2020).

The aim of this study was to use transcriptome RNA-seq data from Nelore cattle liver tissue to identify hub genes based on co-expression networks and network analysis with public databases using BioGRID using the 88 differentially expressed genes (DEGs) common in both genetic groups reported in Chapter 2. We hope to elucidate the most relevant biological functions, including hub genes and transcription factors which play an important role in liver metabolism associated with RFI in the biology of the Nelore.

2. Methods

Animals, RNA-Seq library preparation and differential analysis in two genetics groups were reported in Serna-García et al. (Chapter 2). The raw data of G2 were chosen from ENA repository (EMBL-EBI). The 88 common DEGs between LRFI and HRFI (Table 1) in both genetics groups (genetic group 1 , G1, genetic group 2 , G2) were used to perform the analysis as described below.

Table 1. 88 common DEGs in genetic group 1 and genetic group 2 for LRFI *vs* HRFI.

		Genetic group 1	**Genetic group 2**
Gene Name	**Gene Symbol**	**Fold change (log2)***	**Fold change (log2)***
Vascular cell adhesion molecule 1-like	*LOC534578*	0.8064	0.4097
Contactin 4	*CNTN4*	0.9359	0.6749
Myosin binding protein H	*MYBPH*	0.9146	0.6306
Antigen-presenting glycoprotein CD1d-like	*LOC788175*	0.5953	0.3008
Mannose receptor C type 2	*MRC2*	0.5687	0.3
Cytochrome b-245 beta chain	*CYBB*	0.5508	0.4176
Toll like receptor 5	*TLR5*	0.5309	0.8541
Joining chain of multimeric IgA and IgM	*JCHAIN*	0.5275	0.3372
Lymphocyte cytosolic protein 1	*LCP1*	0.5266	0.4347
Solute carrier family 16 member 9	*SLC16A9*	0.5142	0.5586
CD74 molecule	*CD74*	0.5054	0.2425
Cathepsin S	*CTSS*	0.4946	0.3211

Helicase, lymphoid-specific	*HELLS*	0.4887	0.4734
Major histocompatibility complex, class II, DR alpha	*BoLA-DRA*	0.4832	0.2185
Marginal zone B and B1 cell specific protein	*MZB1*	0.4825	0.57
Serine palmitoyltransferase small subunit B	*SPTSSB*	0.4723	0.5892
Major histocompatibility complex, class II, DR beta 3	*BoLA-DRB3*	0.4489	0.3182
IFI30, lysosomal thiol reductase	*IFI30*	0.4426	0.2357
Lumican	*LUM*	0.4358	0.3917
Integrin subunit beta 2	*ITGB2*	0.4358	0.2014
Stathmin 1	*STMN1*	0.4345	0.2428
Protein regulator of cytokinesis 1	*PRC1*	0.4199	0.1908
Nuclear receptor subfamily 4 group A member 2	*NR4A2*	0.4061	0.5458
Toll like receptor 2	*TLR2*	0.3937	0.3998
Pleckstrin homology domain containing B2	*PLEKHB2*	0.3896	0.1984
Moesin	*MSN*	0.3817	0.1346
Annexin A1	*ANXA1*	0.3743	0.2777
Pyruvate kinase, muscle	*PKM*	0.3742	0.1707
FGR proto-oncogene, Src family tyrosine kinase	*FGR*	0.3737	0.202
TNF alpha induced protein 8	*TNFAIP8*	0.3635	0.2624
FYN binding protein 1	*FYB1*	0.3626	0.192
Lymphocyte antigen 9	*LY9*	0.3472	0.3123
Lymphocyte cytosolic protein 2	*LCP2*	0.3456	0.2466
Rho GTPase activating protein 18	*ARHGAP18*	0.3437	0.2553
Amyloid beta precursor protein binding family B member 1 interacting protein	*APBB1IP*	0.3322	0.2228
LRR binding FLII interacting protein 1	*LRRFIP1*	0.327	0.2132
Myosin IF	*MYO1F*	0.3206	0.1629
P21 (RAC1) activated kinase 1	*PAK1*	0.318	0.2334
Parvin gamma	*PARVG*	0.3136	0.2878
Vimentin	*VIM*	0.3079	0.3593
Thymosin beta 4, X-linked	*TMSB4X*	0.3005	0.3199
Rho GTPase activating protein 30	*ARHGAP30*	0.299	0.1761

NPC intracellular cholesterol transporter 2	*NPC2*	0.2977	0.3
Ecotropic viral integration site 2B	*EVI2B*	0.2939	0.1647
Family with sequence similarity 45 member A	*FAM45A*	0.2852	0.1354
Uncharacterized	*LOC783680*	0.2802	0.1726
Adenylosuccinate synthase	*ADSS*	0.276	0.1704
Coactosin like F-actin binding protein 1	*COTL1*	0.2718	0.1888
Actin beta	*ACTB*	0.2713	0.1693
Syntaxin 7	*STX7*	0.2613	0.1475
Beta-2-microglobulin	*B2M*	0.2578	0.2057
Ribosomal protein S6 kinase A1	*RPS6KA1*	0.252	0.252
Calpain 2	*CAPN2*	0.2296	0.1569
Death associated protein kinase 1	*DAPK1*	0.2284	0.2626
Sorting nexin 2	*SNX2*	0.214	0.1328
Methionine adenosyltransferase 2B	*MAT2B*	0.2139	0.1131
Granulin precursor	*GRN*	0.1808	0.1586
Integral membrane protein 2B	*ITM2B*	0.1688	0.1242
Sorting nexin 6	*SNX6*	0.1672	0.1112
Transmembrane protein 50A	*TMEM50A*	0.1392	0.1724
Damage specific DNA binding protein 2	*DDB2*	0.1129	0.2711
CD63 molecule	*CD63*	0.0771	0.0907
LDL receptor related protein 5	*LRP5*	-0.1151	-0.0944
Inositol polyphosphate-5-phosphatase K	*INPP5K*	-0.1223	-0.1309
Kelch like family member 36	*KLHL36*	-0.1304	-0.1356
ATP binding cassette subfamily F member 3	*ABCF3*	-0.1319	-0.1681
GDNF inducible zinc finger protein 1	*GZF1*	-0.1394	-0.2676
Tetratricopeptide repeat domain 19	*TTC19*	-0.1468	-0.1329
TATA-box binding protein associated factor 5 like	*TAF5L*	-0.1487	-0.1746
Ankyrin repeat and SOCS box containing 8	*ASB8*	-0.1495	-0.1683
BTB domain containing 6	*BTBD6*	-0.169	-0.0768
LETM1 domain containing 1	*LETMD1*	-0.1713	-0.1957

Adhesion G protein-coupled receptor A3	*ADGRA3*	-0.1803	-0.1609
ARV1 homolog, fatty acid homeostasis modulator	*ARV1*	-0.1859	-0.175
Transmembrane and coiled-coil domain family 1	*TMCC1*	-0.2	-0.2596
Adaptor related protein complex 4 beta 1 subunit	*AP4B1*	-0.2135	-0.2169
Zinc finger CCHC-type containing 7	*ZCCHC7*	-0.2154	-0.2452
Mitochondrial trans-2-enoyl-CoA reductase	*MECR*	-0.2213	-0.2052
SET domain and mariner transposase fusion gene	*SETMAR*	-0.2426	-0.2787
Tubulin alpha 4a	*TUBA4A*	-0.2453	-0.2353
Rho guanine nucleotide exchange factor 5	*ARHGEF5*	-0.2641	-0.2617
Adrenoceptor alpha 1A	*ADRA1A*	-0.2719	-0.2767
Zinc finger protein 665	*LOC515089*	-0.2801	-0.3338
Zinc finger protein 286A	*ZNF286A*	-0.2862	-0.1821
Cyclin B3	*CCNB3*	-0.3393	-0.437
F-box protein 15	*FBXO15*	-0.4387	-0.4727
Thymus, brain and testes associated	*TBATA*	-0.4652	-0.5368
Uncharacterized	*LOC112446800*	-0.5062	-0.5978

*The Fold-change estimates (relative expression) refer to the LRFI group.

2.1. Gene co-expression network analysis and prediction of hub genes

Co-expression analysis was performed on 88 DEGs in LRFI groups and compared with HRFI animals (Table 1) using plugins and applications from Cytoscape v.3.4 (SHANNON et al., 2003) which is a free software package for visualizing, modeling and analyzing the integration of biomolecular interaction networks with high-throughput expression data and other molecular states.

The co-expression network was constructed using the ExpressionCorrelation plugin (HUI et al., 2015). The similarity matrix of count data from DEGs (normalized by CPMs) was computed using Pearson's correlation. A histogram tool was used for the screening criteria at node score cut-offs > 0.75 and < -0.75 and employed to identify statistical significance of the pairwise correlations. Using the Network Analyzer plugin, regulatory relationships among various genes were

identified by topological properties of the network, such as degree distribution, betweenness centrality, network density and clustering coefficient (BARABÁSI; OLTVAI, 2004; DONG; HORVATH, 2007).

The cytoHubba plugin (CHIN et al., 2014) was used to explore co-expression network nodes (hub genes). This plugin provides a user-friendly interface to explore important nodes in biological networks and computes using eleven methods (degree, edge percolated component, maximum neighborhood component, density of maximum neighborhood component, maximal clique centrality -MCC and six centralities - bottleneck, EcCentricity, closeness, radiality, betweenness and stress). All these parameters can indicate the robustness of the analysis (STEVENS et al., 2012; CHIN et al., 2014). The MCC method performs better on the precision of predicting essential genes from the co-expression network (CHIN et al., 2014) and generating a sub-network (STEVENS et al., 2013). The top three hub genes were chosen to define top modules screened from the gene co-expression network. The top modules could be functionally relevant and therefore are used to assess biological function (STEVENS et al., 2013).

2.2. Analysis of the nodes of physical interactions with BioGRID

An analysis of the physical interactions between the 88 common DEGs between the two genetic groups with public data from the General Biological Repository for Interaction DataSets (BioGRID 4.4.204 , https://downloads.theBioGRID.org/BioGRID/Release-Archive/BIOGRID-4.4.205/) was performed (OUGHTRED et al., 2021). In-house scripts were used for annotation and interaction map building (CytoScape format).

BioGRID is a biomedical interaction repository with data compiled through comprehensive curation efforts. All BioGRID content is curated from primary experimental evidence in the biomedical literature and includes both focused low-throughput studies and large high-throughput datasets (STARK et al., 2006).

2.3. Gene set enrichment analysis

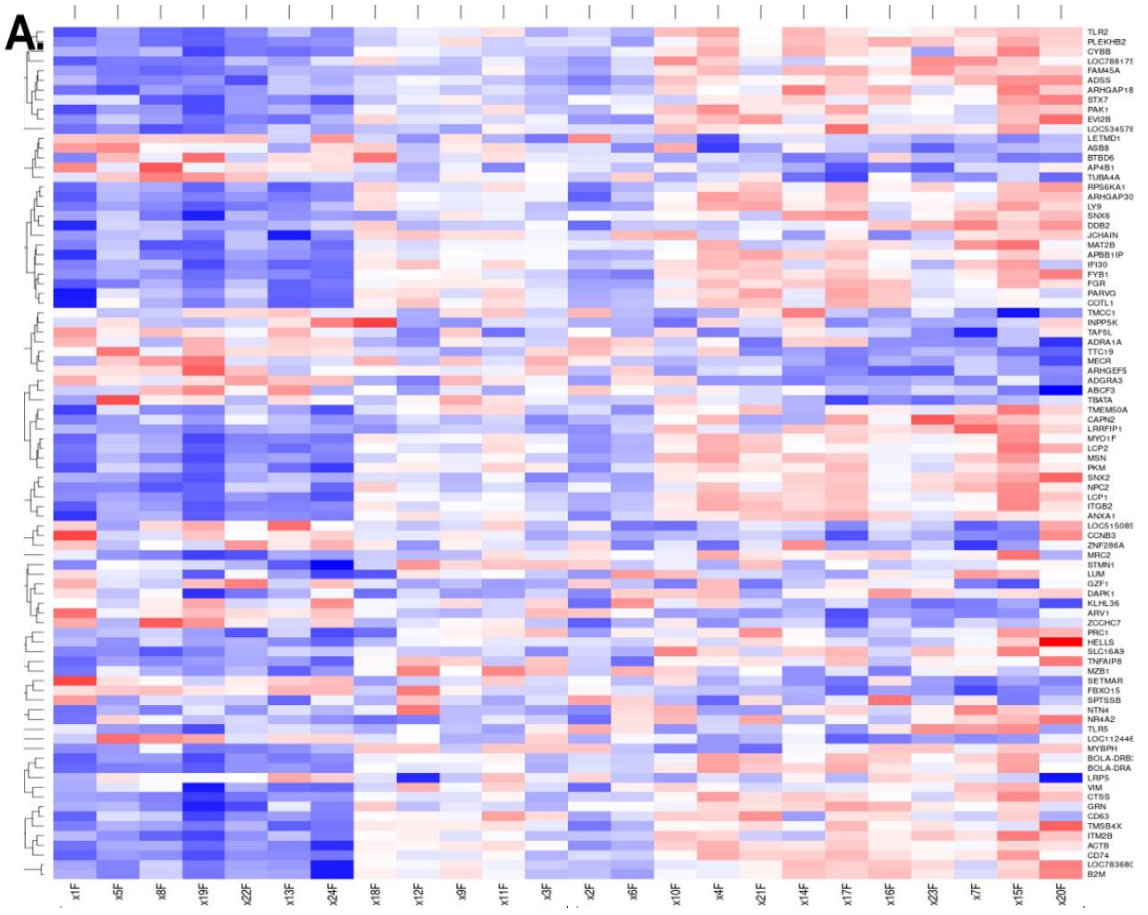
To common nodes obtained in the previously analysis were submitted to enrichment analysis using the Pathway Studio version 10 software (Elsevier Inc., Rockville, MD, USA) with database ResNet v11 and Database for Annotation,

Visualization and Integrated Discovery (DAVID 6.8) (HUANG; SHERMAN; LEMPICKI, 2009). The DAVID Pathway was used to map the enriched pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (KANEHISA; GOTO, 2000). False discovery rates (FDRs) were controlled by the Benjamini-Hochberg method (BENJAMINI; HOCHBERG, 1995) considering a p-value of less than 5%.

3. Results

3.1. Gene co-expression network analysis and prediction of hub genes

Based on differential expression levels of transcripts between LRFI and HRFI animals, 88 DEGs, common to both genetic groups, were identified (Table 1). The expression levels of transcripts changed dynamically between LRFI and HRFI animals in G1 and G2 as shown on the heat map (Figure 1).



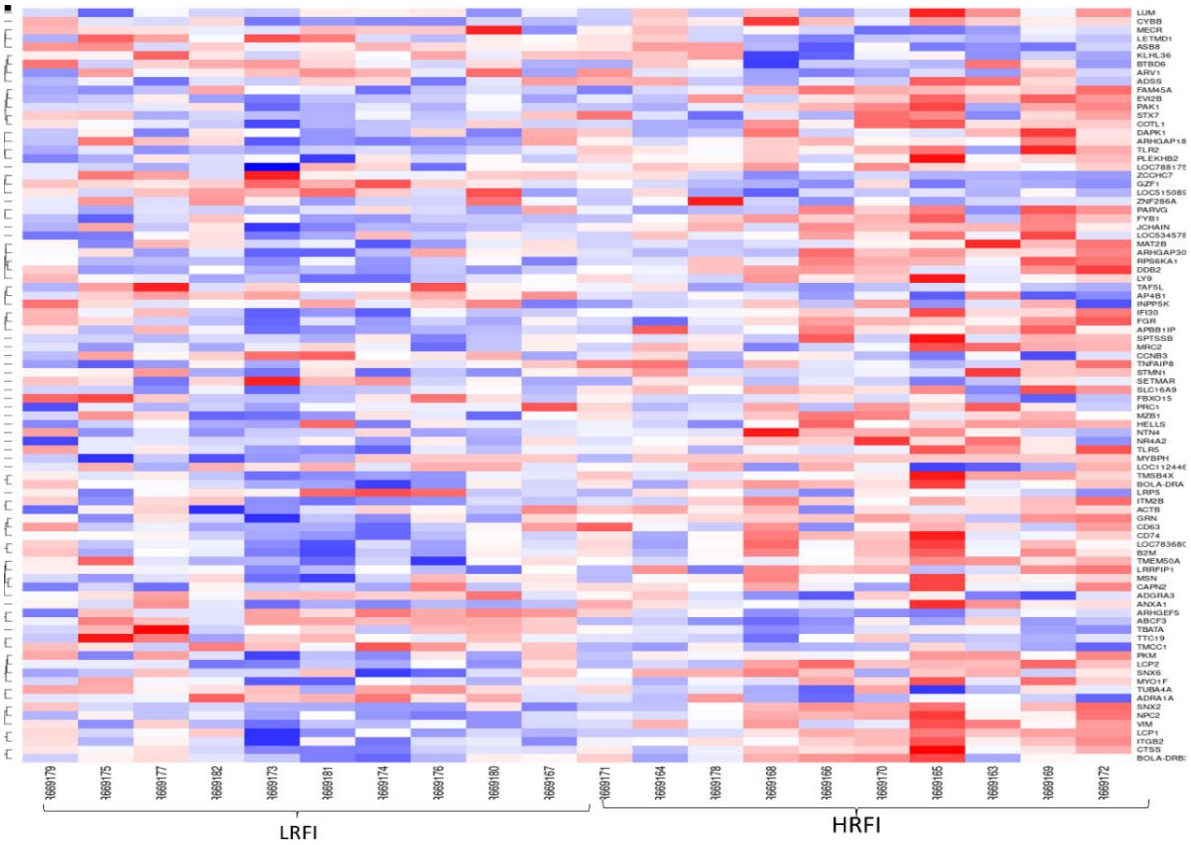


Figure 1. The heat map displays differentially expressed genes found in Nelore cattle divergent for RFI. Blue (under-expressed in LRFI), red (overexpressed in LRFI) and color intensity represent differing levels of expression of the genes. **A.** Heat map for G1. **B.** Heat map for G2.

We used the correlation matrix to construct the co-expression network (Figure 2) from these data. The co-expression network in G1 had 624 gene pairs (edges); 59 genes (nodes); 0.182 density; and 0.342 clustering coefficient (Figure 2A). The co-expression network in G2 had 111 gene pairs (edges); 47 genes (nodes); 0.051 density; and 0.160 clustering coefficient (Figure 2B).

When the network MCC was examined, the top three genes with the highest MCC scores were considered as the hub genes: *Bos taurus* major histocompatibility complex class II DR alpha (*BoLA-DRA*), *Bos taurus* major histocompatibility complex class II DR beta 3 (*BoLA-DRB3*) and *CD74* molecule major histocompatibility complex class II invariant chain (*CD74*) (Figure 2).

We found 36 overlapped genes (Figure 2 and Table 2). The common nodes were ~61% for G1 and ~77% for G2.

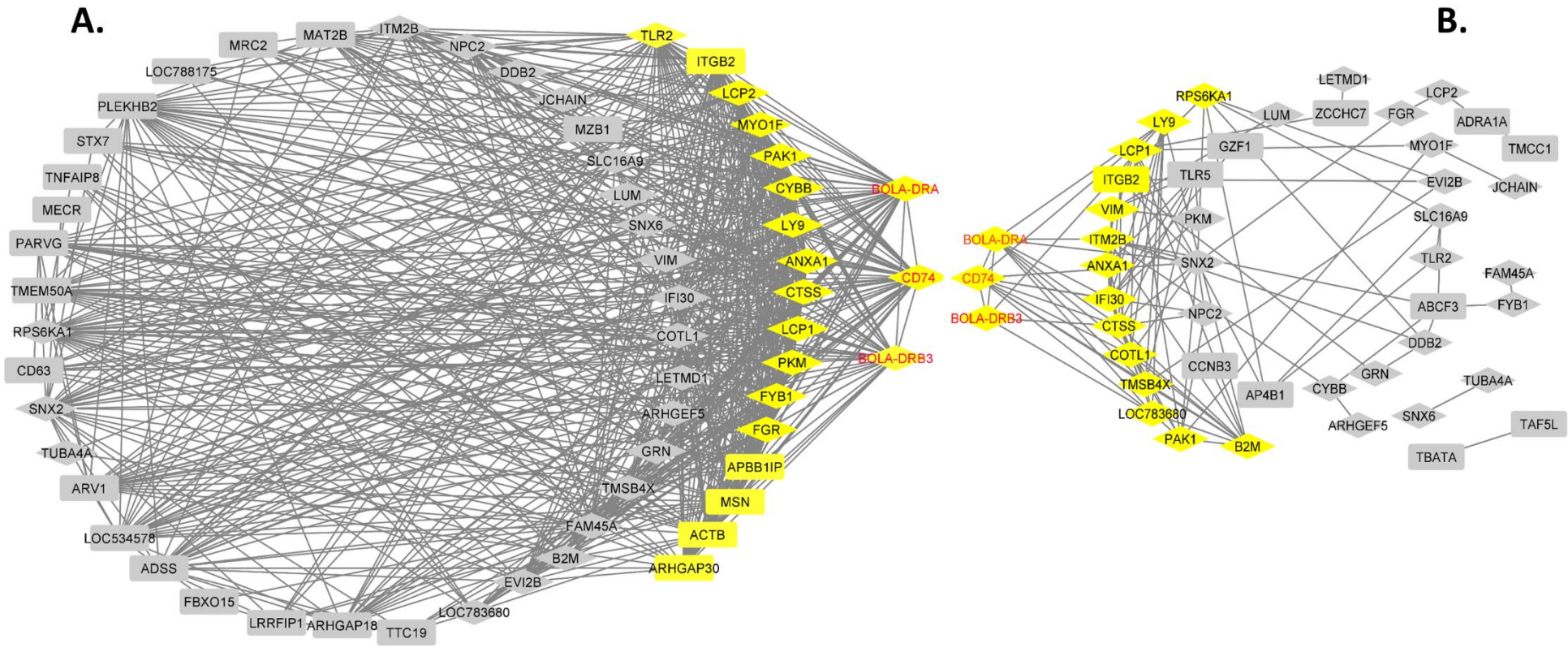


Figure 2. Gene co-expression network analysis for the differentially expressed genes in liver of Nelore cattle divergent for RFI. Genes colored in yellow are the top 20 in network ranked by the MCC method for each genetic group. Genes represented in diamonds were common between both genetic groups, genes highlighted in red letters were predicted as hubs. **A.** Gene co-expression network analysis for G1. **B.** Gene co-expression network analysis for G2.

Table 2. Thirty-six (nodes) overlapped genes in G1 and G2 from co-expression network analysis for the differentially expressed genes in liver of Nelore cattle divergent for RFI.

		Genetic group 1		Genetic group 2	
Nodes	Node Symbol	Fold change (log2)*	p-value	Fold change (log2)*	p-value
Cytochrome b-245 beta chain	*CYBB*	0.5508	0.0338	0.4176	0.0248
Joining chain of multimeric IgA and IgM	*JCHAIN*	0.5275	0.0124	0.3372	0.0146
Lymphocyte cytosolic protein 1	*LCP1*	0.5266	0.0145	0.1908	0.0230
Solute carrier family 16 member 9	*SLC16A9*	0.5142	0.0431	0.5586	0.0058
CD74 molecule	*CD74*	0.5054	0.0158	0.2425	0.0213
Cathepsin S	*CTSS*	0.4946	0.0410	0.3211	0.0209
Major histocompatibility complex, class II, DR alpha	*BoLA-DRA (HLA-DRA)*	0.4832	0.0157	0.2185	0.0499
Major histocompatibility complex, class II, DR beta 3	*BoLA-DRB3*	0.4489	0.0098	0.3182	0.0262
IFI30, lysosomal thiol reductase	*IFI30*	0.4426	0.0206	0.2357	0.0311
Integrin subunit beta 2	*ITGB2*	0.4358	0.0359	0.2014	0.0401
Lumican	*LUM*	0.4358	0.0263	0.3917	0.0390
Toll like receptor 2	*TLR2*	0.3937	0.0337	0.3998	0.0008
Annexin A1	*ANXA1*	0.3743	0.0136	0.2777	0.0171
Pyruvate kinase, muscle	*PKM*	0.3742	0.0274	0.1707	0.0364
FGR proto-oncogene, Src family tyrosine kinase	*FGR*	0.3737	0.0462	0.2020	0.0284
FYN binding protein 1	*FYB1*	0.3626	0.0266	0.1920	0.0240
Lymphocyte antigen 9	*LY9*	0.3472	0.0409	0.3123	0.0053
Lymphocyte cytosolic protein 2	*LCP2*	0.3456	0.0302	0.2466	0.0003
Myosin IF	*MYO1F*	0.3206	0.0366	0.1629	0.0369
P21 (RAC1) activated kinase 1	*PAK1*	0.3180	0.0344	0.2334	0.0356
Vimentin	*VIM*	0.3079	0.0291	0.3593	0.0085
Thymosin beta 4, X-linked	*TMSB4X*	0.3005	0.0350	0.3199	0.0032
Ecotropic viral integration site 2B	*EVI2B*	0.2939	0.0427	0.1647	0.0454
Family with sequence similarity 45 member A	*FAM45A*	0.2852	0.0223	0.1354	0.0270
Uncharacterized	*LOC783680*	0.2802	0.0297	0.1726	0.0247
Coactosin like F-actin binding protein 1	*COTL1*	0.2718	0.0321	0.1888	0.0162
Beta-2-microglobulin	*B2M*	0.2578	0.0460	0.2057	0.0115
Ribosomal protein S6 kinase A1	*RPS6KA1*	0.2520	0.0250	0.2520	0.0250
Sorting nexin 2	*SNX2*	0.2140	0.0231	0.1328	0.0299
Granulin precursor	*GRN*	0.1808	0.0271	0.1586	0.0036
Integral membrane protein 2B	*ITM2B*	0.1688	0.0025	0.1242	0.0308
Sorting nexin 6	*SNX6*	0.1672	0.0234	0.1112	0.0456
Damage specific DNA binding protein 2	*DDB2*	0.1129	0.0264	0.2711	0.0081
LETM1 domain containing 1	*LETMD1*	-0.1713	0.0335	-0.1957	0.0442
Tubulin alpha 4a	*TUBA4A*	-0.2453	0.0461	-0.2353	0.0178

Rho guanine nucleotide exchange factor 5	*ARHGEF5*	-0.2641	0.0240	-0.2617	0.0400

*The Fold-change estimates (relative expression) refer to the LRFI group.

3.1.1. Functional Analysis of 36 common nodes

Significantly enriched gene ontology (GO) terms in the biological processes are shown in Table 3. The most significative biological process of the 36 nodes was innate immune response (GO:0045087).

The three predicted genes as hubs: *BoLA-DRA, BoLA-DRB3* and *CD74* overexpressed in LRFI (Tables 1 and 2 and Figure 2) are included in immune response (GO:0006955) and MHC class II protein complex (GO:0042613). In addition, we found important biological processes related to the immune system for three predicted genes as hubs.

Table 3. Biological Process GO terms and pathways obtained with the DAVID software for 36 hub genes. DEGs in Nelore cattle in liver divergent for RFI phenotype.

GO Term	Biological processe	Number of genes [1]	Genes (nodes)	p-value enricnhect	FDR[2]
GO:0045087	innate immune response	7	*FGR, ANXA1, CYBB, LOC783680, B2M, JCHAIN, TLR2*	6.35E-06	0.001
GO:0001948	glycoprotein binding	4	*ITGB2, LOC783680, VIM, B2M*	1.42E-04	0.009
GO:0050830	defense response to Gram-positive bacterium	4	*FGR, LOC783680, B2M, TLR2*	2.89E-04	0.031
GO:0005925	focal adhesion	6	*PAK1, ANXA1, LOC783680, LCP1, VIM, B2M*	4.55E-04	0.019
GO:0001895	retina homeostasis	3	*LOC783680, B2M, JCHAIN*	5.01E-04	0.036
GO:0042613	MHC class II protein complex	3	*CD74, BOLA-DRB3, BOLA-DRA*	6.94E-04	0.020
GO:0019731	antibacterial humoral response	3	*LOC783680, B2M, JCHAIN*	8.71E-04	0.041
GO:0006955	immune response	5	*BOLA-DRB3, LCP2, BOLA-DRA, B2M, CTSS*	9.96E-04	0.041
GO:0010977	negative regulation of neuron projection development	3	*LOC783680, VIM, B2M*	0.00114	0.041

[1]Number of genes (nodes) related to GO term; [2]FDR: False Discovery Rate <0.05

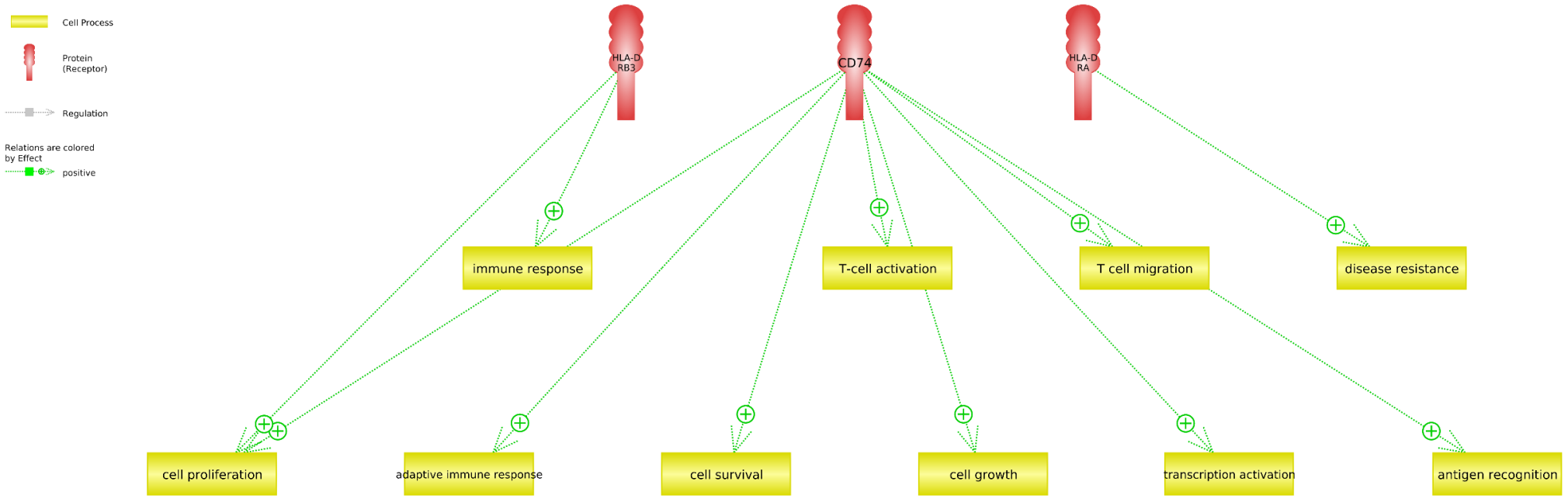


Figure 3. Cell processes associated with three genes predicted as hubs using Pathway Studio. Highlighted in red are genes overexpressed in LRFI group. Green arrows indicate a positive regulation effect.

3.2. Biological processes obtained from a transcription factors study

Within our list of the 88 common DEGs of both genetic groups (Table 1) we found two transcription factors overexpressed genes in the more efficient group using Pathway Studio software: LRR binding FLII interacting protein 1 (*LRRFIP1*) and nuclear receptor subfamily 4 group 2 member (*NR4A2*) (Figure 4). Up-regulated angiogenesis and down-regulated apoptosis were overlapped in both genes, *LRRFIP1* and *NR4A2.* Vital cellular mechanisms were observed, such as inhibition of adipogenesis and adipocyte differentiation, inhibition of inflammatory response, inhibition of cell loss, inhibition of cytotoxicity, activation of cell survival, activation of fatty acid oxidation, activation of macrophage differentiation, activation of b-cell formation, activation of cell maturation, activation of cell differentiation, activation of cell development, activation of regeneration, activation of cell cycle, activation of stem cell differentiation, activation of mitochondrial gene expression, activation of glucose import, activation of cell proliferation, activation of muscle cell differentiation, activation of autophagy and inhibition of hepatic stellate cell activation.

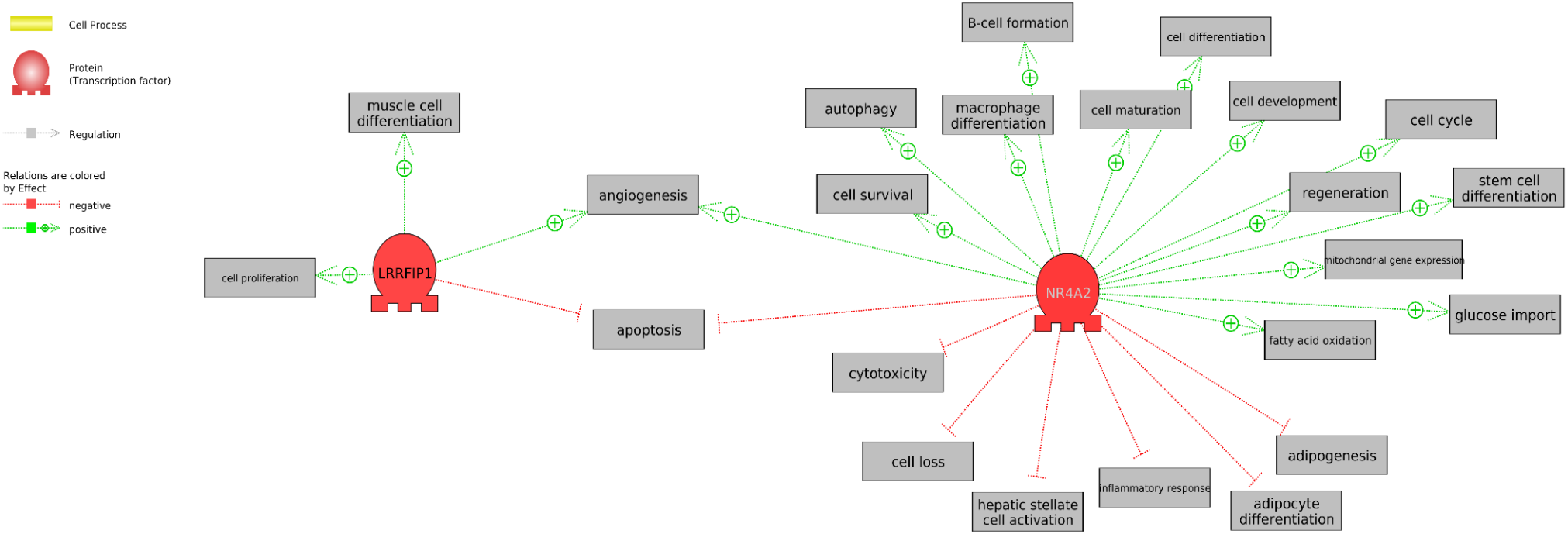


Figure 4. Cell processes associated with two transcription factors using Pathway Studio. Highlighted in red are overexpressed genes in LRFI a group. Green arrows indicate a positive regulation effect, while red arrows has a negative regulation effect.

3.3. Analysis of the nodes of physical interactions with BioGRID

Figures 5 and 6 show the physical interactions between the 88 DEGs (Table 1) with proteins from the General Biological Repository for Interaction DataSets (BioGRID). Fifty-four nodes were found that interact with 358 potential genes (targets) (Figures 5 and 6 and Table 2).

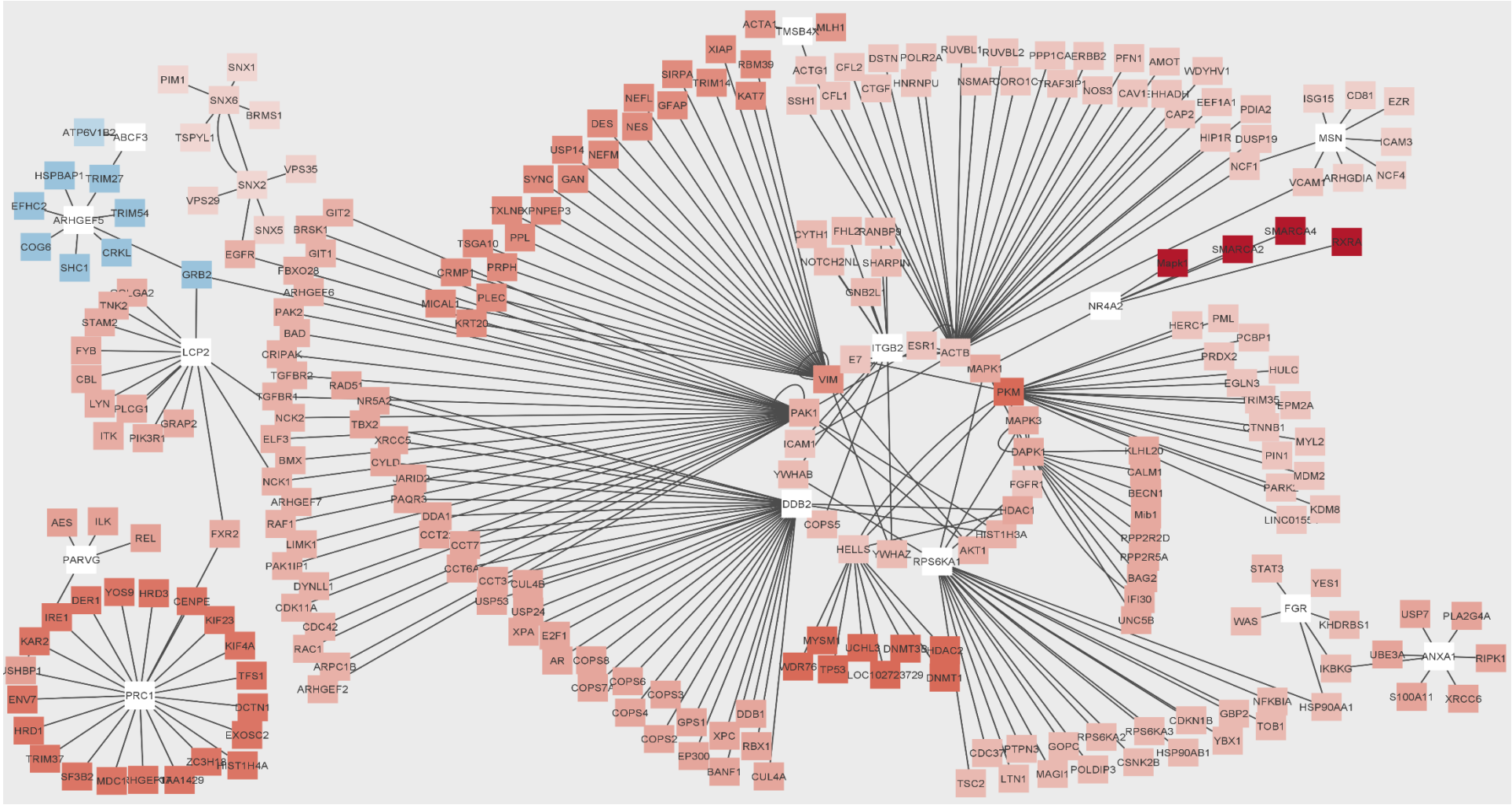


Figure 5. Interaction network for the DEGs in liver of Nelore cattle divergent for RFI by BioGRID. Blue (underexpressed in LRFI), red (overexpressed in LRFI) and color intensity represents differing levels of expression of the genes. White represents nodes upstream.

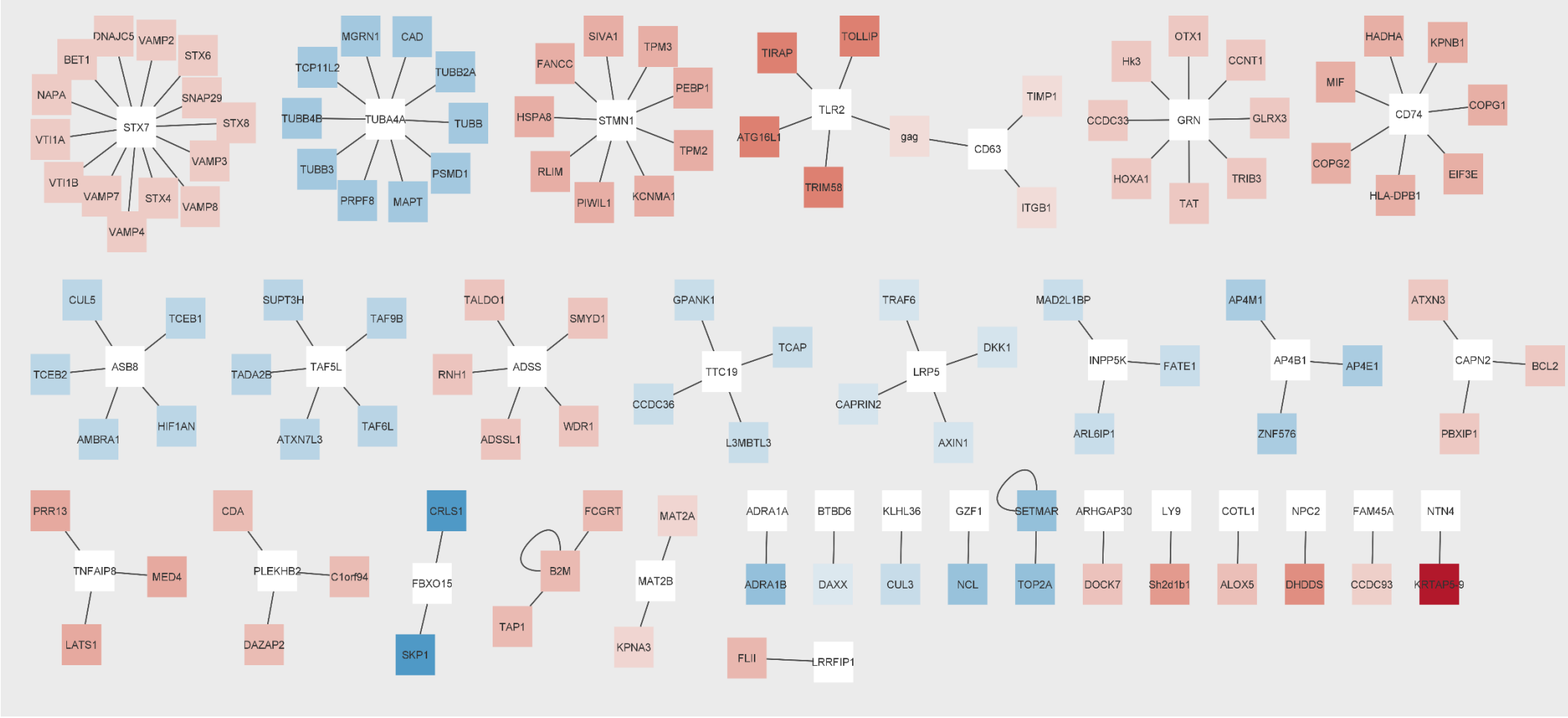


Figure 6. Interaction network for the DEGs in liver of Nelore divergent for RFI by BioGRID. The blue color (underexpressed in LRFI), red color (overexpressed in LRFI) and color intensity represent differing levels of expression of the genes. White color represents nodes upstream.

Table 2. 54 nodes (between the 88 DEGs) with their 358 potential genes through network analysis.

Nodes	Node symbol	Potential Genes (potential targets)	Number of genes
ATP binding cassette subfamily F member 3	*ABCF3*	*ATP6V1B2, TRIM27*	2
Actin beta	*ACTB*	*CFL1, CFL2, ACTB, DSTN, ACTG1, NCF1, SSH1, ESR1, CTGF, POLR2A, HNRNPU, NSMAF, RUVBL1, RUVBL2, CORO1C, PPP1CA, TRAF3IP1, ERBB2, NOS3, ICAM1, VCAM1, PFN1, CAV1, AMOT, EHHADH, CAP2, WDYHV1, EEF1A1, HIP1R, PDIA2, DUSP19*	31
Adrenoceptor alpha 1A	*ADRA1A*	*ADRA1B*	1
Adenylosuccinate synthase	*ADSS*	*RNH1, SMYD1, TALDO1, WDR1, ADSSL1*	5
Annexin A1	*ANXA1*	*S100A11, UBE3A, XRCC6, IKBKG, RIPK1, USP7, PLA2G4A*	7
Adaptor related protein complex 4 beta 1 subunit	*AP4B1*	*AP4E1, AP4M1, ZNF576*	3
Rho GTPase activating protein 30	*ARHGAP30*	*DOCK7*	1
Rho guanine nucleotide exchange factor 5	*ARHGEF5*	*CRKL, GRB2, TRIM54, HSPBAP1, EFHC2, SHC1, COG6, TRIM27*	8
Ankyrin repeat and SOCS box containing 8	*ASB8*	*TCEB2, TCEB1, CUL5, HIF1AN, AMBRA1*	5
Beta-2-microglobulin	*B2M*	*TAP1, B2M, FCGRT*	3
BTB domain containing 6	*BTBD6*	*DAXX*	1
Calpain 2	*CAPN2*	*BCL2, ATXN3, PBXIP1*	3
CD63 molecule	*CD63*	*ITGB1, gag, TIMP1*	3
CD74 molecule	*CD74*	*MIF, HADHA, COPG1, COPG2, HLA-DPB1, EIF3E, KPNB1*	7
Coactosin like F-actin binding protein 1	*COTL1*	*ALOX5*	1
Death associated protein kinase 1	*DAPK1*	*KLHL20, DAPK1, CALM1, MAPK3, MAPK1, BECN1, Mib1, PPP2R2D, PPP2R5A, BAG2, IFI30, UNC5B*	12
Damage specific DNA binding protein 2	*DDB2*	*DDB1, CUL4B, CUL4A, RBX1, BANF1, EP300, XPC, GPS1, COPS2, COPS3, COPS4, COPS5, COPS6, COPS7A, COPS8, AR, E2F1, XPA, USP24, USP53, CCT3, CCT6A, HDAC1, CCT7, CCT2, DDA1, PAQR3, JARID2, CYLD, XRCC5, TBX2, HIST1H3A, NR5A2, RAD51*	35
Family with sequence similarity 45 member A	*FAM45A*	*CCDC93*	1
F-box protein 15	*FBXO15*	*CRLS1, SKP1*	2
FGR proto-oncogene, Src family tyrosine kinase	*FGR*	*WAS, KHDRBS1, HSP90AA1, STAT3, IKBKG, YES1*	6
Granulin precursor	*GRN*	*HOXA1, CCNT1, TAT, TRIB3, Hk3, OTX1, GLRX3, CCDC33*	7
GDNF inducible zinc finger protein 1	*GZF1*	*NCL*	1
Helicase, lymphoid-specific	*HELLS*	*HDAC1, HDAC2, DNMT1, DNMT3B, LOC102723729, UCHL3, TP53, PKM, MYSM1, WDR76*	10
IFI30, lysosomal thiol reductase	*IFI30*	*DAPK1*	1
Inositol polyphosphate-5-phosphatase K	*INPP5K*	*FATE1, MAD2L1BP, ARL6IP1*	3
Integrin subunit beta 2	*ITGB2*	*GNB2L1, ICAM1, CYTH1, YWHAB, COPS5, YWHAZ , RANBP9,NOTCH2NL, SHARPIN,FHL2*	10
Kelch like family member 36	*KLHL36*	*CUL3*	1
Lymphocyte cytosolic protein 2	*LCP2*	*GRAP2,PIK3R1,ITK,NCK1,PLCG1,LYN,CBL,GRB2,FYB,NCK2,FXR2,STAM2,TNK2,GOLGA2*	14
LDL receptor related protein 5	*LRP5*	*AXIN1, CAPRIN2, DKK1, TRAF6*	4
LRR binding FLII interacting protein 1	*LRRFIP1*	*FLII*	1

Lymphocyte antigen 9	*LY9*	*Sh2d1b1*	1
Methionine adenosyltransferase 2B	*MAT2B*	*MAT2A, KPNA3*	2
Moesin	*MSN*	*EZR, NCF1, NCF4, ISG15, CD81, VCAM1, ICAM3, ARHGDIA*	8
NPC intracellular cholesterol transporter 2	*NPC2*	*DHDDS*	1
Nuclear receptor subfamily 4 group A member 2	*NR4A2*	*SMARCA4, SMARCA2, RXRA, Mapk1, MAPK1*	5
Contactin 4	*CNTN4*	*KRTAP5-9*	2
P21 (RAC1) activated kinase 1	*PAK1*	*ARHGEF2, ARPC1B, RAC1, CDC42, CDK11A, DYNLL1, PAK1IP1, LIMK1, RAF1, ARHGEF7, NCK1, BMX, PAK1, ELF3, NCK2, TGFBR1, TGFBR2, CRIPAK, ESR1, BAD, HIST1H3A, PAK2, ARHGEF6, FBXO28, GIT1, AKT1, BRSK1, GIT2, GRB2, EGFR*	30
	PARVG	*USHBP1, ILK, AES, REL*	4
Pyruvate kinase, muscle	*PKM*	*HERC1, PML, PCBP1, PRDX2, EGLN3, TRIM35, FGFR1, EPM2A, CTNNB1, MYL2, MAPK1, PIN1, MDM2, PARK2, KDM8, E7, LINC01554, HELLS, HULC*	19
Pleckstrin homology domain containing B2	*PLEKHB2*	*C1orf94, CDA, DAZAP2*	3
Protein regulator of cytokinesis 1	*PRC1*	*USHBP1, TFS1, KIF4A, KIF23, CENPE, HRD3, YOS9, DER1, IRE1, KAR2, ENV7, HRD1, TRIM37, SF3B2, HIST1H4A, MDC1, FXR2, ARHGEF17, KIAA1429, ZC3H18, DCTN1, EXOSC2*	22
Ribosomal protein S6 kinase A1	*RPS6KA1*	*MAPK3, MAPK1, YWHAB, TSC2, NFKBIA, TOB1, GBP2, YBX1, CDKN1B, HSP90AA1, HSP90AB1, RPS6KA3, CSNK2B, RPS6KA2, POLDIP3, FGFR1, GOPC, MAGI1, PTPN3, LTN1, CDC37*	21
SET domain and mariner transposase fusion gene	*SETMAR*	*SETMAR, TOP2A*	2
Sorting nexin 2	*SNX2*	*EGFR, SNX6, VPS29, VPS35, SNX5*	5
Sorting nexin 6	*SNX6*	*PIM1, SNX2, SNX1, BRMS1, TSPYL*	5
Stathmin 1	*STMN1*	*HSPA8, KCNMA1, TPM2, RLIM, SIVA1, TPM3, PIWIL1, FANCC, PEBP1*	7
Syntaxin 7	*STX7*	*VAMP8, VTI1A, STX8, VTI1B, SNAP29, STX6, NAPA, STX4, VAMP4, VAMP3, VAMP2, VAMP7, BET1, DNAJC5*	14
TATA-box binding protein associated factor 5 like	*TAF5L*	*TADA2B, TAF9B, SUPT3H, TAF6L, ATXN7L3*	5
Toll like receptor 2	*TLR2*	*TOLLIP, TIRAP, gag, ATG16L1, TRIM58*	5
Thymosin beta 4, X-linked	*TMSB4X*	*ACTA1, ACTG1, MLH1*	3
TNF alpha induced protein 8	*TNFAIP8*	*MED4, PRR13, LATS1*	3
Tetratricopeptide repeat domain 19	*TTC19*	*L3MBTL3, CCDC36, TCAP, GPANK1*	4
Tubulin alpha 4a	*TUBA4A*	*TUBB, TUBB2A, TUBB3, TUBB4B, MGRN1, MAPT, CAD, PRPF8, PSMD1, TCP11L2*	10
Vimentin	*VIM*	*KAT7, KRT20, VIM, YWHAZ, MICAL1, PLEC, CRMP1, PRPH, AKT1, TSGA10, PPL, TXLNB, XPNPEP3, SYNC, GAN, USP14, NEFM, DES, NES, NEFL, GFAP, SIRPA, TRIM14, XIAP, RBM39*	25

We found seven nodes that interact with more than 15 genes: adenylosuccinate synthase (*DDB2*), actin beta (*ACTB*) activated kinase 1 (*PAK1*), vimentin (*VIM*), protein regulator of cytokinesis 1 (*PRC1*), ribosomal protein S6 kinase A1 (*RPS6KA1*) and pyruvate kinase, muscle (*PKM*) (Table 2 and Figure 7). All of them were overexpressed in LRFI animals (Figure 8, see letters in red).

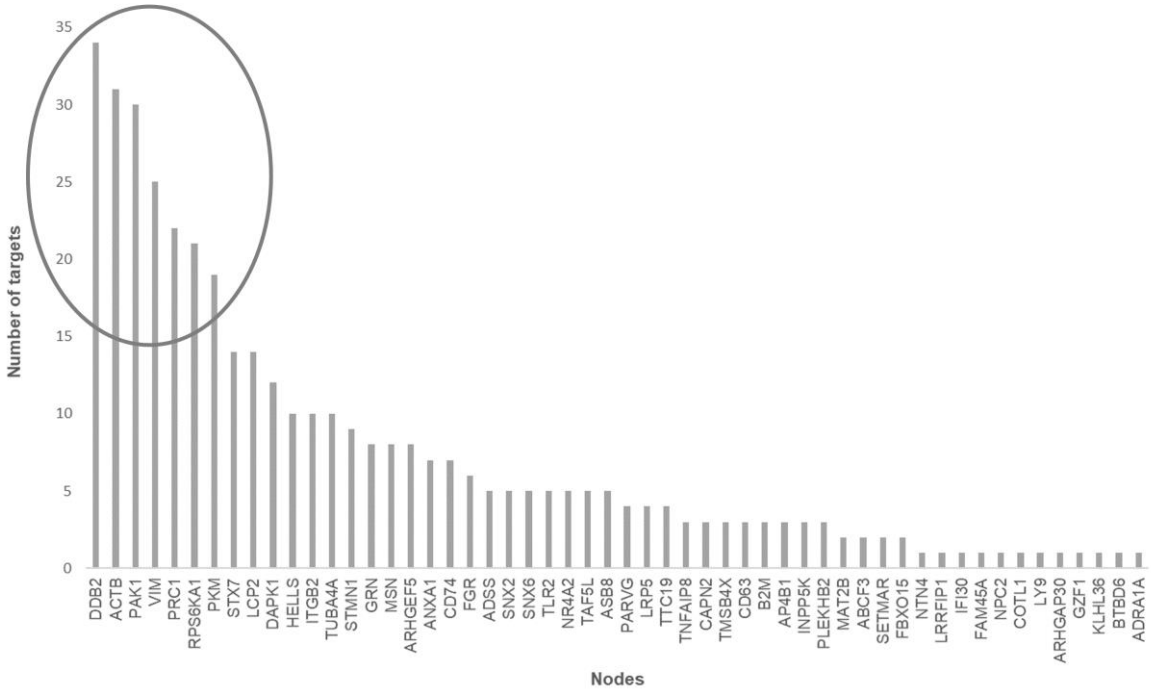


Figure 7. The 54 nodes that interact with genes (targets) using BioGRID. The circle highlights the main nodes that interact with more than 15 potential genes.

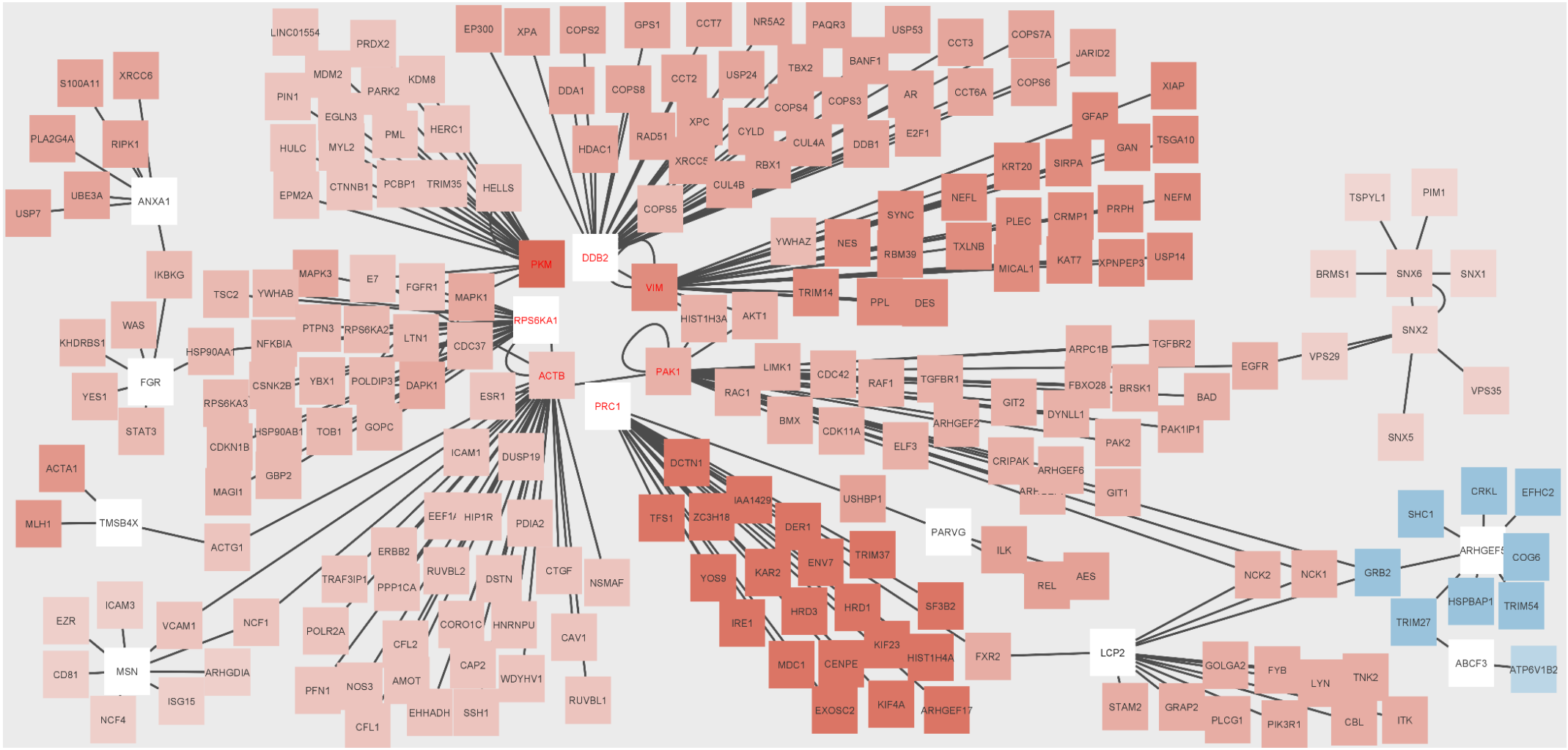


Figure 8. Interaction network for the DEGs in liver of Nelore divergent for RFI by BioGRID. The blue color (underexpressed in LRFI), red color (overexpressed in LRFI) and color intensity represent differing levels of expression of the genes. In white color are represents the upstream nodes and 7 hub genes highlighted with red letters.

3.3.1. Functional Analysis of 54 nodes obtained from networks analysis using BioGRID

We have obtained 101 significant cellular processes (see Additional File 1). Fifteen of the most relevant cellular processes, taking into consideration the number of nodes (genes) regulating these processes, such as T-cell activation (p-value $4.20x10^{-05}$), cytoskeleton organization (p-value $2.56x10^{-03}$), glucose import (p-value $2.82x10^{-03}$) and immune system function (p-value $3.76x10^{-03}$), are presented in Figure 9.

Several nodes are involved in cellular processes of interest for RFI (Figure 10), such as immune system function, cell proliferation, cell survival, angiogenesis, cytoskeleton organization and glycolysis, among others. The main nodes (circled in black) regulate cell proliferation.

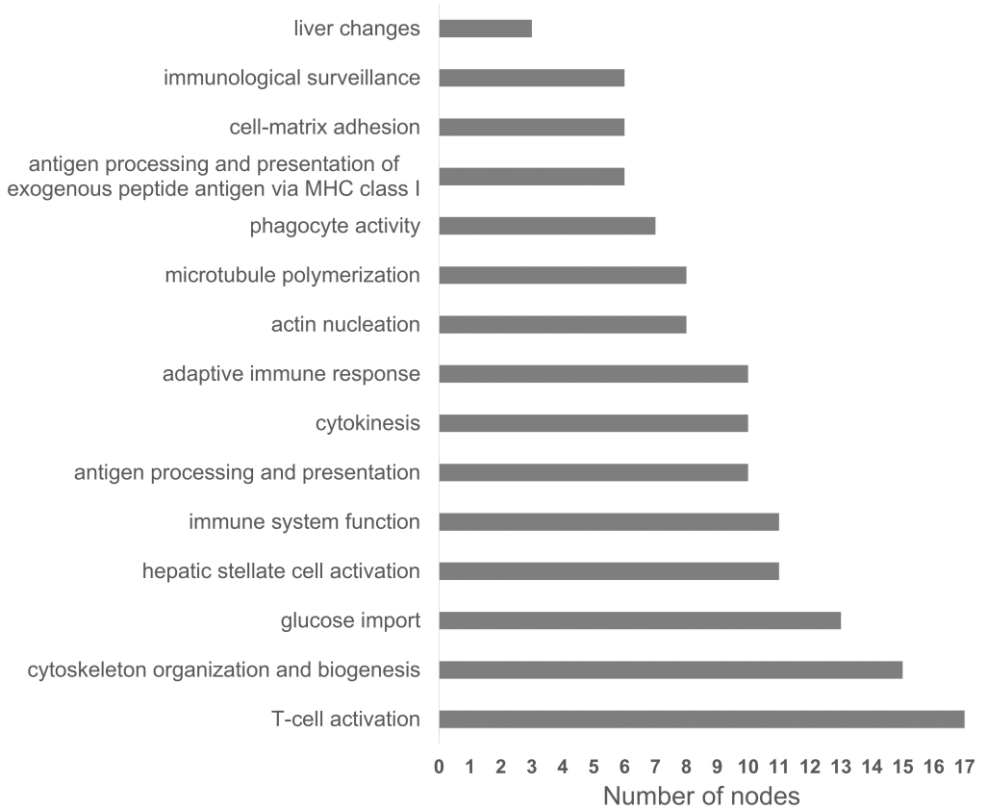


Figure 9. The 15 most significative cell processes and number of nodes (gene participants) in Nelore cattle divergent for RFI using Pathway Studio.

On balance, in Figure 11 showed the principal component analysis (PCA) of the he main nodes in all study: CD74, BoLA-DRA, BoLA-DRB3, LRRFIP1, NR4A2, DDB2, VIM, PAK1, PRC1, ACTB, RPS6KA1 and PKM of LRFI in liver tissue of Nelore cattle in G1 and G2.

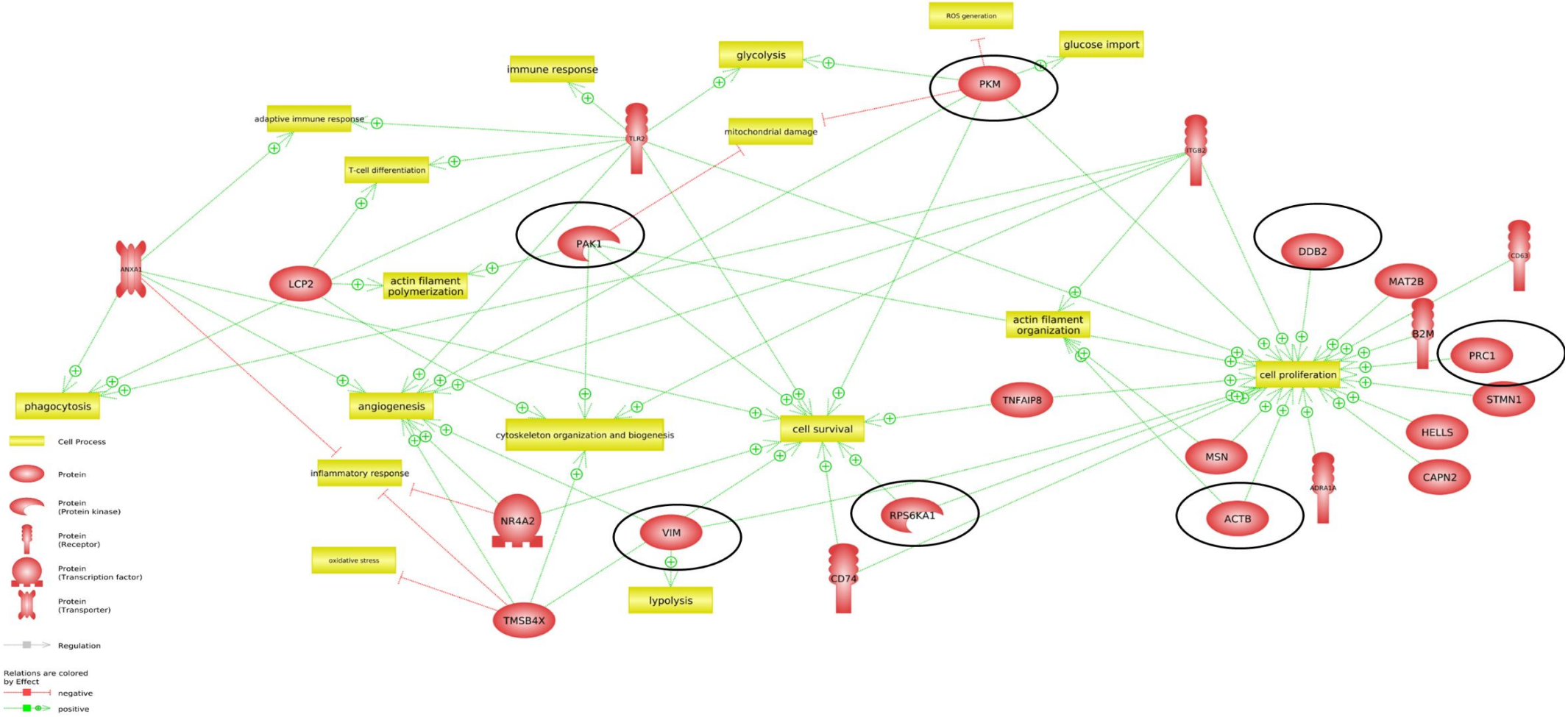


Figure 10. Cell processes associated with main nodes obtained from networks analysis using Pathway Studio. The circled in black represents the 7 main nodes. Highlighted in red are genes overexpressed in LRFI a group. Green arrows indicate a positive regulation effect, while red arrows has a negative regulation effect.

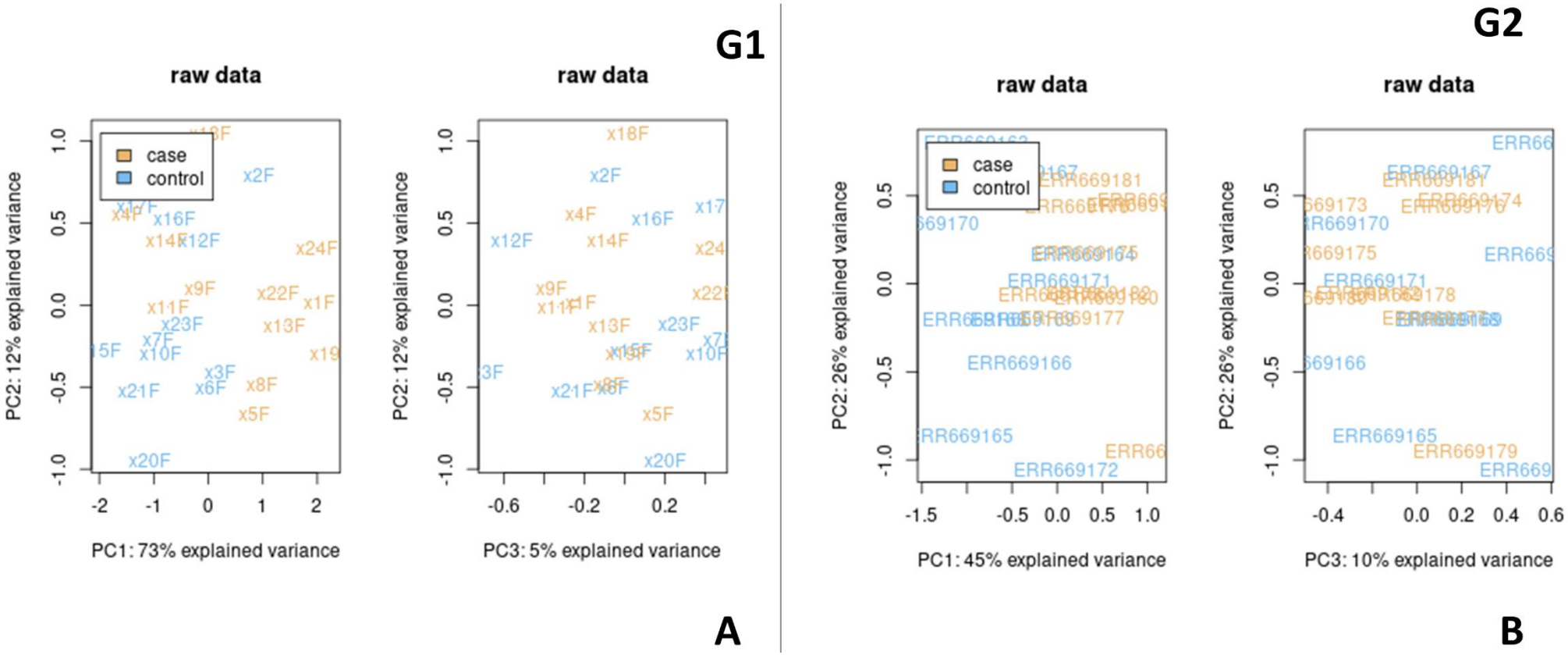


Figure 11. Principal component analysis (PCA) of the he main nodes in study: D74, BoLA-DRA, BoLA-DRB3 as hub genes, LRRFIP1 and NR4A2, as transcription factors and DDB2, VIM, PAK1, PRC1, ACTB, RPS6KA1 and PKM of LRFI (Case, orange color) and HRFI (control, blue color) in liver tissue of Nelore cattle. Each data point represents one sample. **A.** PCA the main nodes in LRFI group in G1, the total percentage of PCA mapping variability is 90%. **B.** PCA the main nodes in LRFI group in G2, the total percentage of PCA mapping variability is 80%.

4. Discussion

In this study we highlighted three main hub genes (C*D74, BoLA-DRA (HLA-DRA*) and *BoLA-DRB3 (HLA-DRB3)*) overexpressed in LRFI *versus* HRFI in both genetic groups (Figure 2 and 3) among the 88 common DEGs.

The *CD74* gene is a cell-surface receptor for the cytokine macrophage migration inhibitory factor, mediating the regulation of inflammatory responses, cell growth (FEX SVENNINGSEN et al., 2017; GIL-YAROM et al., 2017) and T cell activation (MUKHERJEE et al., 2008). It is also crucial for the regulation of B cell survival (SCHWARTZ et al., 2009; COHEN; SHACHAR, 2012). Moreover, CD74 stimulates ERK1/2 and AKT signaling, regulating cell proliferation (ZUO et al., 2011; BUCALA; SHACHAR, 2014; LE HIRESS et al., 2015) and interacts with transcription factors such as NF-κB, joining enriched regulators for genes involved in apoptosis, immune response and cellular migration (GIL-YAROM et al., 2017). *CD74* deficiency decreases adaptive immune responses (SUN et al., 2010). It regulates cell survival (KITANGE et al., 2010; RUVOLO et al., 2019) and its absence leads to cell death (BECKER-HERMAN et al., 2021). According to XU et al. (2013), *CD74* showed a direct association with the *B2M* gene, explained in chapter 2, as a biomarker in this study.

Our results show that *CD74* is overexpressed in LRFI animals and acts by improving the immune response, cell proliferation and cell survival of this group of cattle. In addition, this gene could have a significant action on hepatic cells, exerting an inflammatory state and regulating the accumulation of lipids in the liver thereby preventing obesity in the more efficient group of cattle (MISHRA et al., 2016).

The major cattle histocompatibility complex (MHC) class II DR gene product is encoded by the *BoLA-DRA* and *BoLA -DRB*3 genes and have a crucial role in determining the immune response (FISCH et al., 2021). Furthermore, they are vital specifically in the adaptive immune system that offers a unique alternative for addressing immunological and evolutionary problems (BOHÓRQUEZ et al., 2020; SANCHEZ-MAZAS; LEMAÎTRE; CURRAT, 2012). *BoLA* genes encode the cell surface glycoproteins necessary for T lymphocytes (ROCK et al., 2016) and are expressed by professional antigen-presenting cells, thus playing a key role in the defense against pathogens (BOHÓRQUEZ et al., 2020).

Our results reveal an overexpression of these two genes of the *BoLA* family in

the more efficient group of cattle, indicating that they have a better defensive response to their environment, since a low expression of the *BoLA-DRA* gene contributes to deficient defense (BIRLE; NEBE; GESSLER, 2003).

In cattle, the *BoLA-DRA* overexpressed gene was identified for mastitis (ASSELSTINE et al., 2019) and *BoLA-DRB3* was associated with production and milk quality traits (A; Z; LÓPEZ-HERRERA, 2014). This finding is also observed in ovines (TSHILWANE et al., 2019). In a recent study on Nelore cattle, the BoLA gene was associated with bovine immune response to infections (FERNANDES JÚNIOR et al., 2020).

The results suggest that the immune system is a key factor in explaining the differences we encountered between the LRFI and the HRFI animals in the two genetic groups. Furthermore, *CD74* could help the non-accumulation of lipids, thereby protecting against liver injury.

In addition, this study revealed the overexpression of two important transcription factors in the LRFI group (*LRRFIP1* and *NR4A2*). These factors share vital cellular processes, such as activation of angiogenesis and inhibition of apoptosis (VAN TIEL; DE VRIES, 2012; CABALLERO et al., 2014; BEARD; HAVE; CHEN, 2015; AMOASII et al., 2016; SUN et al., 2016; LÄHTEENVUO et al., 2020; JING et al., 2021; KAWASAKI et al., 2021; WANG et al., 2021; ZHU et al., 2021).

LRRFIP1 may enhance the proliferation and migration of vascular smooth muscle cells (MA; REN; GUAN, 2019) and inhibit apoptosis in endothelial cells (ZHU et al., 2021). LRRFIP1 knockdown in liver tissue results in inhibited cell growth (YAO et al., 2020).

NR4A2 is important for prostaglandin E2-mediated regulation of apoptosis by blocking cleavage of caspase-3 (HOLLA et al., 2006) and regulates mitochondrial survival (GKIKAS; TSAMPOULA; POLITIS, 2017). Moreover, this gene regulates cell differentiation (LÄHTEENVUO et al., 2020) promoting cell survival (HOT et al., 2011) and autophagy (LIU et al., 2018; SHI et al., 2017). The overexpression the *NR4A2* gene was observed to induce cell proliferation and reduce cell apoptosis (BEARD; HAVE; CHEN, 2015; SUN et al., 2016; LÄHTEENVUO et al., 2020). The *NR4A2* knockdown can inhibit cell proliferation, migration (JING et al., 2021) and promote intrinsic apoptosis (CABALLERO et al., 2014; WANG et al., 2021). The *NR4A2* gene also induces endothelial activation (HU; OLSEN, 2016).

NR4A2 is a key transcriptional regulator of energy metabolism in liver functions (HU et al., 2014) and increases fatty acid oxidation (HOLLA et al., 2011; KELLY et al., 2018). In skeletal muscle it can, potently, stimulate muscle glucose uptake, by normalizing glucose and insulin levels and conferring resistance to hepatic steatosis in obese mice (AMOASII et al., 2016, 2019). In humans, this gene inhibits adipogenesis (CHAO et al., 2008) and inflammatory responses (BONTA et al., 2006). It also interferes in *NR4A2* overexpression and blocks the activation of hepatic stellate cells (CHEN et al., 2016; TSUCHIDA; FRIEDMAN, 2017). Furthermore, it has a neuroprotective role (HAWK; ABEL, 2011; XU et al., 2015).

The transcription factors in the efficient group could be mobilizing the energy and not accumulating lipids, activating angiogenesis, inhibiting apoptosis, protecting toxicity in the liver and silencing the activation of stellate cells as they are the boosters of central hepatic fibrosis (a consequence of fatty liver) and hepatic injury in humans (Figure 4).

The 54 nodes were found in the network analysis using a free database (BioGRID). The most relevant seven nodes were overexpressed in the LRFI animals: *DDB2, VIM, PAK1, PRC1, ACTB, RPS6KA1* and *PKM*. This study found relevant cell processes; cell proliferation, cytoskeleton organization, energy metabolism and angiogenesis (Figure 10). The seven main nodes activate the cellular process of cell proliferation.

Cell proliferation

The *DDB2* gene activates cell proliferation, which coincides with the findings of PERUCCA et al., (2015). It stimulates the activity of the cellular cycle and global genomic repair (NAG et al., 2001). The knockdown of *DDB2* inhibits cell proliferation and migration (DONIZY; MARCZUK, 2019).The *PRC1* gene also activates cell proliferation (HERNÁNDEZ-ORTEGA et al., 2019). *ACTB* appears to have a role in cellular proliferation as studies report that its underexpression decreases the mobility and proliferation in fibroblasts giving way to proliferation and migration defects and embryonic lethality. (JOSEPH; SRIVASTAVA; PFISTER, 2012). Knockouts of *ACTB* bring about proliferation defects (ALMUZZAINI et al., 2016) and decreased fitness (PATRINOSTRO et al., 2018). The *PKM* gene activates proliferation cells and generates more ATP and fewer metabolic intermediates (NANDI; DEY, 2020). *RPS6KA1* is involved in controlling

proliferation and differentiation of stem cells (CARNEVALLI et al., 2010); it is also essential for protein synthesis and cell growth (WEISS et al., 2019). *VIM* influences signal transduction pathways and participates in regulating proliferation, adhesion and migration of cells (YAO et al., 2013; MOYO; NICHOLSON; ARBUTHNOT, 2016; HU et al., 2019), apoptosis inhibition (TRAVERSA et al., 2021) and contractility (LI et al., 2020). This gene is also up-regulated in LRFI animals (KERN et al., 2016). *PAK* is involved in regulating cell motility, survival, cell proliferation (HE; BALDWIN, 2008; POORGHOLI BELVERDI et al., 2012; ZHANG et al., 2017) and cell growth (YUAN et al., 2016). PAK in a transcriptomic study in cattle also showed cell proliferation in the LRFI group (KERN et al., 2016). These genes could be regulating different mechanisms of cell proliferation and can help other cell processes, among others, turn-over protein and the immune system in the most efficient animals.

Cytoskeletal organization

We highlight three within the seven most relevant nodes for cytoskeletal organization: *VIM, PAK1 and ACTB,* which are overexpressed in the efficient group. The *FYB1* and *TUBA4A* genes also stand out as common in the two genetic groups in the co-expression analysis (Figure 2).

It was observed that the better the bovine feed efficiency, the better the cytoskeletal organization (MUKIIBI et al., 2018). Our results are reinforced as we have various overexpressed genes that are associated with cytoskeletal regulation, such as the *VIM* gene, which is a cytoskeletal protein that is mainly expressed in the cytoskeleton of mesenchymal cells and functions to maintain the stability of cell structure (LUND et al., 2010; HU et al., 2019). *PAK1* is involved in regulating cytoskeletal organization (KICHINA et al., 2010; KE et al., 2010; POORGHOLI BELVERDI et al., 2012; ZHANG et al., 2017), cytoskeleton remodeling by affecting the processes of actin depolymerization and nucleation (CHENG et al., 2021) and actin phosphorylation (ZHANG et al., 2010). The *ACTB* gene is essential for normal cell migration and facilitates normal actin assembly (CORNACHIONE et al., 2014; CHIA et al., 2020). FYB1 gene also acts as regulator the cytoskeleton of actin (KUROPKA et al., 2016). Another gene related to the organization of the extracellular network was the TUBA4A gene, this gene was found increased in the rumen of steers with LRFI (MUKIIBI et al., 2018) but the other study with hepatic tissue with steers was up-regulated in an efficient

group (HIGGINS et al. al., 2019), which coincides with our results.

Energy metabolism

We found genes that participate in metabolic processes, such as glycolysis and those with a tendency to lipolysis, namely P*KM, DDB2, RPS6KA1* and *VIM,* overexpressed in the LRFI group. The *LCP1, LETMD1, SNX1, ANXA1, GRN* and *LUM* genes were also highlighted as common in the two genetic groups in the co-expression analysis (Figure 2).

Glycolysis is a catabolic process of glucose hydrolysis necessary for energy and biosynthetic intermediates, while gluconeogenesis is an important glucose production process for maintaining blood glucose levels. The *PKM* gene induces glycolysis which facilitates the production of ATP and hepatic cell proliferation observing its contribution in the most feed efficient group (KITAMURA et al., 2011; XIONG et al., 2011; LEE et al., 2015; QI et al., 2020). The *PKM* gene is up-regulated in LRFI animals and it could be increasing glucose uptake (LEE et al., 2015; LU; HUNTER, 2018; CHEN et al., 2021; YAN et al., 2021). The *DDB2* gene promotes metabolism of glucose (LAI; HU, 2019). The *RPS6KA1* gene is dispensable for terminal adipocyte differentiation, but plays a critical role in early adipocyte commitment (CARNEVALLI et al., 2010). The *VIM* gene facilitates lipolysis and this absence results in smaller adipocytes (KUMAR et al., 2007; SHEN et al., 2012; WILHELMSSON et al., 2019). We detected another overexpressed gene in LRFI animals, *LCP1*; its deficiency promotes lipid catabolism and suppresses adipogenesis and lipogenesis (SUBRAMANI; YUN, 2021). *LETMD1* plays a crucial role in energy homeostasis by regulating mitochondrial structures and functions (SONG et al., 2020). It activates the formation of brown fat (KONG et al., 2021). This gene is underexpressed in LRFI animals and also helps in lipolysis. *SNX1* overexpressed in LRFI animals regulates growth through the EGF signaling pathway (GULLAPALLI et al., 2004). We found *ANXA1* overexpressed in LRFI animals which regulates body weight, glucose metabolism, fatty acid uptake in adipose tissue and metabolic health (ZHANG et al., 2021); it can control lipolysis of adipocytes (GREWAL et al., 2019). In addition, in cattle, HASANKHANI et al., (2021) found *ANXA1* as a hub gene and related it to respiratory disease. The *GRN* gene, which is overexpressed in LRFI animals, was considered a candidate marker for feed efficiency in a bovine study on muscle (MUKHERJEE et al., 2020). Another

overexpressed gene in LRFI animals, *LUM,* was identified as the central gene in fatty liver (CHANG et al., 2021) and requires further research in the gene and its function in cattle. This gene was already related to energy metabolism in chickens in LRFI (LIU et al., 2019). Energy generating pathways, such as glycolysis and oxidative phosphorylation, (KONG et al., 2016) were found in LRFI bovines. The efficient group could be mobilizing energy, as well as not accumulating lipids in the liver, which is beneficial for avoiding fatty liver.

Angiogenesis

Alterations in angiogenesis may have implications in feed efficiency, as vascularity of the liver is crucial to carry out its functions adequately (LORENTE; HAUTEFEUILLE; SANCHEZ-CEDILLO, 2020). ALEXANDRE et al. (2015) studying liver and DE OLIVEIRA et al. (2018) studying liver and muscle, reported differentially co-expressed genes related to angiogenesis in the Nelore breed divergent for feed efficiency.

In our study we found that the *PKM* gene facilitates endothelial cell angiogenesis (LIEW; KUBES, 2019; KIM et al., 2020) and *PAK1,* a well-known regulator of angiogenesis, stimulated endothelial proliferation (PUTO et al., 2003; DIEBOLD et al., 2009).

Oxidative stress

We observed several genes regulating oxidative stress and decreasing apoptosis, such as *PKM* and *PAK1.* The *COTL1* and *CYBB* genes were common in both genetic groups in the co-expression analysis (Figure 2).

PKM, overexpressed in LRFI animals, stabilizes the production of ROS by mitochondria contributing the oxidative potential of these macrophages and improving the immunological state (IQBAL et al., 2014; GROH et al., 2018). *PKM* overexpressed in LRFI activation prevented high glucose-induced elevation of ROS production (QI et al., 2017) and reduced mitochondrial damage (HAMMER; DIAKONOVA, 2015). The *PAK1* gene overexpressed in the efficient group, this gene plays a role in reducing oxidative stress and apoptosis (AHN et al., 2021). *COTL1,* also mentioned in chapter two, overexpressed in LRFI animals and was involved in the improvement of oxidative stress and helping to improve the immune system (XIA et al., 2018). Another gene that helps the immune system and improves oxidative stress is *CYBB* overexpressed in LRFI animals; it is down-regulated in situations of oxidative stress (GIEFING et al., 2013; EGUCHI

et al., 2021). Moreover, reduced oxidative stress of the liver was found in the efficient steers (CASAL et al., 2020). Efficient Nelore cattle seem to have a greater adaptation capacity to oxidative stress (TIZIOTO et al., 2015, 2016; DE OLIVEIRA et al., 2018; FONSECA et al., 2019).

Cell survival

Participant nodes increased cell survival in a process in which we highlight the participation of our overexpressed genes in LRFI animals: *PKM, RPS6KA1, PAK1* and *PRC1*.

PKM protects cell survival (TAYLOR et al., 2021; WU et al., 2021) and boosts IL-10 production (MARELLI-BERG; JANGANI, 2018). *RPS6KA1* also promotes cell survival (SHIMAMURA et al., 2000; FENTON; GOUT, 2011; CRONIN; BROOKE; PRISCHI, 2021) and controls cell growth and size (RICHARDS et al., 1999; OKAYAMA, 2012).This gene intensifies protective autophagy under stress conditions (HAĆ et al., 2015) and it inhibits apoptosis (CRONIN; BROOKE; PRISCHI, 2021). *PAK* plays an important role in promoting cell survival (DAN; WATANABE; KUSUMI, 2001; RIDER et al., 2007; AHN et al., 2016) and inhibits apoptosis of human stem cells (MA et al., 2020). This finding was specifically observed in hepatocytes (HAO et al., 2015). In addition, the *PRC1* gene in liver was blocked, thereby inducing apoptosis (LI et al., 2021).

Figure 12 summarizes the most important cellular processes and associated genes highlighted in our study in the LRFI group of the two genetic groups. We observed increases in the immune system which reduced ROS, thereby decreasing apoptosis, mobilizing energy as well as avoiding the accumulation of lipids and improving the organization of the cytoskeleton and thus increasing cell survival.

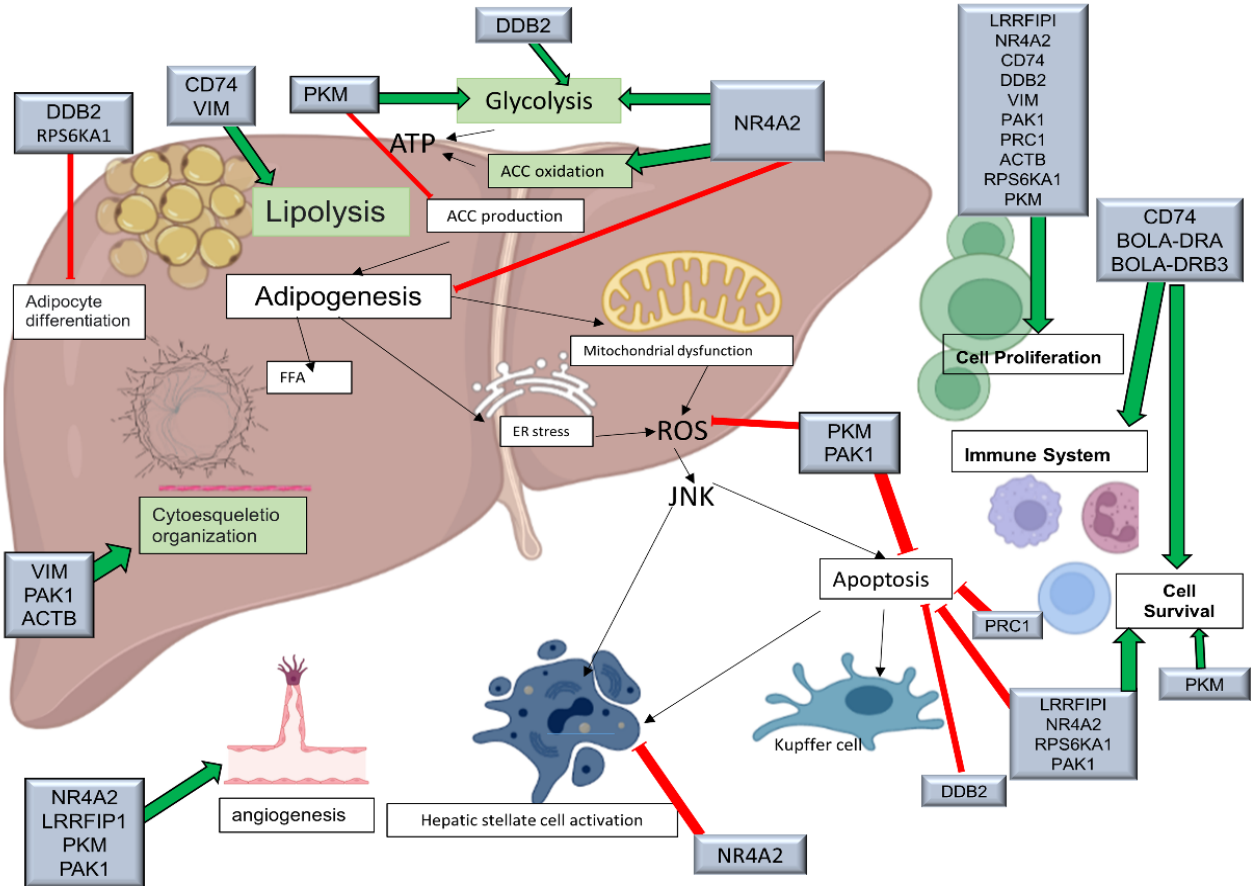


Figure 12. Summary of the principal cellular processes and genes that could be bearing an influence on the LRFI group in the liver of Nelore cattle. The green arrows indicate a positive regulation effect, while the red arrows show a negative regulation effect.

5. Conclusion

Three transcripts were determinate as hub (CD74, BoLA-DRA and BoLA-DRB3); and seven, as the most relevant nodes in interaction network and co-expression analysis (DDB2, VIM, PAK1, PRC1, ACTB, RPS6KA1 and PKM). The direct association between B2M gene, explained in chapter 2, as a biomarker in this study, and CD74, a hub gene, both more expressed in LRFI group, can indicate important candidate genes to select more efficient animals.

List of abbreviations

Differentially expressed genes (DEGs), Residual Feed Intake (RFI), major histocompatibility complex (MHC), ribonucleic acid sequencing RNA-Seq, Database for Annotation, Visualization and Integrated Discovery (DAVID), maximal clique centrality

(MCC), Low Residual Intake (LRFI), High Low Residual Intake (HRFI), Serine/threonine kinase 1 and 2 (ERK1/2), and Serine-Threonine Protein Kinase (AKT), Epidermal growth factor (EGF) and Reactive oxygen species (ROS).

Declarations

Ethics approval and consent to participate

All experimental procedures were approved by Ethics Committee of the School of Agricultural and Veterinarian Sciences, São Paulo State University (UNESP), Jaboticabal, SP, Brazil (protocol number 18.340/16).

Funding

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and São Paulo Research Foundation – FAPESP (grant #2009/16118-5, #2017/10630-2 e #2018/20026-8).

Acknowledgements

The authors thank the Qualitas Nelore breeding program company for providing the tissue samples and database used in this study.

References

A, J. Z.; Z, J. E.; LÓPEZ-HERRERA, A. Association of gene BoLA DRB3.2 with production traits in a dairy herd of Antioquia, Colombia. **Revista MVZ Córdoba**, v. 19, n. 2, p. 4116–4129, 2014.

AHN, M. et al. The p21-activated kinase (PAK1) is involved in diet-induced beta cell mass expansion and survival in mice and human islets. **Diabetologia**, v. 59, n. 10, p. 2145–2155, 2016.

AHN, M. et al. A requirement for PAK1 to support mitochondrial function and maintain cellular redox balance via electron transport chain proteins to prevent β-cell apoptosis. **Metabolism: clinical and experimental**, v. 115, p. 154431, 2021.

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

ALMUZZAINI, B. et al. In β-actin knockouts, epigenetic reprogramming and rDNA transcription inactivation lead to growth and proliferation defects. **FASEB journal**, v. 30, n. 8, p. 2860–2873, 2016.

AMOASII, L. et al. A MED13-dependent skeletal muscle gene program controls systemic glucose homeostasis and hepatic metabolism. **Genes & development**, v. 30, n. 4, p. 434–446, 2016.

AMOASII, L. et al. NURR1 activation in skeletal muscle controls systemic energy homeostasis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 116, n. 23, p. 11299–11308, 2019.

ASSELSTINE, V. et al. Genetic mechanisms regulating the host response during mastitis. **Journal of Dairy Science**, v. 102, n. 10, p. 9043–9059, 2019.

AYDIN, B.; ARGA, K.Y. Co-expression Network Analysis Elucidated a Core Module in Association With Prognosis of Non-functioning Non-invasive Human Pituitary Adenoma. **Frontiers in Endocrinology**, v.10, 2019 .

BARABÁSI, A. L.; OLTVAI, Z. N. Network biology: understanding the cell's functional organization. **Nature reviews. Genetics**, v. 5, n. 2, p. 101–113, 2004.

BEARD, J. A.; TENGA, A.; CHEN, T. The interplay of NR4A receptors and the oncogene-tumor suppressor networks in cancer. **Cellular signalling**, v. 27, n. 2, p. 257–266, 2015.

BECKER-HERMAN, S. et al. CD74 is a regulator of hematopoietic stem cell maintenance. **PLoS biology**, v. 19, n. 3, 2021.

BENJAMINI, Y.; HOCHBERG, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. **Source: Journal of the Royal Statistical Society**, v. 57, n. 1, p. 289–300, 1995.

BIRLE, A.; NEBE, C. T.; GESSLER, P. Age-related Low Expression of HLA-DR Molecules on Monocytes of Term and Preterm Newborns With and Without Signs of Infection. **Journal of Perinatology**, v. 23, n. 4, p. 294–299, 2003.

BRANCO, R. et al. Consumo alimentar residual de machos nelore selecionados para peso pós-desmame. **46° Reunião Anual da Sociedade Brasileira de Zootecnia**, p. 3, 2009.

BOHÓRQUEZ, M. D. et al. Major Histocompatibility Complex Class II (DRB3) Genetic Diversity in Spanish Morucha and Colombian Normande Cattle Compared to Taurine and Zebu Populations. **Frontiers in Genetics**, v. 10, p. 1293, 10 jan. 2020.

BONTA, P. I. et al. Nuclear receptors Nur77, Nurr1, and NOR-1 expressed in atherosclerotic lesion macrophages reduce lipid loading and inflammatory responses. **Arteriosclerosis, thrombosis, and vascular biology**, v. 26, n. 10, p. 2288–2294, out. 2006.

BUCALA, R.; SHACHAR, I. The integral role of CD74 in antigen presentation, MIF signal transduction, and B cell survival and homeostasis. **Mini reviews in medicinal chemistry**, v. 14, n. 14, p. 1132–1138, 26 jan. 2014.

CABALLERO, I. S. et al. Lassa and Marburg viruses elicit distinct host

transcriptional responses early after infection. **BMC Genomics**, v. 15, n. 1, 2014.

CARNEVALLI, L. S. et al. S6K1 plays a critical role in early adipocyte differentiation. **Developmental cell**, v. 18, n. 5, p. 763–774, 2010.

CASAL, A. et al. Differential hepatic oxidative status in steers with divergent residual feed intake phenotype. **Animal**, v. 14, n. 1, p 78-85, 2020.

CEACERO, T. M. et al. Phenotypic and Genetic Correlations of Feed Efficiency Traits with Growth and Carcass Traits in Nelore Cattle Selected for Postweaning Weight. **PLOS ONE**, v. 11, n. 8, p. 66,2016.

CHANG, Y. et al. LUM is the hub gene of advanced fibrosis in nonalcoholic fatty liver disease patients. **Clinics and research in hepatology and gastroenterology**, v. 45, n. 1, 2021.

CHAO, L. C. et al. Inhibition of adipocyte differentiation by Nur77, Nurr1, and Nor1. **Molecular endocrinology**, v. 22, n. 12, p. 2596–2608, 2008.

CHEN, P. et al. Adenovirus-mediated expression of orphan nuclear receptor NR4A2 targeting hepatic stellate cell attenuates liver fibrosis in rats. **Scientific Reports**, v. 6, n. 1, p. 1–13, 2016.

CHEN, R. L. et al. Isovitexin potentiated the antitumor activity of cisplatin by inhibiting the glucose metabolism of lung cancer cells and reduced cisplatin-induced immunotoxicity in mice. **International immunopharmacology**, v. 94, 2021.

CHENG, C. et al. Disruption of the autism-related gene Pak1 causes stereocilia disorganization, hair cell loss, and deafness in mice. **Journal of genetics and genomics**, v. 48, n. 4, p. 324–332, 2021.

CHIA, S. B. et al. Dysregulation of the glutaredoxin/ S- glutathionylation redox axis in lung diseases. **American journal of physiology,** v. 318, n. 2, p. C304–C327, 2020.

CHIN, C.H. et al. cytoHubba: identificación de objetos hub y subredes a partir de interactomas complejos. **BMC Systems Biology**. v.8 n.4, p.11, 2014.

COHEN, S.; SHACHAR, I. Cytokines as regulators of proliferation and survival of healthy and malignant peripheral B cells. **Cytokine**, v. 60, n. 1, p. 13–22, 2012.

CORNACHIONE, A. S. et al. Arginylation of myosin heavy chain regulates skeletal muscle strength. **Cell reports**, v. 8, n. 2, p. 470–476, 2014.

CRONIN, R.; BROOKE, G. N.; PRISCHI, F. The role of the p90 ribosomal S6 kinase family in prostate cancer progression and therapy resistance. **Oncogene**, v. 40, n. 22, p. 3775–3785, 2021.

DAN, I.; WATANABE, N. M.; KUSUMI, A. The Ste20 group kinases as regulators of MAP kinase cascades. **Trends in cell biology**, v. 11, n. 5, p. 220–230, 2001.

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. 1, p. 1–12, 2018.

DE OLIVEIRA et al. An integrative transcriptome analysis indicates regulatory mRNA-miRNA networks for residual feed intake in Nelore cattle. **Scientific Reports**. v,8, p.17072, 2018.

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

DIEBOLD, I. et al. Phosphodiesterase 2 mediates redox-sensitive endothelial cell proliferation and angiogenesis by thrombin via Rac1 and NADPH oxidase 2. **Circulation research**, v. 104, n. 10, p. 1169–1177, 2009.

DONG, J.; HORVATH, S. Understanding network concepts in modules. **BMC systems biology**, v. 1, 4 jun. 2007. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17547772/>.

DONIZY, P.; MARCZUK, J. Selected Golgi-Localized Proteins and Carcinogenesis: What Do We Know? **Results and problems in cell differentiation**, v. 67, p. 487–529, 2019.

EGUCHI, N. et al. The Role of Oxidative Stress in Pancreatic β Cell Dysfunction in Diabetes. **International Journal of Molecular Sciences**, v. 22, n. 4, p. 1509, 2021.

FENTON, T. R.; GOUT, I. T. Functions and regulation of the 70kDa ribosomal S6 kinases. **The international journal of biochemistry & cell biology**, v. 43, n. 1, p. 47–59, 2011.

FERNANDES JÚNIOR, G. A. et al. Whole-genome sequencing provides new insights into genetic mechanisms of tropical adaptation in Nelore (Bos primigenius indicus). **Scientific Reports ,**v. 10, n. 1, p. 1–7, 2020.

FEX SVENNINGSEN, Å. et al. Macrophage migration inhibitory factor (MIF) modulates trophic signaling through interaction with serine protease HTRA1. **Cellular and molecular life sciences**, v. 74, n. 24, p. 4561–4572, 2017.

FILTEAU, M. et al. Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. **Molecular Biology and Evolution**, v. 30, n. 6, p. 1384–1396, 2013.

FISCH, A. et al. Integral Use of Immunopeptidomics and Immunoinformatics for the Characterization of Antigen Presentation and Rational Identification of BoLA-DR–Presented Peptides and Epitopes. **The Journal of Immunology**, v. 206, n. 10, p. 2489–2497, 2021.

FONSECA et al. Liver proteomics unravel the metabolic pathways related to feed efficiency in beef cattle. **Scientific Reports**. v. 9, n.5364, 2019.

GERACI, F.; SAHA, I.; BIANCHINI, M. Editorial: RNA-Seq Analysis: Methods,

Applications and Challenges. **Frontiers in Genetics**, v. 11, p. 220, 2020.

GIEFING, M. et al. Hodgkin-Reed-Sternberg Cells in Classical Hodgkin Lymphoma Show Alterations of Genes Encoding the NADPH Oxidase Complex and Impaired Reactive Oxygen Species Synthesis Capacity. **PLOS ONE**, v. 8, n. 12, p. e84928, 2013.

GIL-YAROM, N. et al. CD74 is a novel transcription regulator. **Proceedings of the National Academy of Sciences of the United States of America**, v. 114, n. 3, p. 562–567, 2017.

GOV, E.; ARGA, K.Y. Differential co-expression analysis reveals a novel prognostic gene module in ovarian cancer. S**cientific Reports**, v. 7, p.4996, 2017,

GREWAL, T. et al. Annexins in Adipose Tissue: Novel Players in Obesity. **International journal of molecular sciences**, v. 20, n. 14, 2019.

GRION, A. L. et al. Selection for feed efficiency traits and correlated genetic responses in feed intake and weight gain of Nelore cattle. **Journal of Animal Science**, v. 92, n. 3, p. 955–965, 2014.

GROH, L. et al. Monocyte and macrophage immunometabolism in atherosclerosis. **Seminars in immunopathology**, v. 40, n. 2, p. 203–214, 2018.

HAĆ, A. et al. S6K1 controls autophagosome maturation in autophagy induced by sulforaphane or serum deprivation. **European journal of cell biology**, v. 94, n. 10, p. 470–481, 2015.

HAMMER, A.; DIAKONOVA, M. Tyrosyl phosphorylated serine-threonine kinase PAK1 is a novel regulator of prolactin-dependent breast cancer cell motility and invasion. **Advances in experimental medicine and biology**, v. 846, p. 97–137, 2015.

HASANKHANI, A. et al. Integrated Network Analysis to Identify Key Modules and Potential Hub Genes Involved in Bovine Respiratory Disease: A Systems Biology Approach. **Frontiers in genetics**, v. 12, 2021.

HAWK, J. D.; ABEL, T. The role of NR4A transcription factors in memory formation. **Brain Research Bulletin**, v. 85, n. 1–2, p. 21–29, 2011.

HERNÁNDEZ-ORTEGA, S. et al. Phosphoregulation of the oncogenic protein regulator of cytokinesis 1 (PRC1) by the atypical CDK16/CCNY complex. **Experimental and molecular medicine**, v. 51, n. 4, 2019.

HIGGINS, M. G. et al. The effect of breed and diet type on the global transcriptome of hepatic tissue in beef cattle divergent for feed efficiency. **BMC Genomics**, v. 20, n. 1, p. 1–11, 26 jun. 2019. Disponível em: <https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-019-5906-8>. Acesso em: 29 dez. 2021.

HOLLA, V. R. et al. Prostaglandin E2 Regulates the Nuclear Receptor NR4A2 in

Colorectal Cancer *. **Journal of Biological Chemistry**, v. 281, n. 5, p. 2676–2682, 2006.

HOLLA, V. R. et al. Nuclear orphan receptor NR4A2 modulates fatty acid oxidation pathways in colorectal cancer. **The Journal of biological chemistry**, v. 286, n. 34, p. 30003–30009, 2011.

HU, J. et al. Effects of menstrual blood-derived stem cells on endometrial injury repair. **Molecular medicine reports**, v. 19, n. 2, p. 813–820, 2019.

HU, K.; OLSEN, B. R. The roles of vascular endothelial growth factor in bone repair and regeneration. **Bone**, v. 91, p. 30, 2016.

HU, Y. et al. Accelerated partial hepatectomy-induced liver cell proliferation is associated with liver injury in Nur77 knockout mice. **The American journal of pathology**, v. 184, n. 12, p. 3272–3283, 2014.

HUANG, D. W.; SHERMAN, B. T.; LEMPICKI, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. **Nature protocols**, v. 4, n. 1, p. 44–57, 2009.

HUI, S. et al. ExpressionCorrelation. 2015. Available online at: http://apps.cytoscape.org/apps/expressioncorrelation.

IQBAL, M. A. et al. Missense Mutations in Pyruvate Kinase M2 Promote Cancer Metabolism, Oxidative Endurance, Anchorage Independence, and Tumor Growth in a Dominant Negative Manner. **The Journal of Biological Chemistry**, v. 289, n. 12, p. 8098, 2014.

JING, C. Y. et al. Hepatic stellate cells promote intrahepatic cholangiocarcinoma progression via NR4A2/osteopontin/Wnt signaling axis. **Oncogene**, v. 40, n. 16, p. 2910–2922, 2021.

JOSEPH, R.; SRIVASTAVA, O. P.; PFISTER, R. R. Downregulation of β-actin gene and human antigen R in human keratoconus. **Investigative Ophthalmology and Visual Science**, v. 53, n. 7, p. 4032–4041, 2012.

KANEHISA, M.; GOTO, S. KEGG: kyoto encyclopedia of genes and genomes. **Nucleic Acids Research**, v.1, n.28, p. 27-30, 2000.

KAWASAKI, S. et al. Silencing of LRRFIP1 enhances the sensitivity of gemcitabine in pancreatic cancer cells by activating JNK/c-Jun signaling. **Pancreatology** , v. 21, n. 4, p. 771–778, 2021.

KE, Y. et al. Novel bradykinin signaling in adult rat cardiac myocytes through activation of p21-activated kinase. **American journal of physiology. Heart and circulatory physiology**, v. 298, n. 4, 2010.

KELLY, M. E. et al. Orphan Nuclear Receptors in Colorectal Cancer. **Pathology oncology research**, v. 24, n. 4, p. 815–819, 2018.

KERN, R.J. Transcriptome differences in the rumen of beef steers with variation

in feed intake and gain. **Gene**, v15, n. 586, p. 12-26, 2016.

KICHINA, J. V. et al. PAK1 as a therapeutic target. **Expert opinion on therapeutic targets**, v. 14, n. 7, p. 703–725, 2010.

KIM, H. et al. Extracellular pyruvate kinase M2 facilitates cell migration by upregulating claudin-1 expression in colon cancer cells. **Biochemistry and cell biology**, v. 98, n. 2, p. 219–226, 2020.

KITAMURA, K. et al. Proliferative activity in hepatocellular carcinoma is closely correlated with glucose metabolism but not angiogenesis. **Journal of hepatology**, v. 55, n. 4, p. 846–857, 2011.

KITANGE, G. J. et al. Expression of CD74 in high grade gliomas: a potential role in temozolomide resistance. **Journal of neuro-oncology**, v. 100, n. 2, p. 177–186, 2010.

KOGELMAN, L. J. A. et al. Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA Sequencing in a porcine model. **BMC Medical Genomics**, v. 7, n. 1, 2014.

KONG, R.S. et al. Transcriptome profiling of the rumen epithelium of beef cattle differing in residual feed intake. BMC Genomics. v. 17, p. 592, 2016.

KONG, X. et al. Defective brown adipose tissue thermogenesis and impaired glucose metabolism in mice lacking Letmd1. **Cell Reports**, v. 37, n. 11, p. 110104, 2021.

KUMAR, N. et al. Requirement of vimentin filament assembly for beta3-adrenergic receptor activation of ERK MAP kinase and lipolysis. **The Journal of biological chemistry**, v. 282, n. 12, p. 9244–9250, 2007.

KUROPKA, B. et al. Tyrosine-phosphorylation of the scaffold protein ADAP and its role in T cell signaling. **Expert review of proteomics**, v. 13, n. 6, p. 545–554, 2016.

LÄHTEENVUO, J. et al. Susceptibility to Cardiac Arrhythmias and Sympathetic Nerve Growth in VEGF-B Overexpressing Myocardium. **Molecular therapy,** v. 28, n. 7, p. 1731–1740, 2020.

LAI, T. C.; HU, M. C. Regulation of liver receptor homologue-1 by DDB2 E3 ligase activity is critical for hepatic glucose metabolism. **Scientific reports**, v. 9, n. 1, 2019.

LANGFELDER, P.; MISCHEL, P. S.; HORVATH, S. When Is Hub Gene Selection Better than Standard Meta-Analysis?. **PLoS ONE**, v. 8, n. 4, 2013.

LE HIRESS, M. et al. Proinflammatory Signature of the Dysfunctional Endothelium in Pulmonary Hypertension. Role of the Macrophage Migration Inhibitory Factor/CD74 Complex. **American journal of respiratory and critical care medicine**, v. 192, n. 8, p. 983–997, 2015.

LEE, K. min et al. ECM1 promotes the Warburg effect through EGF-mediated activation of PKM2. **Cellular signalling**, v. 27, n. 2, p. 228–235, 2015.

LI, C. et al. Quantitative SUMO proteomics identifies PIAS1 substrates involved in cell migration and motility. **Nature communications**, v. 11, n. 1, 2020a.

LI, J. et al. Therapeutic targeting of the PLK1-PRC1-axis triggers cell death in genomically silent childhood cancer. **Nature communications**, v. 12, n. 1, 2021.

LIEW, P. X.; KUBES, P. The Neutrophil's Role During Health and Disease. **Physiological reviews**, v. 99, n. 2, p. 1223–1248, 2019.

LIM, D. et al. Gene co-expression analysis to characterize genes related to marbling trait in Hanwoo (Korean) cattle. **Asian-Australasian Journal of Animal Sciences**, v. 26, n. 1, p. 19–29, 2013.

LIU, R. et al. Relevance of the intestinal health-related pathways to broiler residual feed intake revealed by duodenal transcriptome profiling. **Poultry Science**, v. 98, n. 3, p. 1102–1110, 2019.

LORENTE, S.; HAUTEFEUILLE, M.; SANCHEZ-CEDILLO, A. The liver, a functionalized vascular structure. **Scientific Reports**, v. 10, n. 1, p. 1–10, 2020.

LU, Z.; HUNTER, T. Metabolic Kinases Moonlighting as Protein Kinases. **Trends in biochemical sciences**, v. 43, n. 4, p. 301–310, 2018.

LUND, N. et al. Vimentin expression influences flow dependent VASP phosphorylation and regulates cell migration and proliferation. **Biochemical and biophysical research communications**, v. 395, n. 3, p. 401–406, 2010.

MA, W. et al. Up-regulation of P21-activated kinase 1 in osteoarthritis chondrocytes is responsible for osteoarthritic cartilage destruction. **Bioscience reports**, v. 40, n. 1, 2020.

MA, Y.; REN, Y.; GUAN, J. Knockdown of GC binding factor 2 by RNA interference inhibits invasion and migration of vascular smooth muscle cells. **Molecular medicine reports**, v. 20, n. 2, p. 1781–1789, 2019.

MARELLI-BERG, F. M.; JANGANI, M. Metabolic regulation of leukocyte motility and migration. **Journal of leukocyte biology**, v. 104, n. 2, p. 285–293, 2018.

MAY, R. C. Fuel-injected cell motility. **Trends in Cell Biology**, v. 11, n. 7, p. 281, 2001.

MISHRA, A. et al. Role of antigen presenting cell invariant chain in the development of hepatic steatosis in mouse model. **Experimental cell research**, v. 346, n. 2, p. 188–197, 2016.

MITSIS T. et al. Transcription factors and evolution: An integral part of gene expression (Review). **World Academy of Sciences Journal,** n. 2, p. 3-8, 2020

MOYO, B.; NICHOLSON, S. A.; ARBUTHNOT, P. B. The role of long non-coding

RNAs in hepatitis B virus-related hepatocellular carcinoma. **Virus research**, v. 212, p. 103–113, 2016.

MUKHERJEE, R. et al. Exogenous CLIP localizes into endocytic compartment of cells upon internalization: implications for antigen presentation by MHC class II molecules. **Molecular immunology**, v. 45, n. 8, p. 2166–2176, 2008.

MUKHERJEE, S. et al. Muscle transcriptome signature and gene regulatory network analysis in two divergent lines of a hilly bovine species Mithun (Bos frontalis). **Genomics**, v. 112, n. 1, p. 252–262, 2020.

MUKIIBI, R. et al. Transcriptome analyses reveal reduced hepatic lipid synthesis and accumulation in more feed efficient beef cattle. **Scientific reports**, v. 8, n. 1, 2018.

NAG, A. et al. The Xeroderma Pigmentosum Group E Gene Product DDB2 Is a Specific Target of Cullin 4A in Mammalian Cells. **Molecular and Cellular Biology**, v. 21, n. 20, p. 6738, 2001.

NANDI, S.; DEY, M. Biochemical and structural insights into how amino acids regulate pyruvate kinase muscle isoform 2. **The Journal of biological chemistry**, v. 295, n. 16, p. 5390–5403, 2020.

NEPH, S. et al. Circuitry and Dynamics of Human Transcription Factor Regulatory Networks. **Cell.** n. 150, v.6, p 1274–1286, 2012.

OKAYAMA, H. Cell cycle control by anchorage signaling. **Cellular signalling**, v. 24, n. 8, p. 1599–1609, 2012.

OUGHTRED, R. et al. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. **Protein science : a publication of the Protein Society**, v. 30, n. 1, p. 187–200, 2021.

PATRINOSTRO, X. et al. Essential nucleotide- and protein-dependent functions of Actb/β-actin. **Proceedings of the National Academy of Sciences of the United States of America**, v. 115, n. 31, p. 7973–7978, 2018.

PERUCCA, P. et al. A DDB2 mutant protein unable to interact with PCNA promotes cell cycle progression of human transformed embryonic kidney cells. **Cell cycle ,** v. 14, n. 24, p. 3920–3928, 2015.

POORGHOLI BELVERDI, M. et al. Comprehensive analysis of TGF-β and BMP receptor interactomes. **European journal of cell biology**, v. 91, n. 4, p. 287–293, 2012.

PUTO, L. A. et al. p21-activated kinase 1 (PAK1) interacts with the Grb2 adapter protein to couple to growth factor signaling. **The Journal of biological chemistry**, v. 278, n. 11, p. 9388–9393, 2003.

QI, W. et al. Pyruvate kinase M2 activation may protect against the progression of diabetic glomerular pathology and mitochondrial dysfunction. **Nature medicine**, v. 23, n. 6, p. 753, 2017. D

QI, Y. et al. Kinome Profiling Reveals Abnormal Activity of Kinases in Skeletal Muscle From Adults With Obesity and Insulin Resistance. **The Journal of clinical endocrinology and metabolism**, v. 105, n. 3, 2020.

RICHARDS, S. A. et al. Ribosomal S6 kinase 1 (RSK1) activation requires signals dependent on and independent of the MAP kinase ERK. **Current biology,** v. 9, n. 15, p. 810–820, 1999.

RIDER, L. et al. JAK2 tyrosine kinase phosphorylates PAK1 and regulates PAK1 activity and functions. **The Journal of biological chemistry**, v. 282, n. 42, p. 30985–30996, 2007.

ROCK, K.L et al. Present Yourself! By MHC Class I and MHC Class II Molecules. **Trends in Immunology**, v. 37, p. 724-737, 2016.

RUVOLO, P. P. et al. LGALS3 is connected to CD74 in a previously unknown protein network that is associated with poor survival in patients with AML. **EBioMedicine**, v. 44, p. 126–137, 2019.

SERIN, E.A. R. et al. Learning from Co-expression Networks: Possibilities and Challenges. **Frontiers in Plant Science**, v. 7, 2016.

STARK, C. et al. BioGRID: a general repository for interaction datasets. **Nucleic Acids Res**, 1;34(Database issue):D535-9, 2006.

SCHWARTZ, V. et al. A functional heteromeric MIF receptor formed by CD74 and CXCR4. **FEBS Letters**, v. 583, n. 17, p. 2749–2757, 3 set. 2009.

SHANNON, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome research**, v. 13, n. 11, p. 2498–2504, 2003.

SHEN, W. J. et al. Ablation of vimentin results in defective steroidogenesis. **Endocrinology**, v. 153, n. 7, p. 3249–3257, 2012. D

SHIMAMURA, A. et al. Rsk1 mediates a MEK-MAP kinase cell survival signal. **Current biology**, v. 10, n. 3, p. 127–135, 2000.

SKINKYTE-JUSKIENE, R.; KOGELMAN, L.J.A.; KADARMIDEEN, H.N. Transcription factor co-expression networks of adipose RNA-Seq data reveal regulatory mechanisms of obesity. **Current Genomics, v. 29, n.4, p.**289-299, 2018.

SONG, R. et al. Deletion of Letmd1 leads to the disruption of mitochondrial function in brown adipose tissue. **bioRxiv**, p. 2020.11.18.388199, 2020.

STEVENS, A. et al. Análisis de redes: un nuevo enfoque para estudiar los trastornos endocrinos. **Journal of Molecular Endocrinology**, v. 52, p 79–93, 2013.

SUBRAMANI, M.; YUN, J. W. Loss of lymphocyte cytosolic protein 1 (LCP1) induces browning in 3T3-L1 adipocytes via β3-AR and the ERK-independent

signaling pathway. **The international journal of biochemistry & cell biology**, v. 138, 2021.

SUN, J. et al. Deficiency of antigen-presenting cell invariant chain reduces atherosclerosis in mice. **Circulation**, v. 122, n. 8, p. 808–820, 2010.

SUN, L. et al. Notch Signaling Activation in Cervical Cancer Cells Induces Cell Growth Arrest with the Involvement of the Nuclear Receptor NR4A2. **Journal of Cancer**, v. 7, n. 11, p. 1388–1395, 2016.

TAYLOR, S. R. et al. Dietary fructose improves intestinal cell survival and nutrient absorption. **Nature**, v. 597, n. 7875, p. 263–267, 2021.

TIZIOTO et al. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC Genomics**, v. 16, p.242, 2015.

TIZIOTO et al. Gene expression differences in Longissimus muscle of Nelore steers genetically divergent for residual feed intake. **Scientific Reports**, v. 6, p. 39493, 2016.

WRAY, G.A. et al. The evolution of transcriptional regulation in eukaryotes. **Molecular Biology and Evolution**, v. 20, n. 9, p.1377-419, 2003

ADDITIONAL FILES

Additional File 1. Cellular processes obtained with the Pathway Studio software for 54 nodes (genes) DEGs in Nelore cattle in liver divergent for RFI obtained for networks analysis in BioGRID.

Gene Set Seed	Overlap	Overlapping Entities	p-value
drug resistance	19	B2M;CD74;PKM;DAPK1;TMSB4X;VIM;TNFAIP8;CAPN2;TLR2;IFI30;FBXO15;PRC1;RPS6KA1;MSN;ITGB2;CD63;STMN1;NR4A2;ANXA1	0.004084188
phagocytosis	18	B2M;PKM;DAPK1;TMSB4X;VIM;TNFAIP8;CAPN2;TLR2;LCP2;FGR;LRP5;MSN;ITGB2;CD63;STMN1;ANXA1;STX7;NPC2	1.85246E-05
T-cell activation	17	B2M;CD74;PKM;DAPK1;TMSB4X;VIM;TNFAIP8;CAPN2;TLR2;IFI30;LCP2;MSN;ITGB2;CD63;COTL1;LY9;ANXA1	4.20245E-05
endocytosis	16	B2M;CD74;TMSB4X;VIM;SNX2;TLR2;IFI30;LRP5;SNX6;MSN;ITGB2;CD63;STMN1;ANXA1;STX7;NPC2	0.005335817
cytoskeleton organization and biogenesis	15	DAPK1;TMSB4X;VIM;CAPN2;TLR2;LCP2;ARHGAP30;RPS6KA1;MSN;TUBA4A;ITGB2;CD63;STMN1;COTL1;ANXA1	0.002562034
glucose import	13	DDB2;B2M;CD74;PKM;DAPK1;INPP5K;CAPN2;TLR2;FGR;LRP5;RPS6KA1;ITGB2;NR4A2	0.00282266
cell interaction	12	B2M;CD74;PKM;TMSB4X;VIM;TLR2;LCP2;MSN;ITGB2;CD63;STMN1;ANXA1	0.004493346
endothelial cell migration	12	PKM;DAPK1;TMSB4X;VIM;CAPN2;TLR2;LRP5;MSN;ITGB2;CD63;NR4A2;ANXA1	0.005855637
cell polarity	12	PKM;DAPK1;TMSB4X;VIM;TNFAIP8;CAPN2;TLR2;MSN;ITGB2;CD63;STMN1;ANXA1	0.008685034
hepatic stellate cell activation	11	CD74;PKM;MSN;TMSB4X;VIM;MAT2B;CD63;STMN1;NR4A2;TLR2;NPC2	1.72082E-05
locomotion	11	LCP2;B2M;LRP5;MSN;TMSB4X;ITGB2;STMN1;NR4A2;CAPN2;ANXA1;TLR2	0.003719093
immune system function	11	LCP2;PKM;MSN;TMSB4X;VIM;ITGB2;STMN1;TNFAIP8;ANXA1;TLR2;NPC2	0.003758141
endothelial cell adhesion	10	CD74;PKM;MSN;TMSB4X;VIM;ITGB2;CD63;CAPN2;ANXA1;TLR2	8.793E-05
platelet activation	10	LCP2;FGR;RPS6KA1;MSN;TMSB4X;VIM;ITGB2;CD63;ANXA1;TLR2	0.000309043
cell shape	10	LCP2;FGR;MSN;TMSB4X;VIM;ITGB2;STMN1;CAPN2;ANXA1;TLR2	0.00154084
antigen processing and presentation	10	B2M;CD74;TMSB4X;VIM;ITGB2;CD63;ANXA1;TLR2;NPC2;IFI30	0.001691566
cytokinesis	10	PKM;PRC1;RPS6KA1;MSN;TMSB4X;VIM;ITGB2;STMN1;TTC19;HELLS	0.00195269
NK cell mediated cytotoxicity	10	LCP2;B2M;CD74;TMSB4X;VIM;ITGB2;LY9;ANXA1;TLR2;IFI30	0.00406424

microglial activation	10	B2M;CD74;PKM;MSN;TMSB4X;VIM;ITGB2;NR4A2;ANXA1;TLR2	0.004249746
cell motion	10	CD74;PKM;MSN;TMSB4X;VIM;CD63;STMN1;CAPN2;ANXA1;TLR2	0.006115691
cell regeneration	10	DDB2;B2M;CD74;LRP5;PKM;TMSB4X;VIM;CAPN2;ANXA1;TLR2	0.006470572
adaptive immune response	10	B2M;CD74;RPS6KA1;VIM;ITGB2;NR4A2;TNFAIP8;ANXA1;TLR2;IFI30	0.008578196
integrin activation	9	LCP2;DAPK1;MSN;VIM;ITGB2;CD63;CAPN2;ANXA1;TLR2	2.93595E-05
platelet aggregation	9	LCP2;B2M;MSN;VIM;ITGB2;CD63;COTL1;ANXA1;TLR2	0.002966296
superoxide anion generation	9	LCP2;B2M;RPS6KA1;TMSB4X;VIM;ITGB2;ANXA1;TLR2;IFI30	0.005016428
cell recognition	8	B2M;CD74;PKM;MSN;ITGB2;ANXA1;TLR2;IFI30	0.000240928
cell fusion	8	MSN;ITGB2;CD63;STMN1;NR4A2;CAPN2;ANXA1;TLR2	0.001226804
leukocyte recruitment	8	LCP2;CD74;MSN;TMSB4X;ITGB2;CD63;ANXA1;TLR2	0.001606086
actin nucleation	8	LCP2;MSN;TMSB4X;VIM;ITGB2;CD63;CAPN2;ANXA1	0.001866016
drug susceptibility	8	CD74;PKM;DAPK1;TMSB4X;VIM;STMN1;LY9;TLR2	0.005909105
microtubule polymerization	8	DAPK1;PRC1;MSN;TMSB4X;TUBA4A;VIM;MAT2B;STMN1	0.006061974
neuronal plasticity	8	B2M;CD74;PKM;VIM;STMN1;NR4A2;CAPN2;TLR2	0.006417058
lymphocyte activation	7	LCP2;B2M;CD74;RPS6KA1;ITGB2;LY9;TLR2	0.000569666
mast cell degranulation	7	LCP2;PKM;VIM;ITGB2;CD63;ANXA1;TLR2	0.00084979
regulatory T-cell function	7	B2M;MSN;VIM;ITGB2;NR4A2;ANXA1;TLR2	0.001715809
focal adhesion assembly	7	DAPK1;MSN;TMSB4X;VIM;STMN1;CAPN2;ANXA1	0.001749493
intracellular transport	7	B2M;CD74;PKM;VIM;CD63;STMN1;NPC2	0.00229865
ubiquitin-dependent protein degradation	7	DDB2;PKM;DAPK1;VIM;NR4A2;CAPN2;ASB8	0.002947869
T cell migration	7	CD74;DAPK1;MSN;ITGB2;COTL1;CAPN2;TLR2	0.00414392
phagocyte activity	7	B2M;PKM;VIM;ITGB2;CAPN2;ANXA1;TLR2	0.004663544
cell phagocytosis	7	B2M;LRP5;TMSB4X;ITGB2;STMN1;ANXA1;TLR2	0.005431894
leukocyte cell adhesion	7	LCP2;MSN;TMSB4X;VIM;ITGB2;ANXA1;TLR2	0.005722294
antigen processing and presentation of exogenous peptide antigen via MHC class I	6	B2M;CD74;ITGB2;ANXA1;TLR2;IFI30	0.000678167
cell-matrix adhesion	6	PKM;DAPK1;MSN;VIM;ITGB2;ANXA1	0.001366325
platelet adhesion	6	VIM;ITGB2;CD63;COTL1;ANXA1;TLR2	0.00143567

ruffle assembly	6	RPS6KA1;MSN;VIM;STMN1;INPP5K;ANXA1	0.00298624
T-helper 1 cell differentiation	6	PKM;VIM;NR4A2;LY9;ANXA1;TLR2	0.003761296
dendritic cell migration	6	LCP2;CD74;ITGB2;CD63;TLR2;ARHGEF5	0.003986152
regulatory T-cell differentiation	6	PKM;MSN;ITGB2;NR4A2;LY9;TLR2	0.004342059
immunological surveillance	6	B2M;CD74;DAPK1;ITGB2;TLR2;IFI30	0.008097566
granulocyte transendothelial migration	5	VIM;ITGB2;CAPN2;ANXA1;TLR2	4.09643E-05
granulocyte migration	5	CD74;TMSB4X;ITGB2;ANXA1;TLR2	0.000430239
pseudopodium formation	5	MSN;TMSB4X;VIM;ITGB2;CAPN2	0.000782392
phagosome maturation	5	SNX6;MSN;ANXA1;TLR2;STX7	0.001425981
RNA replication	5	PKM;VIM;CD63;CAPN2;ANXA1	0.004212643
T-cell adhesion	5	LCP2;B2M;MSN;VIM;ITGB2	0.004262846
T-helper 17 cell differentiation	5	PKM;DAPK1;RPS6KA1;NR4A2;TLR2	0.006574824
neutrophil apoptosis	5	DAPK1;VIM;ITGB2;ANXA1;TLR2	0.008211283
suppressor T-cell differentiation	4	MSN;NR4A2;LY9;TLR2	4.24193E-05
neutrophil rolling	4	MSN;ITGB2;ANXA1;TLR2	0.000245553
lymphocyte transendothelial migration	4	MSN;TMSB4X;VIM;ITGB2	0.000375598
T-cell recognition	4	B2M;CD74;TLR2;IFI30	0.000731415
myeloid cell migration	4	FGR;ITGB2;ANXA1;TLR2	0.001601918
neutrophil survival	4	ITGB2;NR4A2;ANXA1;TLR2	0.002605355
membrane organization and biogenesis	4	MSN;CD63;ANXA1;NPC2	0.003128199
leukocyte tethering / rolling	4	LCP2;MSN;ITGB2;TLR2	0.006650308
eosinophil accumulation	4	FGR;ITGB2;ANXA1;TLR2	0.00674335
lymphocyte adhesion	4	LCP2;MSN;VIM;ITGB2	0.006837256
establishment of T-cell polarity	4	LCP2;ITGB2;TNFAIP8;TLR2	0.007221578
autophagosome maturation	4	RPS6KA1;VIM;ANXA1;STX7	0.008244459
granulocyte motility	3	DAPK1;ITGB2;ANXA1	2.80555E-05
endosome localization	3	PLEKHB2;VIM;ANXA1	0.000108092
neutrophil diapedesis	3	ITGB2;ANXA1;TLR2	0.000480174

monocyte extravasation	3	VIM;ITGB2;ANXA1	0.001068117
tissue migration	3	B2M;ITGB2;ANXA1	0.002050196
monocyte phenotype	3	B2M;CD74;ITGB2	0.003364767
CD8+ T-cell priming	3	B2M;CD74;IFI30	0.004979354
wound contraction	3	TMSB4X;VIM;ANXA1	0.00627584
granulocyte infiltration	3	TMSB4X;ITGB2;ANXA1	0.007141832
monocyte apoptosis	3	B2M;ANXA1;TLR2	0.007600147
thymic T-cell selection	3	B2M;CD74;NPC2	0.007915173
liver changes	3	PKM;TMSB4X;NPC2	0.008568181
alveolar macrophage chemotaxis	2	ITGB2;TLR2	0.000184477
lymphocyte costimulation	2	LCP2;ITGB2	0.000908353
granulocyte spreading	2	LCP2;ITGB2	0.000908353
effector T-cell adhesion	2	LCP2;ITGB2	0.001316022
myeloid dendritic cell activation	2	PKM;TLR2	0.002155457
phagocyte recognition	2	ITGB2;TLR2	0.002753038
dendritic cell recognition	2	ITGB2;TLR2	0.002967606
eosinophil homing	2	ITGB2;TLR2	0.00341953
monocyte spreading	2	ITGB2;CD63	0.003656794
attachment to substrate	2	TMSB4X;ITGB2	0.003901532
thymocyte expansion	2	LCP2;HELLS	0.003901532
conventional dendritic cell function	2	PKM;TLR2	0.004153699
activated T-cell adhesion	2	MSN;ITGB2	0.004954328
peripheral nerve function	2	TMSB4X;TLR2	0.006750172
hematopoietic cell count	2	CD74;IFI30	0.006750172
endothelial cell transendothelial migration	2	ITGB2;ANXA1	0.007743029
CD4+ T-cell proliferative response	2	TMSB4X;ANXA1	0.008087759
neutrophil lifespan	2	ITGB2;NR4A2	0.008087759

CHAPTER 4 - Alternatively spliced genes in hepatic tissue associated with RFI in two Nelore cattle genetics group

ABSTRACT – RNA-Seq data were used to identify differentially expressed alternatively spliced transcripts in Nelore cattle liver tissue from two genetic groups with residual feed intake measured. A total of 67 differentially expressed alternatively spliced transcripts (DAS genes) in genetic group 1 (G1), and 75 DAS genes in genetic group 2 (G2) ($p \leq 0.05$) in liver between the highest and lowest RFI groups were identified from the exon-centric approach. Complete and exact match of intron chain was the most frequently found alternative splicing event were the most frequently found for two genetics groups. For both genetics' groups, nine common differentially expressed genes were predicted to produce alternatively spliced isoforms between LRFI *vs* HRFI. In general, genes found for both traits are involved in biological process and pathways related to lipid metabolism and immune system. Eight DAS detected were considered biomarkers (AUC > 0.75). Five of them were potential nodes in the physical interactions analysis and six were as common nodes in the co-expression analysis, highlighting *SAT1* as hub gene. The *SAT1* play a role in innate immune response and lipid metabolism. The results reinforce the biological importance of these known processes but also reveal new insights into the complexity of the liver tissue transcriptome of Nelore cattle.

Keywords: bovine, genes, Hub gene, nodes, RFI, splicing alternative

1. Background

The cellular functions are underpinned by the expression of protein-coding and non-coding genes, termed the transcriptome (ROBERT; WATSON, 2016). The protein-coding genes express multiple RNA isoforms via processes such as alternative transcriptional start sites, termination sites and splicing, greatly increasing the diversity of the transcriptome and proteome within cells (PAN et al., 2008; WANG et al., 2008). Gene regulation through alternative splicing can create different functions for a single gene and increase protein-coding capacity (ROBERT; WATSON, 2015; TUNG; SHAO; KINGSFORD, 2019).

The alternative processing of mRNAs (messenger RNAs) produces multiple mRNA and protein isoforms that may have similar or quite different functions (FANG et al., 2021). In bovines, isoforms and alternative splicing were studied in numerous tissues as muscle (FANG et al., 2021), fatty acid composition in various organs, included liver (SUN et al., 2021), in carcass composition and affecting meat quality (SILVA et al., 2020; MUNIZ et al., 2021). Although, studies aiming to find alternative splicing events for RFI trait differences in cattle remains limited, particularly for Nelore cattle.

The aim of this study was to utilize RNA-Seq to identify candidate genes that are differentially expressed as alternatively spliced isoforms in the liver tissue of Nelore animals that differed for RFI in two genetics groups.

2. Material and methods

To animals' phenotypes description, RNA-Seq library preparation, sequences analysis and aligner, see material and methods in the Serna-García et al., (chapter 2).

2.1. Identification of differentially expressed alternatively spliced transcripts (DAS)

Based on splice junction reads predicted using the Cufflinks2 (v.2.1.1) suite of tools (Cufflinks2, and Cuffmerge2) (TRAPNELL et al., 2012), the JuncBASE (v.0.9) software (Junction-Based Analysis of Splicing Events) (LERCH et al., 2012) was used to identify and classify exon-centric alternative splicing events. For each RNA-Seq sample, the spliced alignment is given as input, BAM files aligned with HISAT2 mapper against the Bos taurus assembly (ARS-UCD1.2). Splice junctions passing an entropy score threshold were combined with exon coordinates from transcript reference annotations and optional novel

transcripts to identify alternative splicing events. Only concordant new possible transcripts among the samples analyzed were selected from the GTF files obtained, before building the normalized table used in the differential expression analysis EdgeR. For more details, the following six steps were used.

Step 1 Building annotation database: the annotation database was used to identify all exons in the transcriptome, and it was derived from a GTF file. The first, was GTF file that combined Cufflinks2 and Cuffmerge2 identified transcripts with those that are annotated using ARS-UCD1.2 (genome *Bos taurus*) and it was used to identify all internal exons (not the first or last exon in a transcript).

Step 2 Determination of junction entropy cutoff: for all subset of samples, was built an entropy plot of annotated and novel junctions. Cut off of 1.1 was used to reduce the number of potential false positive novel junctions. The new transcripts found in all samples were selected for further analysis.

Step 3 Identification of all junctions: this step created one BED file (that combined junctions from all samples.

Step 4 Building of exon–intron junction count files: this step created a counts file of the number of reads aligning to every exon–intron junction.

Step 5 Alternative splicing events identification and quantification of the events from each sample: this step identified, classified, and quantified alternative splicing events. The collapsed isoform GTF files were compared to annotated transcripts using gffcompare (v0.11.2) (TILGNER et al., 2014). GffCompare is a generic, standalone tool for merging and tracking transcript structures across multiple samples and comparing them to a reference annotation (TUNG; SHAO; KINGSFORD, 2019).

Gffcompare assigns each transcript a class code based on how it compares to annotated reference transcripts. The class codes were grouped into the following categories. Known: full (=) or partial (c) intron chain match to annotated transcript isoform. Novel isoforms of known genes: novel splice isoform (j), splice chain match with additional terminal exon(s) (k), retained intron isoforms (m,n). Novel transcripts potentially representing novel genes or transcriptional units: intronic transcript (i), antisense transcript (x), intergenic transcript (u), 5′ distal overlapping RNA (y), other exonic overlap (o). Other: including RNA fragments and potential sequencing artifacts (p,e,s,r) (GLEESON et al., 2022) (Figure 1).

Step 6 Differential splicing analysis: this last step identified differentially spliced events between two specified groups (Highest and Lowest) of animals (RFI) in two genetics groups. The program was used to estimation of normalized

reads and identify of differentially expressed alternatively spliced transcripts between RFI groups that was performed using EdgeR software (OSHLACK; ROBINSON; YOUNG, 2010; SMYTH et al., 2016). Only alternatively spliced transcripts that were supported by at least 25 mean reads per sample between the respective high and low groups of greater than 5% were considered significant.

Figure 1. Transcript classification codes based on their relationship to reference transcripts, asgenerated by GffCompare. Reference exons and transcripts are shown in black, transcripts to be classified are shown in blue, and hashed regions represent repeated regions in the genome. For example, the transcript in blue on the uppermost left panel is labeled “=” because all of its introns precisely match the annotation in black (PERTEA; PERTEA, 2020).

2.2. Gene set enrichment analyses

The gene sets in which alternatively spliced transcripts and genes set DAS transcripts associated with RFI were identified, and then the enrichment and pathway analyses of these gene sets were performed using the GSEA software (SUBRAMANIAN et al., 2005) using MSigDB database v7.5.

The enrichment and pathway analyses of genes set DAS transcripts associated with RFI were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID v.6.8) (HUANG; SHERMAN; LEMPICKI, 2009) Fisher’s exact test was used and p values were adjusted for

multiple testing using the Benjamini–Hochberg method (BENJAMINI; HOCHBERG, 1995) considering significantly values of less than 5%. DAVID Pathway was used to map the gene enriched pathways using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (KANEHISA; GOTO, 2000).

Pathway Studio software v.10 (Elsevier Inc., Rockville, MD, USA) was used to identify the main biological processes and pathways of genes set DAS transcripts associated with RFI with database ResNet v11.

2.3. Co-expression networks analyses

For co-expression network analyses, the ExpressionCorrelation plugin (HUI et al., 2015) within Cytoscape v.3.4 (SHANNON et al., 2003) was used to identify consensus networks of gene sets in which differentially expressed alternatively spliced transcripts were associated with RFI. Similarity matrices were computed using the Pearson correlations among CPMs values.

A histogram tool was used for the screening criteria at node score cut-offs > 0.77 and < − 0.77 and employed to identify statistical significance of the pairwise correlations. We established a co-expression network between the significantly co-expressed DAS genes. The Network Analyzer plugin (ASSENOV et al., 2008) in Cytoscape software, regulatory relationships among various genes were identified by topological properties of the network such as degree distribution, betweenness centrality, network density, and clustering coefficient (BARABÁSI; OLTVAI, 2004; DONG; HORVATH, 2007).

The cytoHubba plugin (CHIN et al., 2014) was used to explore co-expression network nodes (hub genes). This plugin provides a user-friendly interface to explore important nodes in biological networks and computes using eleven methods (degree, edge percolated component, maximum neighborhood component, density of maximum neighborhood component, maximal clique centrality -MCC and six centralities - bottleneck, EcCentricity, closeness, radiality, betweenness, and stress). All these parameters can indicate the robustness of the analysis (STEVENS et al., 2013; CHIN et al., 2014). The maximal clique centrality (MCC) has a better performance on the precision of predicting essential genes from the co-expression network (CHIN et al., 2014) and to generate a subnetwork (STEVENS et al., 2013).

2.4. Prediction of footprint gene profile

To obtain RFI predictors or biomarkers receiver operating characteristic curves (ROC) were generated. For analysis used genefilters package (Bioconductor) with the function "rowpAUCs". The function has applied on the normalized values (counts) in two genetic groups (independent). Statistical correction (FDR) is not applied. Rowwise calculation of Receiver Operating Characteristic (ROC) curves and the corresponding partial area under the curve (pAUC) for a given data matrix or ExpressionSet. Cutpoints (theta are calculated before the first, in between and after the last data value. By default, both classification rules x > theta and x < theta are tested and the (partial) area under the curve of the better one of the two i returned. This is only valid for symmetric cases, where the classification is independent of the magnitude of x (e.g., both over- and under-expression of different genes in the same class).

An AUC (Area under the ROC Curve) of 0.5 represents a test with no discriminating ability (ie, no better than chance), while an AUC of 1.0 represents a test with perfect discrimination (HOO; CANDLISH; TEARE, 2017). AUC of the proposed test was estimated to be >0.75.

3. Results

3.1. Differentially expressed alternatively spliced transcripts

There were found 768 alternatively spliced transcripts produced from 748 genes, and 67 genes that were differentially expressed alternatively spliced transcripts (DAS) (p-value ≤ 0.05) between the LRFI and HRFI groups in genetic group 1 (G1) and 676 alternatively spliced transcripts produced from 661 genes , and 75 DAS (p-value ≤ 0.05) in genetic group 2 (G2) (the comprehensive list of the differentially expressed alternatively spliced transcripts associated with RFI in G1 is provided in Supplementary Tables: Table S1) and in G2 (Supplementary Tables: Table S2).

For RFI in G1 and G2, complete, exact match of intron chain was the most frequently found alternative splicing event (Figure 2). Of the differentially expressed alternatively spliced transcripts, 55 were known and 14 were novel transcripts in G1. For G2 group, 65 were known and 13 were novel transcripts (Table S1).

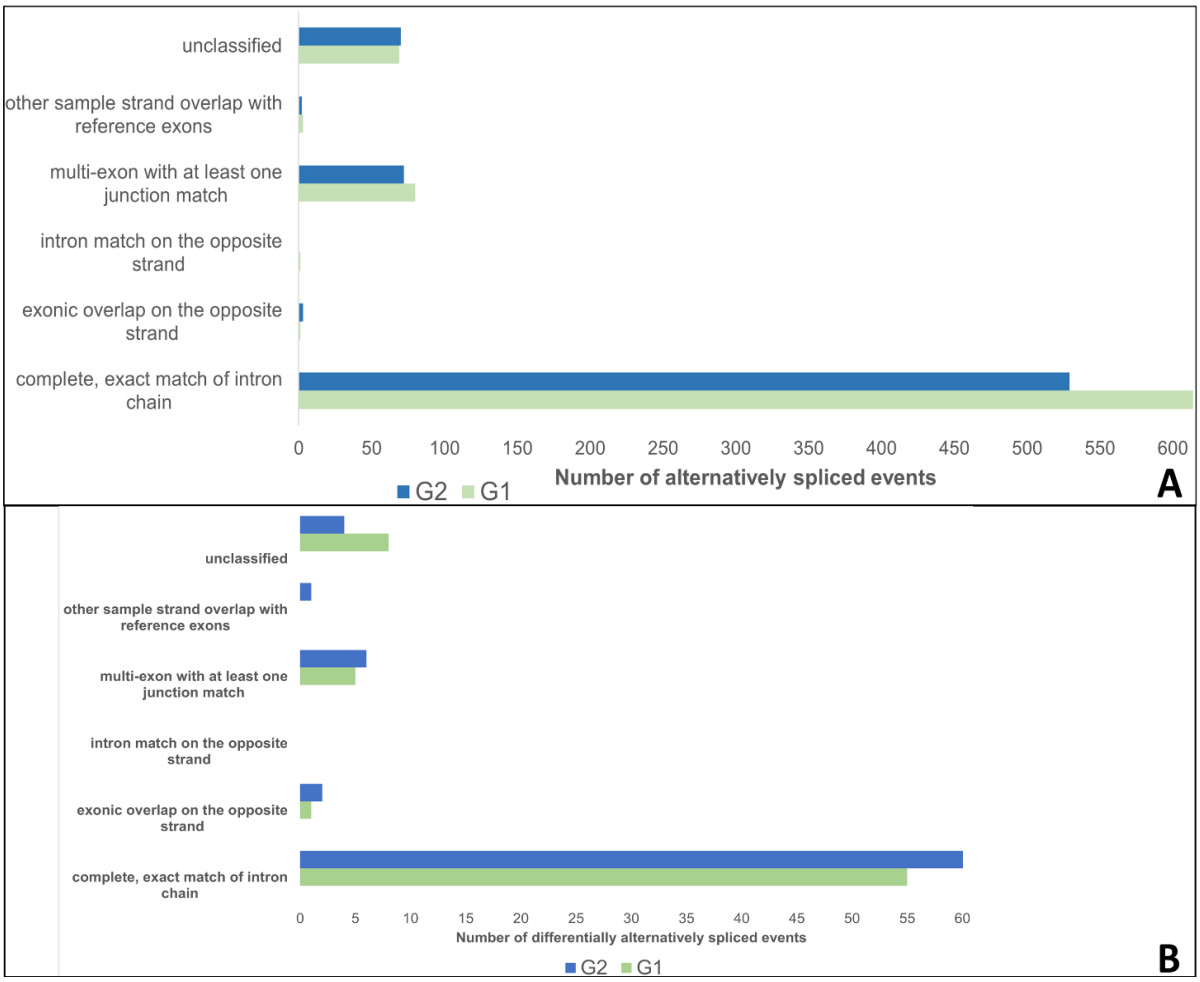


Figure 2. Transcript classification based on their relationship to reference transcripts, as generated by GffCompare. Transcripts identified in the RFI of Nelore cattle with the lowest and highest RFI in G1 (genetic group 1) (green color) and genetic group 2 (genetic group 2) (blue color). **A.** Alternatively spliced transcripts identified in G1. **B.** Alternatively spliced and differentially expressed ($p < 0.05$) transcripts identified in G2.

For genetic group 1, some differentially expressed alternatively spliced transcripts belong to important gene families, SERPINA down-regulated in LRFI (serine protease inhibitors or classified inhibitor family I4) such as *SERPINA3-8* and *SERPINA3-7*, It is predominantly expressed in the liver and secreted into plasma. Serpins are a broadly distributed family of protease inhibitors that use a conformational change to inhibit target enzymes (LAW et al., 2006). The transcript isoforms produced by these gene families differ by complete, exact match of intron chain (see Supplementary Tables: Table S1).

Other genes were down-regulated in LRFI vs HRFI such as inter-alpha-trypsin inhibitors (ITI) i.e *ITIH2* and *ITIH3*, which are a family of plasma protease inhibitors (HAMM et al., 2008). We showed *HMGCL* gene was overexpressed (3-(Hydroxy-3-Methylglutaryl-CoA Lyase). The transcription factor *HMGN2* was overexpressed in LRFI. It modulates chromatin structure in a variety of pathways (SUBRAMANIAN et al., 2009; GARZA-MANERO et al., 2019). Its main role is in global DNA repair and maintenance of genome integrity.

For genetic group 2, some differentially expressed alternatively spliced transcripts belong to important gene families as ATP synthase subunit 5, mitochondrial (*ATP5*) up-regulated in LRFI such as (*ATP5F1D, ATP5MD, ATP5ME, ATP5MPL, ATP5PB*). Other gene family overexpressed in LRFI were Cytochrome c oxidase (*COX*). It is the terminal enzyme of the mitochondrial respiratory chain such as (*COX4I1, COX6A1, COX6B1, COX7C*). Other gene family up-regulated in LRFI was *H3F3A* and *H3F3B* up-regulated in LRFI genes encode core histone proteins. (ALDERA; GOVENDER, 2022). Furthermore, we showed numerous genes up-regulated in LRFI related to the RPs family, which are involved in assembly of either the small 40S (RPS) or the large 60S ribosomal subunits (RPL) (PROVOST; WEIER; LEACH, 2013). The ribosomal protein R (RPS); *RPS13, RPS14, RPS16, RPS17. RPS21, RPS25, RPS28, RPS29, RPS9, RPSA* and *RPL18A* and Ribosomal protein L (RPL); *RPL27, RPL28, RPL29, RPL34* and *RPL39*) were detected. This family was found also in G1 such as *RPS12* and *RPSL23*.

3.2. Differentially expressed alternatively spliced transcripts common

The differentially expressed alternatively spliced genes from both genetics groups between LRFI and HRFI animals were eleven (Table 1). We observed that most of them (nine) have the same biological orientation in G1 and G2 (up-regulated or down-regulated).

Table 1. The eleven common genes set DAS transcripts from both genetics groups (G1 and G2) between LRFI vs HRFI.

		G1			G2		
Gene name	**Gene Symbol**	**Fold change (log2***	**p-value**	**padj**	**Fold change (log2)***	**p-value**	**padj**
Beta-2-microglobulin	*B2M*	0.25	0.005	0.541	0.24	0.013	0.270
Major histocompatibility complex, class II, DR alpha	*BOLA-DRA*	0.46	0.006	0.541	0.34	0.022	0.301
Complement C1q C Chain	*C1QC*	0.43	0.037	0.544	0.41	0.003	0.165
CD74 molecule	*CD74*	0.46	0.006	0.541	0.31	0.041	0.386
H3 Histone, Family 3B	*H3F3B*	0.19	0.016	0.541	0.19	0.025	0.329
High Mobility Group Nucleosomal Binding Domain 2	*HMGN2*	0.19	0.028	0.544	0.18	0.020	0.300
NPC Intracellular Cholesterol Transporter 2	*NPC2*	0.31	0.004	0.541	0.37	0.001	0.158
Spermidine/Spermine N1-Acetyltransferase 1	*SAT1*	0.24	0.006	0.541	0.24	0.017	0.283
Spondin 2	*SPON2*	-0.26	0.025	0.541	0.24	0.002	0.158
Transcobalamin 2	*TCN2*	-0.22	0.026	0.544	0.25	0.005	0.165
Thymosin Beta 10	*TMSB10*	0.34	0.024	0.541	0.33	0.004	0.165

*The fold-change estimates (relative expression) refer to the LRFI group.

The *B2M, BOLA-DRA* and *CD74* genes expressed alternatively spliced transcripts were up-regulated in LRFI group. The bovine MHC (major histocompatibility complex class II), called bovine leucocyte antigen (BoLA), that exhibits very little sequence variation, especially at the protein level (KLEIN et al., 1990). Previously, *BOLA-DRA, B2M* and *CD74* genes were detected and discussed in the gene expression analysis shown in Serna-García et al. (chapter 2 and 3).

3.3. Enrichment biological analysis

To approach the biological analysis, we followed two strategies: one using of gene sets in which alternatively spliced transcripts for each genetic group and the other using of genes set DAS transcripts.

3.3.1. Biological analysis of the alternatively spliced transcripts

Gene enrichment analysis (padj<0.05) with gene sets in which alternatively spliced transcripts showed 76 biological processes in G1 with up-regulation comparing LRFI vs HRFI in GSEA software. For G2, 54 up-regulated biological processes were detected comparing LRFI vs HRFI (Supplementary Tables: Table S4)

For RFI- selected bovine in G1 the Gene Ontology (GO) terms of biological process for the gene sets in which differentially expressed alternatively spliced transcripts immune system were identified, for example, immune system process (GO:0002376), immune response (GO:0006955). Organization of the cytoskeleton: actin cytoskeleton organization (GO:0030036), actin filament organization (GO:0007015) and protein synthesis: cellular protein metabolic process (GO:0044267), peptide biosynthetic process GO:0043043. For RFI-selected bovine in G2 the Gene Ontology (GO) terms of biological process process; SRP-dependent cotranslational protein targeting to membrane (GO:0006614), peptide metabolic process (GO:0006518), protein targeting to ER (GO:0045047), amide biosynthetic process (GO:0043604), ribosome biogenesis (GO:0042254), ribosomal small subunit biogenesis (GO:0042274).

We highlight common biological processes such as peptide biosynthetic process (GO:0043043), mRNA catabolic process (GO:0006402), cellular

nitrogen compound catabolic process (GO:0044270), cellular nitrogen compound biosynthetic process (GO:0044271) and nucleic acid metabolic process (GO:0090304).

3.3.2. The enrichment analyses of genes set DAS transcripts

Gene sets in which alternatively spliced transcripts are differentially expressed (67 for G1 and 75 for G2) were used for analysis in GSEA software. The gene enrichment analysis showed 11 biological processes in genetic group 1 (10 up-regulated and 1 down-regulated) comparing LRFI vs HRFI. For genetic group 2, 27 biological processes were detected (18 up-regulated and 9 down-regulated) comparing LRFI vs HRFI (Supplementary Tables: Table S3A).

For bovine selected on RFI in G1 Gene Ontology (GO) terms of the biological process were identified for the gene sets in which alternatively spliced transcripts were differentially expressed immune system, for example, myeloid leukocyte activation (GO:0002274) and neutrophil mediated immunity (GO:0002446). In genetic group 1 and 2 found for RLFI animals in G1 were associated immune effector process (GO:0002252). In G1 SRP was found to be a cytosolic particle that transiently binds to the endoplasmic reticulum (ER) signal sequence in a nascent protein, the large ribosomal unit, and the SRP receptor on the ER membrane. SRP-dependent cotranslational protein targeting to membrane (GO:0006614), protein targeting to ER (GO:0045047), protein localization to endoplasmic reticulum (GO:0070972).

Gene enrichment analysis with DAVID software (Supplementary Tables: Table S3B) for RFI selected bovine in G1 (GO) in terms of the biological processes was performed and immune system was identified: innate immune response (GO:0045087), antigen processing and presentation of exogenous protein antigen via MHC class Ib (GO:0002481), positive regulation of T cell cytokine production (GO:0002726), MHC class I protein complex (GO:0042612). Furthermore, we identified processes associated with arginine such as arginine biosynthetic process via ornithine (GO:0042450) and arginine biosynthetic process (GO:0006526). The most biological terms found for animals selected to differ for RFI in G2 were associated with were identified immune system and Respiratory chain, for example, positive regulation of macrophage cytokine production (GO:0060907), intracellular ribonucleoprotein complex (GO:0030529)

, ubiquinol-cytochrome-c reductase activity (GO:0008121), cytochrome-c oxidase activity (GO:0004129), oxidative phosphorylation, mitochondrial electron transport, cytochrome c to oxygen (GO:0006123), respiratory chain complex IV (GO:0045277), ribosomal small subunit assembly (GO:0000028), mitochondrial respiratory chain complex III (GO:0005750).

In the Figure 3, we showed the main cellular processes for nine common genes set DAS transcripts from both genetics groups (G1 and G2) between LRFI vs HRFI. The enrichment biological processes identified immune system terms such as: MHC class II protein complex (GO:0042613) and antigen processing and presentation (bta04612) (Supplementary Tables: Table S5).

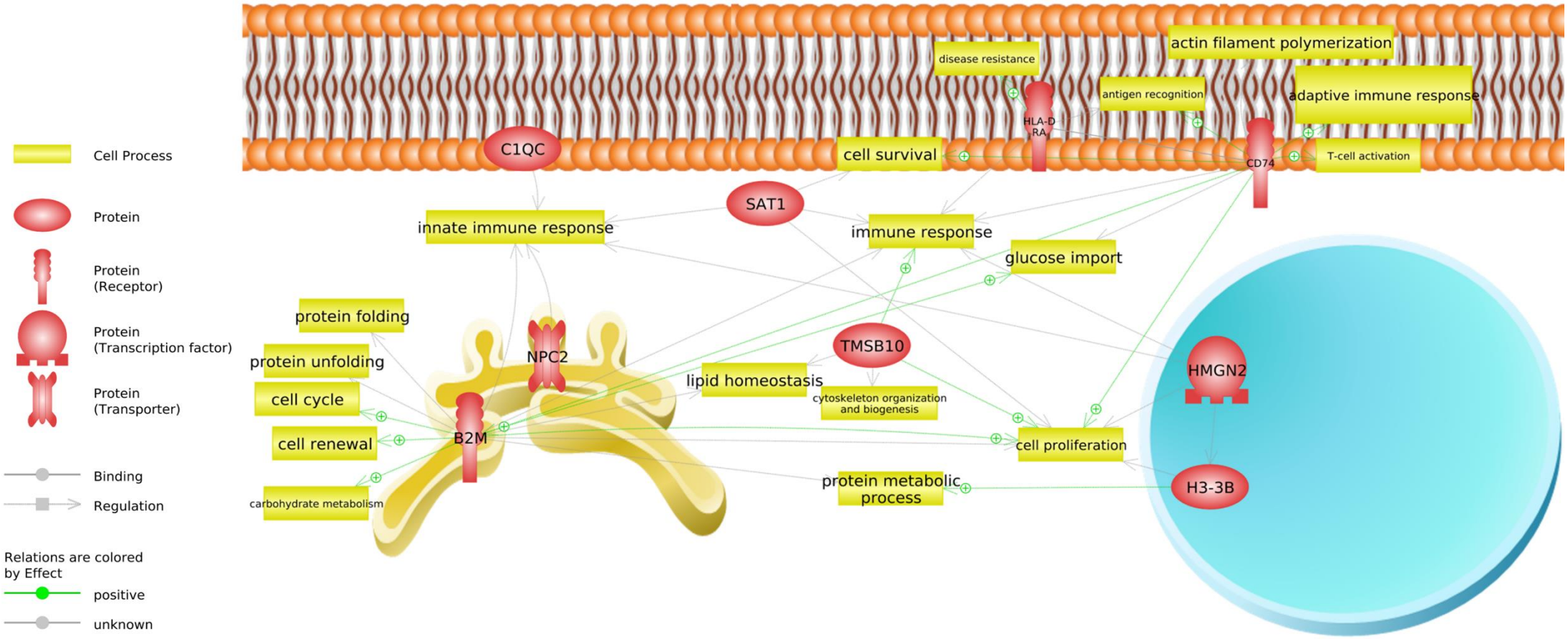


Figure 3. The main cellular processes for the common genes set DAS transcripts associated to RFI trait using Pathway studio. The genes highlighted in red are genes upregulated in LRFI, the blue circles present the downregulated genes in LRFI group compared with HRFI; green arrows indicate a positive regulation effect, while red arrows have a negative regulation effect.

3.4. Gene co-expression network analysis and prediction of hub genes

We used the correlation matrix to construct the co-expression network (Figure 4). In G1, 65 significantly correlated gene pairs (edges) were discovered for 26 genes (nodes). The density and the clustering coefficient were 0.102 and 0.224, respectively. For G2, there were 76 significantly correlated gene pairs (edges) involving 41 genes (nodes). The density and the clustering coefficient were 0.047 and 0.228, respectively. Networks with clustering and density values close to 1 may contain many edges and nodes connected (DONG; HORVATH, 2007). We showed 6 common nodes: *B2M, TMSB10, SAT1, BOLA-DRA, NPC2* and *C1QC* (Figure 2, yellow color).

When network maximal clique centrality (MCC) was examined, the genes with the highest MCC score were considered hub genes. In this study, ten top hub genes for G1 with the largest numbers of interactions were *ARPC2, AIF1, TMSB10, LOC783680, B2M, SAT1, ACTG1, BOLA-DRA, GNAI2 and CD74*. The top ten hub genes for G2 with the largest numbers of interactions were *RPL39, TMSB4X, C1QA, RPL34, ATP5ME, UQCRH, RPS16, RPS25, SAT1 and C1QC.* We considered as main hub gene to SAT1 with the highest MCC score common in G1 and G2 (Figure 4). The high interconnection between hub genes and nodes within network, have been shown to be biologically significant (LERCH et al., 2012). The hub genes found may have a putative effect on lipid metabolism in Nelore cattle.

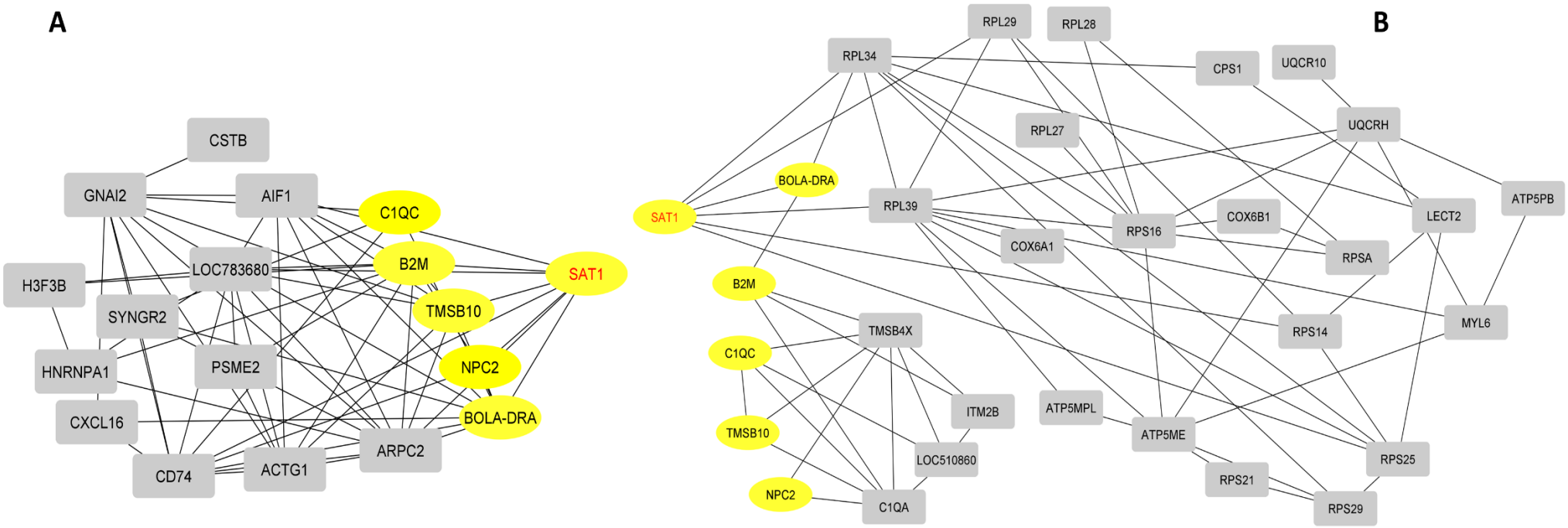


Figure 4. Transcript correlation network for the genes with differentially expressed isoforms in liver of Nelore cattle divergent for RFI. Hub gene are letters in red. **A.** Gene co-expression network analysis for G1. **B.** Gene co-expression network analysis for G2. Six common nodes are circled in yellow.

3.5. Physical interactions of nodes

We show the physical interactions between the 65 and 76 genes DAS transcripts (Table additional) with proteins from the General Biological Repository for Interaction DataSets (BioGRID). We found 37 nodes interacting with 419 potential genes (targets) for G1 and 39 nodes with their 783 potential genes (targets) for G2 (Figure 5,6,7 and 8). We highlight 5 nodes up-regulated in LRFI for G1 and G2: *HMGN2, B2M, CD74, SAT1* and *NPC2* (Table 1, Figure 8).

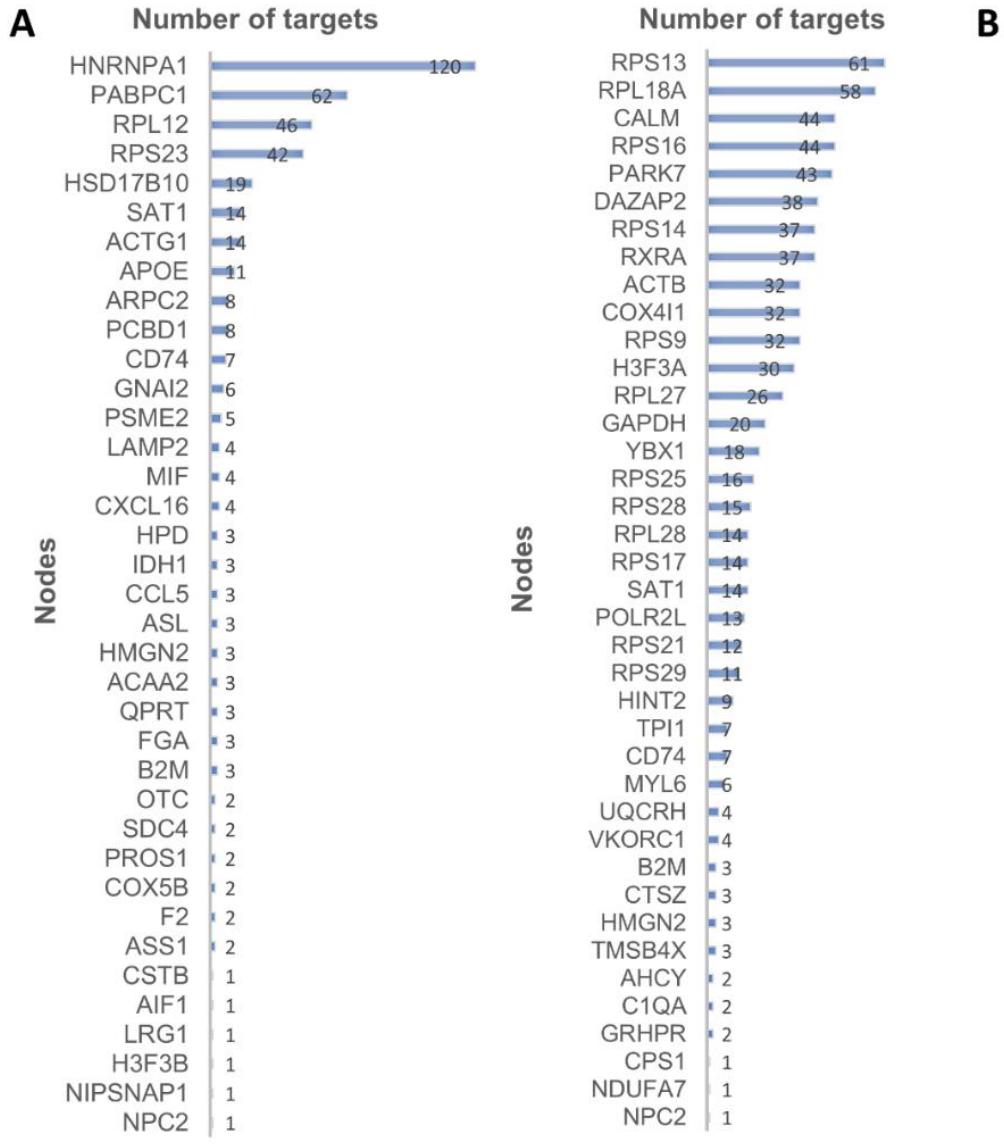


Figure 5. Transcript correlation network for the genes with differentially expressed isoforms in liver of Nelore cattle divergent for RFI. A. Analysis of the nodes of physical interactions for genetic group 1. B. Analysis of the nodes of physical interactions for genetic group 2.

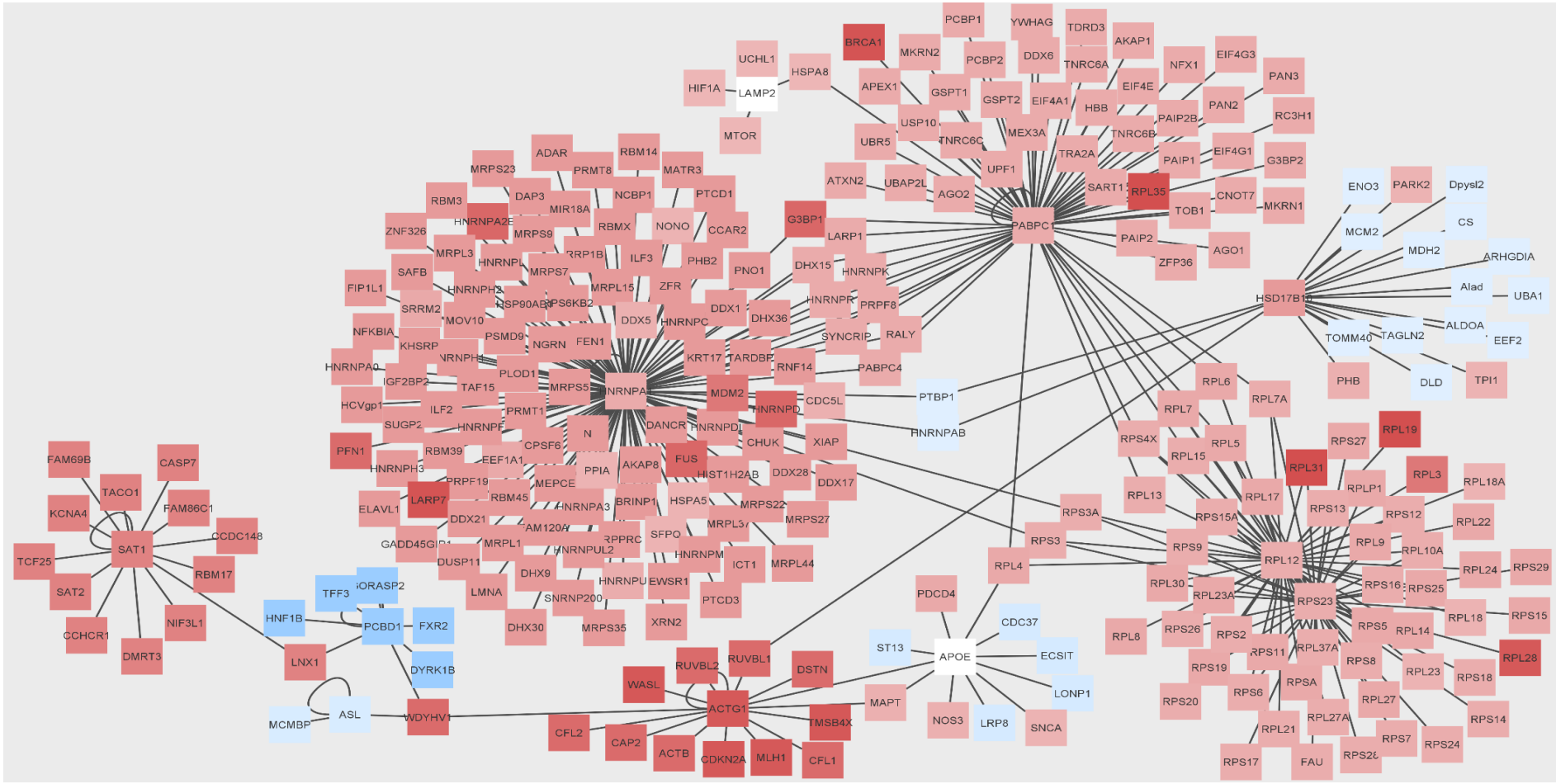


Figure 6. A. Interaction network for the genes DAS transcripts in liver of Nelore divergent for RFI by BioGRID for G1. The blue color (underexpressed in LRFI), red color (overexpressed in LRFI) and color intensity represent differing levels of expression of the genes. White color represents nodes upstream.

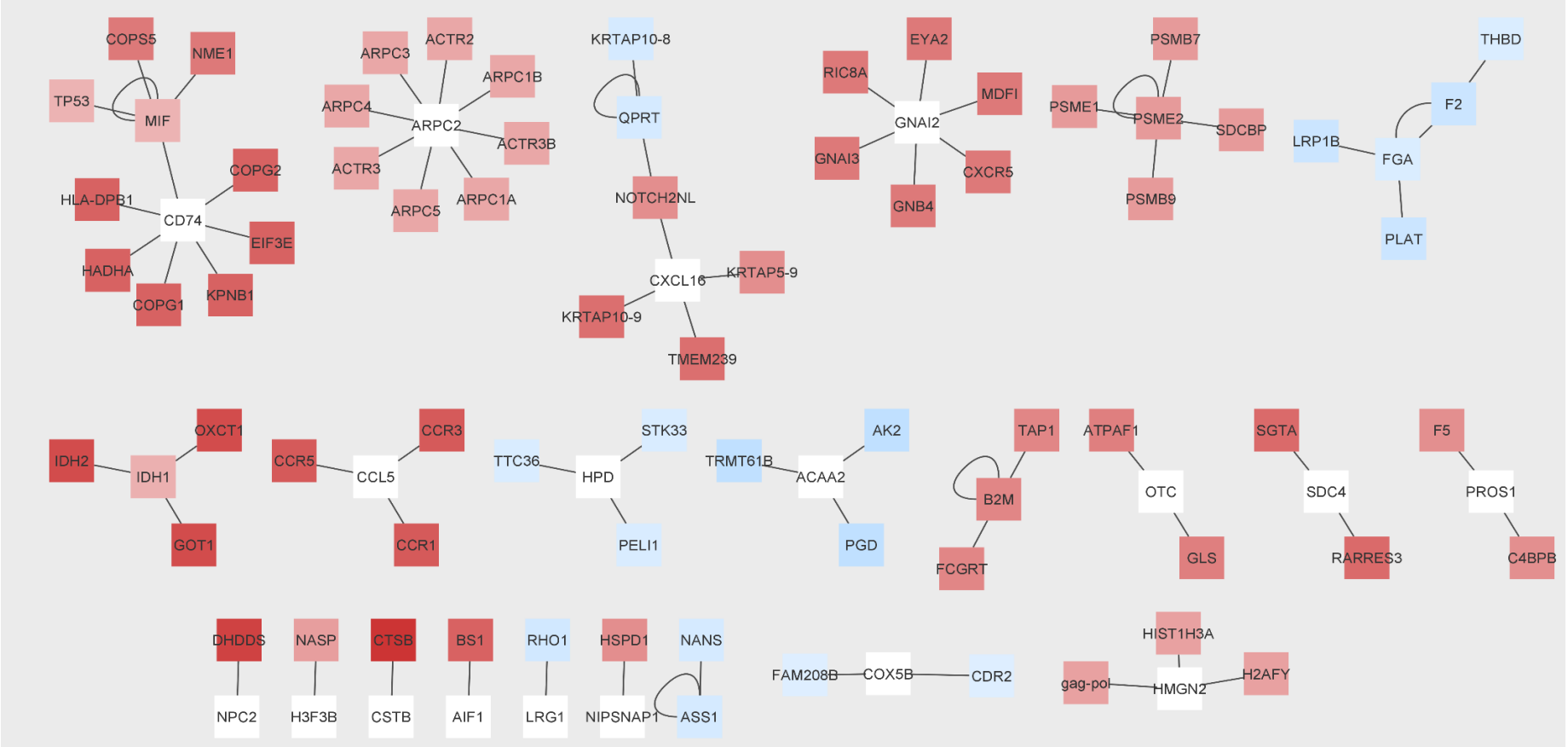


Figure 6.B. Interaction network for the genes DAS transcripts in liver of Nelore divergent for RFI by BioGRID for G1. The blue color (underexpressed in LRFI), red color (overexpressed in LRFI) and color intensity represent differing levels of expression of the genes. White color represents nodes upstream.

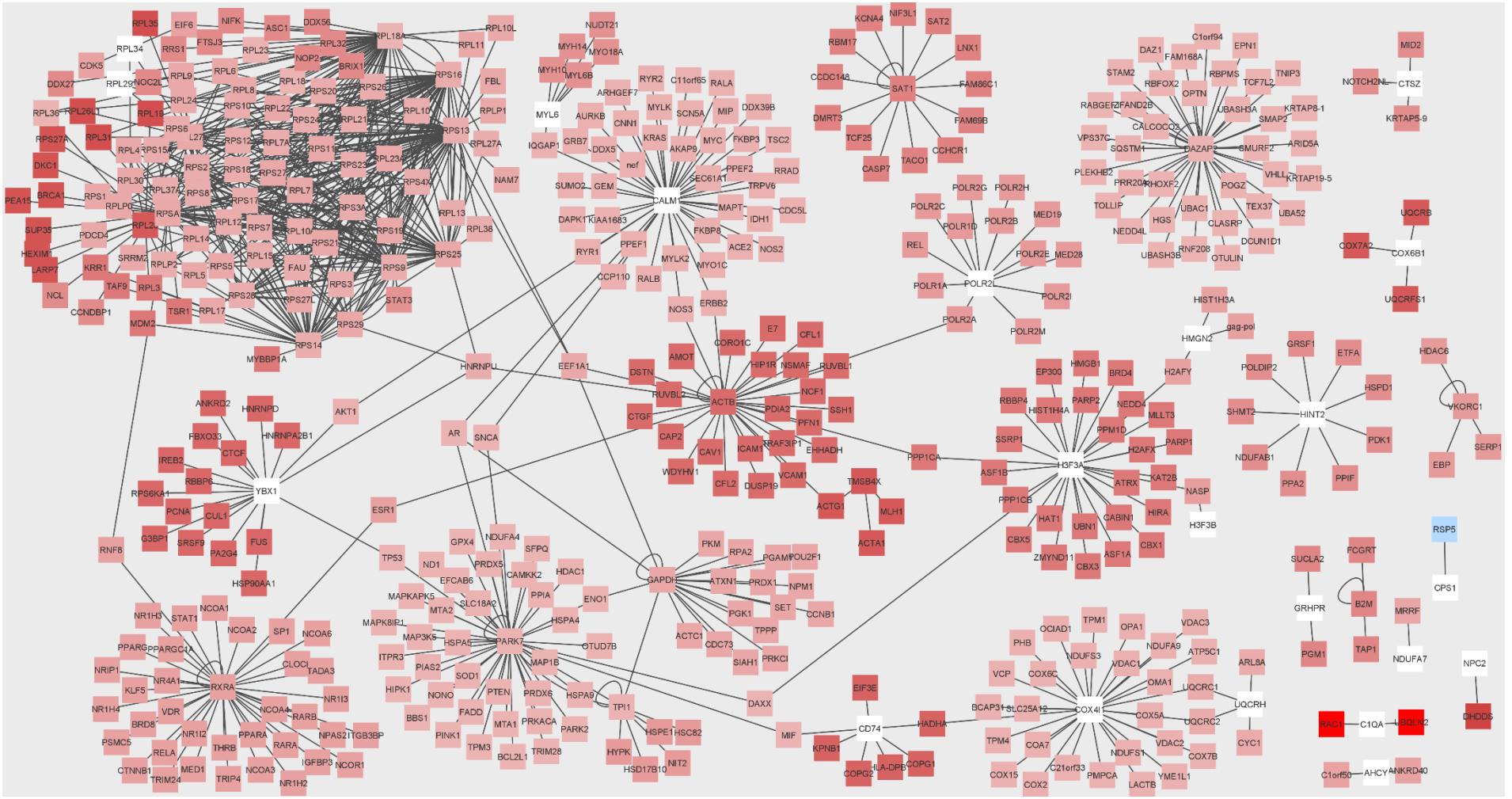


Figure 7. Transcript correlation network for the genes with differentially expressed isoforms in liver of Nelore cattle differing for RFI for G2.

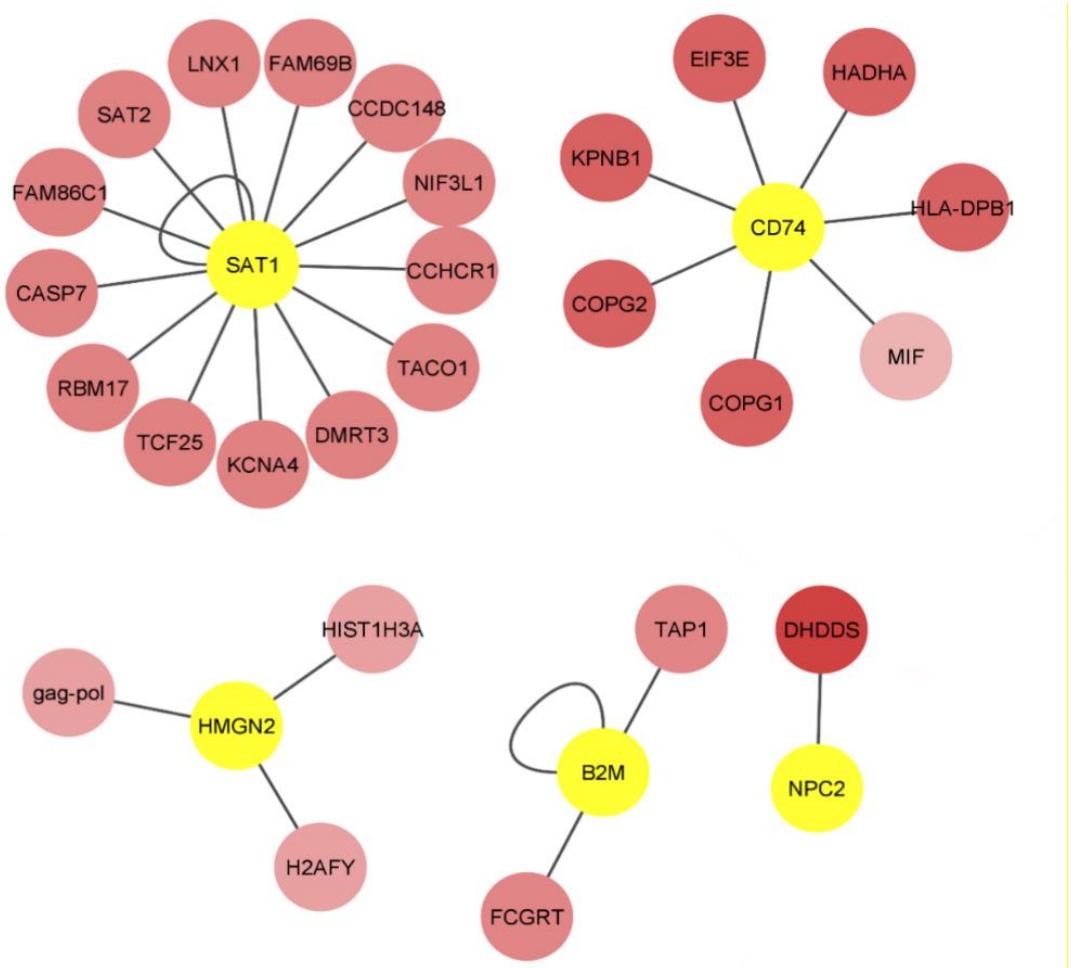


Figure 8. Interaction network for the gene DAS transcripts in liver of Nelore divergent for RFI by BioGRID. The red color (overexpressed in LRFI) and color intensity represent differing levels of expression of the genes.

3.6. Biomarkers

To find the most important biomarkers that describe our trait, we performed ROC curves using AUC >0.75. We found 10 overlapped genes in both genetics groups, G1 and G2: *B2M, BOLA-DRA, CD74, C1QC, H3F3B, NPC2, SAT1 a*nd *TMSB10,* which were integrated in Table 1. The predicted biomarkers coincide with those detected in the subsequent (S*AT1, B2M, CD74, NPC2*) (Figure 4 and 8). These biomarkers were associated with important cellular processes such as like Cell proliferation, lipid homeostasis, immune response among others (Figure 1).

Discussion

The nine common genes DAS transcripts

We have found nine common genes DAS transcripts up-regulated that produced transcripts between LRFI vs HRFI in both genetic groups (Table 1), of which eight were detected and considered biomarkers (ROC AUC > 0.75). Five were potential nodes in the physical interactions analysis and six were as common nodes in the co-expression analysis, highlighting *SAT1* as hub gene.

SAT1 is a key enzyme in the polyamine metabolic pathway gene, modulates the onset of the innate immune response and reduces the extent of apoptotic cell death (ZAHEDI et al., 2014). It regulates cell growth, proliferation and cell death (TAPPIA et al., 2019). In addition, *SAT1* overexpression enhances anti-inflammatory actions in the acute phase of the immune response (PIRNES-KARHU et al., 2012). It may play a helpful role in cell survival, it is induced by AMPK signaling and it has cardio protectant function (RYU et al., 2008). *SAT1* is involved in chromatin remodeling and regulation of gene expression (BRETT-MORRIS et al., 2014).

SAT1 mediated protective impacts involve the regulation of lipid metabolism, inflammation response, gut barrier function and thermogenesis. These findings demonstrate that spermidine has potential to treat obesity (MA et al., 2021a). SAT1 is involved in calorie intake and fatty acid metabolism (JELL et al., 2007; PIRINEN et al., 2007; KRAMER et al., 2008).

Overexpression *SAT1* had reduced white adipose tissue mass, high basal metabolic rate, improved glucose tolerance, high insulin sensitivity, coordinated by increased levels of PGC-1α and 5′-AMP-activated protein kinase (AMPK) (PIRINEN et al., 2007) and protects against obesity in mice (CASTOLDI et al., 2020), whereby lack of *SAT1* causes body fat accumulation with an increase in white adipose tissue (JELL et al., 2007; PEGG; PEGG, 2008). Therefore, this gene could be regulating lipid metabolism in LRFI.

Furthermore, *SAT1* was also found to be associated with a diamine exporter *SLC3A2* (UEMURA et al., 2008) and may enhance acetylpolyamine secretion through this exporter and potential role in acetylation of other proteins (LEE et al., 2010). However, it was found underexpressed for the LRFI group in chicken (PRAKASH et al., 2020). *SAT1* appears to be associated with *SLC3A2* (LIU et

al., 2021) and may help improve amino acid uptake and protein synthesis in the LRFI group. However, the function of this gene is not well understood.

The H3.3 histone is a replacement histone subtype that is expressed in all phases of the cell cycle and is overexpressed in transcriptionally active regions, promoter regions, and intergenic or intragenic regulatory elements. This histone encoded by *H3F3B* (AYOUBI; MAHJOUBI; MIRZAEI, 2017) were overexpressed our study for G1 and G2. Alterations in this gene are associated with diseases (ALDERA; GOVENDER, 2022) and some mutation in this gene would be expected to affect almost 25% of the total H3.3 protein in the cells depending on the differential gene expression (FANG et al., 2016). A unique feature of H3.3 histone genes is that they contain introns and the mRNA is polyadenylated, unlike most other histone genes that are polycistronic and intronless (SZENKER; RAY-GALLET; ALMOUZNI, 2011). In addition, reduced expression of *H3F3B* leads to decreased cell proliferation (HAUSER et al., 2015). In bovine, *H3F3B* has been described as a regulator of embryonic development (ZHANG et al., 2018). This gene could be helping to activate cell proliferation in LRFI group.

Other common genes for DAS transcripts were *HMGN2,* is a member of HMG-nucleosome binding family, a group of non-histone nuclear proteins that associate with the core nucleosome and alter the chromatin fiber structure (LUCEY et al., 2008). This gene and *H3F3B* gene could be involved in epigenetic regulation of gene expression. *HMGN2* regulates cell proliferation of neuronal progenitor cells in telencephalon, induces autophagy by activating AMPK/ULK1 pathway and thereby boosts UPEC proliferation within bladder epithelial cells (MA et al., 2021b). *HMGN2* acts as potentially essential factors in the innate immunity of human reproductive tract, and regulates lipopolysaccharide induced β-defensin expression in rat genital tract inflammation model (XIE et al., 2011). HMGN2 may be an important defense factor in innate immune response (WANG et al., 2015; ZHANG et al., 2019b) *HMGN2* could play a vital role immune response in RFI.

Another gene of interest for RFI is *TMSB10*, which is up-regulated in fast-growing chickens (CLAIRE D'ANDRE et al., 2013). This gene is related to lipid homeostasis that may involve subcellular transport (BARTZ et al., 2009), cytoskeleton organization, (NUKALA et al., 2019) and promotion of cell proliferation (SONG; SU; GUO, 2019) cell morphology, and expression of the

anti-apoptotic protein bcl-2 (GUTIÉRREZ-PABELLO; MCMURRAY; ADAMS, 2002). In the liver, loss of TMSB10 expression reduces proliferation and cell migration (SONG; SU; GUO, 2019). KONG et al.; 2016 studying rumen epithelial in cattle also found this gene upregulating the cytoskeletal organization in LRFI.

In summary, all nine genes are found regulating the immune system (Figure 1) and also the genes to those mentioned above (*SAT1* and *HMGN2*). *C1QC* gene has been reported to be essential for development of host innate immune response (PERGANDE et al., 2019) and *NCP2* gene which could modulate inflammation and innate immunity (FROLOV et al., 2013). In addition to the *NCP2* which is related to positively regulated lipid transport, it was found overexpressed in the less efficient group in Nelore (TIZIOTO et al., 2016). In general it has been speculated that by stimulating the immune response, the body's energy consumption increases, thus lowering feed efficiency (ZHANG et al., 2019a). Nevertheless, the relationship between feed efficiency and immunity in cattle is not clarified, needing further investigation.

Moreover, *B2M, BOLA-DRA* and *CD74* genes expressed alternatively spliced transcripts were up-regulated in LRFI group were discussed in previous Serna-García et al. (Chapter 2 and 3).

In the chapter 2, we highlighted *B2M* as un possible biomarker of RFI. It has been identified in more efficient pigs relating it to lipid metabolism (HORODYSKA et al., 2016). Furthermore, it plays a key role in lipid homeostasis (HUANG; ZHAU; CHUNG, 2010). It is also involved in protein turnover (TAURINES et al., 2011). This gene regulates vital processes such as inhibition of cell death (LEÓN-LETELIER; BONIFAZ; FUENTES-PANANÁ, 2019) and reduces infections (CHEVALIEZ et al., 2008), having an important role in the regulation of the immune system (KOLLNBERGER, 2016).

In the chapter 3 ,we showed that *CD74* gene plays a role in regulating inflammatory responses, cell growth (FEX SVENNINGSEN et al., 2017; GIL-YAROM et al., 2017), stimulates cell proliferation (ZUO et al., 2011; BUCALA; SHACHAR, 2014; LE HIRESS et al., 2015) and immune response and cellular migration (GIL-YAROM et al., 2017). *CD74* deficiency decreases adaptive immune responses (SUN et al., 2010). It regulates cell survival (KITANGE et al., 2010; RUVOLO et al., 2019) and its absence leads to cell death (BECKER-HERMAN et al., 2021). In addition, this gene could have a significant action on

hepatic cells exerting an inflammatory state and regulating the accumulation of lipids in the liver preventing obesity in the most efficient group of cattle (MISHRA et al., 2016). Other genes cited in chapter 3 were; the major histocompatibility complex (MHC) class II DR gene product is encoded by the *BoLA-DRA* and *BoLA -DRB3* genes and have a crucial role in determining the immune response (FISCH et al., 2021). The BoLA gene was associated with cattle immune response to infections in Nelore (FERNANDES JÚNIOR et al., 2020). Therefore, in the process of feeding, ensuring body health is an important basis to maximize the growth performance of animals.

Conclusion

We found nine genes that were predicted to produce alternatively spliced isoforms that were differentially expressed have also been implicated in lipid metabolism and immune system. All of these processes could promote changes in RFI phenotypes. Several of the genes that were found to produce differentially expressed splice transcripts in this study; have not previously been reported to be associated with RFI trait of beef cattle.

References

ALDERA, A. Pietro; GOVENDER, D. Gene of the month: H3F3A and H3F3B. **Journal of Clinical Pathology**, v. 75, n. 1, p. 1–4, 1 jan. 2022.

ASSENOV, Y. et al. Computing topological parameters of biological networks. **Bioinformatics (Oxford, England)**, v. 24, n. 2, p. 282–284, 15 jan. 2008.

AYOUBI, H. A.; MAHJOUBI, F.; MIRZAEI, R. Investigation of the human H3.3B (H3F3B) gene expression as a novel marker in patients with colorectal cancer. **Journal of Gastrointestinal Oncology**, v. 8, n. 1, p. 64–69, 2017.

BARABÁSI, A. L.; OLTVAI, Z. N. Network biology: understanding the cell's functional organization. **Nature reviews. Genetics**, v. 5, n. 2, p. 101–113, fev. 2004.

BARTZ, F. et al. Identification of Cholesterol-Regulating Genes by Targeted RNAi Screening. **Cell Metabolism**, v. 10, n. 1, p. 63–75, 8 jul. 2009.

BECKER-HERMAN, S. et al. CD74 is a regulator of hematopoietic stem cell maintenance. **PLoS biology**, v. 19, n. 3, 4 mar. 2021.

BENJAMINI, Y.; HOCHBERG, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. **Source: Journal of the Royal Statistical Society. Series B (Methodological)**, v. 57, n. 1, p. 289–300, 1995.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics (Oxford, England)**, v. 30, n. 15, p. 2114–2120, 1 ago. 2014.

BRETT-MORRIS, A. et al. The Polyamine Catabolic Enzyme SAT1 Modulates Tumorigenesis and Radiation Response in GBM. **Cancer Research**, v. 74, n. 23, p. 6925–6934, 1 dez. 2014.

BUCALA, R.; SHACHAR, I. The integral role of CD74 in antigen presentation, MIF signal transduction, and B cell survival and homeostasis. **Mini reviews in medicinal chemistry**, v. 14, n. 14, p. 1132–1138, 26 jan. 2014.

CASTOLDI, F. et al. Chemical activation of SAT1 corrects diet-induced metabolic syndrome. **Cell Death & Differentiation 2020 27:10**, v. 27, n. 10, p. 2904–2920, 6 maio 2020.

CHEVALIEZ, S. et al. Role of class I human leukocyte antigen molecules in early steps of echovirus infection of rhabdomyosarcoma cells. **Virology**, v. 381, n. 2, p. 203–214, 25 nov. 2008.

CHIN, C. H. et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. **BMC systems biology**, v. 8 Suppl 4, n. Suppl 4, 8 dez. 2014.

CLAIRE D'ANDRE, H. et al. Identification and characterization of genes that control fat deposition in chickens. **Journal of Animal Science and Biotechnology**, v. 4, n. 1, p. 43, 9 nov. 2013.

DONG, J.; HORVATH, S. Understanding network concepts in modules. **BMC systems biology**, v. 1, 4 jun. 2007.

FANG, D. et al. The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. **Science**, v. 352, n. 6291, p. 1344–1348, 10 jun. 2016.

FANG, X. et al. Comparative Genome-Wide Alternative Splicing Analysis of Longissimus Dorsi Muscles Between Japanese Black (Wagyu) and Chinese Red Steppes Cattle. **Frontiers in veterinary science**, v. 8, 29 abr. 2021.

FERNANDES JÚNIOR, G. A. et al. Whole-genome sequencing provides new insights into genetic mechanisms of tropical adaptation in Nellore (Bos primigenius indicus). **Scientific Reports**, v. 10, n. 1, p. 1–7, 10 jun. 2020.

FEX SVENNINGSEN, Å. et al. Macrophage migration inhibitory factor (MIF) modulates trophic signaling through interaction with serine protease HTRA1. **Cellular and molecular life sciences : CMLS**, v. 74, n. 24, p. 4561–4572, 1 dez. 2017.

FISCH, A. et al. Integral Use of Immunopeptidomics and Immunoinformatics for the Characterization of Antigen Presentation and Rational Identification of BoLA-DR–Presented Peptides and Epitopes. **The Journal of Immunology**, v. 206, n. 10, p. 2489–2497, 15 maio 2021.

FROLOV, A. et al. Niemann-pick type C2 deficiency in human fibroblasts confers robust and selective activation of prostaglandin E2 biosynthesis. **The Journal of biological chemistry**, v. 288, n. 33, p. 23696–23703, 16 ago. 2013.

GARZA-MANERO, S. et al. Maintenance of active chromatin states by HMGN2 is

required for stem cell identity in a pluripotent stem cell model. **Epigenetics & Chromatin**, v. 12, n. 1, 12 dez. 2019.

GIL-YAROM, N. et al. CD74 is a novel transcription regulator. **Proceedings of the National Academy of Sciences of the United States of America**, v. 114, n. 3, p. 562–567, 17 jan. 2017.

GLEESON, J. et al. Accurate expression quantification from nanopore direct RNA sequencing with NanoCount. **Nucleic Acids Research**, v. 50, n. 4, p. e19–e19, 28 fev. 2022.

GUTIÉRREZ-PABELLO, J. A.; MCMURRAY, D. N.; ADAMS, L. G. Upregulation of Thymosin β-10 by Mycobacterium bovis Infection of Bovine Macrophages Is Associated with Apoptosis. **Infection and Immunity**, v. 70, n. 4, p. 2121, 2002.

HAMM, A. et al. Frequent expression loss of Inter-alpha-trypsin inhibitor heavy chain (ITIH) genes in multiple human solid tumors: a systematic expression analysis. **BMC cancer**, v. 8, 28 jan. 2008.

HAUSER, B. et al. Functions of MiRNA-128 on the regulation of head and neck squamous cell carcinoma growth and apoptosis. **PloS one**, v. 10, n. 3, 12 mar. 2015.

HOO, Z. H.; CANDLISH, J.; TEARE, D. What is an ROC curve? **Emergency medicine journal : EMJ**, v. 34, n. 6, p. 357–359, 1 jun. 2017.

HORODYSKA, J. et al. P3024 Transcriptome analysis of longissimus thoracis et lumborum from pigs divergent in residual feed intake. **Journal of Animal Science**, v. 94, n. suppl_4, p. 63–64, 1 set. 2016.

HUANG, D. W.; SHERMAN, B. T.; LEMPICKI, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. **Nature protocols**, v. 4, n. 1, p. 44–57, 2009.

HUANG, W. C.; ZHAU, H. E.; CHUNG, L. W. K. Androgen receptor survival signaling is blocked by anti-β2- microglobulin monoclonal antibody via a MAPK/lipogenic pathway in human prostate cancer cells. **Journal of Biological Chemistry**, v. 285, n. 11, p. 7947–7956, 12 mar. 2010.

HUI, S. et al. **Cytoscape App Store - ExpressionCorrelation**. Disponível em: <https://apps.cytoscape.org/apps/expressioncorrelation>.

JELL, J. et al. Genetically altered expression of spermidine/spermine N1-acetyltransferase affects fat metabolism in mice via acetyl-CoA. **The Journal of biological chemistry**, v. 282, n. 11, p. 8404–8413, 16 mar. 2007.

KANEHISA, M.; GOTO, S. KEGG: kyoto encyclopedia of genes and genomes. **Nucleic acids research**, v. 28, n. 1, p. 27–30, 1 jan. 2000.

KIM, D.; LANGMEAD, B.; SALZBERG, S. L. HISAT: a fast spliced aligner with low memory requirements. **Nature methods**, v. 12, n. 4, p. 357–360, 31 mar. 2015.

KITANGE, G. J. et al. Expression of CD74 in high grade gliomas: a potential role in temozolomide resistance. **Journal of neuro-oncology**, v. 100, n. 2, p. 177–186, nov. 2010.

KLEIN, J. et al. Nomenclature for the major histocompatibility complexes of different species: a proposal. **Immunogenetics 1990 31:4**, v. 31, n. 4, p. 217–219, 1990.

KOLLNBERGER, S. The Role of HLA-Class I Heavy-Chain Interactions with Killer-Cell Immunoglobulin-Like Receptors in Immune Regulation. **Critical reviews in immunology**, v. 36, n. 3, p. 269–282, 2016.

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. 1, p. 1–16, 9 ago. 2016.

KRAMER, D. L. et al. Polyamine acetylation modulates polyamine metabolic flux, a prelude to broader metabolic consequences. **The Journal of biological chemistry**, v. 283, n. 7, p. 4241–4251, 15 fev. 2008.

LAW, R. H. P. et al. An overview of the serpin superfamily. **Genome Biology**, v. 7, n. 5, p. 216, 30 maio 2006.

LE HIRESS, M. et al. Proinflammatory Signature of the Dysfunctional Endothelium in Pulmonary Hypertension. Role of the Macrophage Migration Inhibitory Factor/CD74 Complex. **American journal of respiratory and critical care medicine**, v. 192, n. 8, p. 983–997, 15 out. 2015.

LEE, S. B. et al. Suppression of Exogenous Gene Expression by Spermidine/Spermine N1-Acetyltransferase 1 (SSAT1) Cotransfection. **The Journal of Biological Chemistry**, v. 285, n. 20, p. 15548, 14 maio 2010.

LEÓN-LETELIER, R. A.; BONIFAZ, L. C.; FUENTES-PANANÁ, E. M. OMIC signatures to understand cancer immunosurveillance and immunoediting: Melanoma and immune cells interplay in immunotherapy. **Journal of leukocyte biology**, v. 105, n. 5, p. 915–933, 1 maio 2019.

LERCH, J. K. et al. Isoform diversity and regulation in peripheral and central neurons revealed through RNA-Seq. **PloS one**, v. 7, n. 1, 17 jan. 2012.

LIU, Y. X. et al. Effects of Intensive Fattening With Total Mixed Rations on Carcass Characteristics, Meat Quality, and Meat Chemical Composition of Yak and Mechanism Based on Serum and Transcriptomic Profiles. **Frontiers in veterinary science**, v. 7, 21 jan. 2021.

LUCEY, M. M. et al. Differential expression of the HMGN family of chromatin proteins during ocular development. **Gene expression patterns : GEP**, v. 8, n. 6, p. 433–437, jul. 2008.

MA, L. et al. Spermidine ameliorates high-fat diet-induced hepatic steatosis and adipose tissue inflammation in preexisting obese mice. **Life sciences**, v. 265, 15 jan. 2021a.

MA, L. et al. Developmental programming and lineage branching of early human telencephalon. **The EMBO journal**, v. 40, n. 21, 2 nov. 2021b.

MISHRA, A. et al. Role of antigen presenting cell invariant chain in the development of hepatic steatosis in mouse model. **Experimental cell research**, v. 346, n. 2, p. 188–197, 15 ago. 2016.

MUNIZ, M. M. M. et al. Identification of novel mRNA isoforms associated with meat tenderness using RNA sequencing data in beef cattle. **Meat science**, v. 173, 1 mar.

2021.

NUKALA, S. B. et al. Differentially Expressed Proteins in Primary Endothelial Cells Derived From Patients With Acute Myocardial Infarction. **Hypertension (Dallas, Tex. : 1979)**, v. 74, n. 4, p. 947–956, 1 out. 2019.

OSHLACK, A.; ROBINSON, M. D.; YOUNG, M. D. **From RNA-seq reads to differential expression resultsGenome Biology**, 22 dez. 2010. .

PAN, Q. et al. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. **Nature genetics**, v. 40, n. 12, p. 1413–1415, dez. 2008.

PEGG, A. E.; PEGG, A. E. Spermidine/spermine-N 1-acetyltransferase: a key metabolic regulator Pegg AE. Spermidine/spermine-N 1-acetyltransferase: a key metabolic regulator. **Am J Physiol Endocrinol Metab**, v. 294, p. 995–1010, 2008.

PERGANDE, M. R. et al. Differential Proteomics Reveals miR-155 as a Novel Indicator of Liver and Spleen Pathology in the Symptomatic Niemann-Pick Disease, Type C1 Mouse Model. **Molecules (Basel, Switzerland)**, v. 24, n. 5, 2019.

PERTEA, G.; PERTEA, M. GFF Utilities: GffRead and GffCompare. **F1000Research**, v. 9, 2020.

PIRINEN, E. et al. Enhanced polyamine catabolism alters homeostatic control of white adipose tissue mass, energy expenditure, and glucose metabolism. **Molecular and cellular biology**, v. 27, n. 13, p. 4953–4967, jul. 2007.

PIRNES-KARHU, S. et al. Lipopolysaccharide-induced anti-inflammatory acute phase response is enhanced in spermidine/spermine N1-acetyltransferase (SSAT) overexpressing mice. **Amino acids**, v. 42, n. 2–3, p. 473–484, fev. 2012.

PRAKASH, A. et al. Molecular basis of residual feed intake in broilers. **https://doi.org/10.1080/00439339.2020.1789534**, v. 77, n. 1, p. 175–189, 2020.

PROVOST, E.; WEIER, C. A.; LEACH, S. D. Multiple Ribosomal Proteins Are Expressed at High Levels in Developing Zebrafish Endoderm and Are Required for Normal Exocrine Pancreas Development. **Zebrafish**, v. 10, n. 2, p. 161, 1 jun. 2013.

ROBERT, C.; WATSON, M. Errors in RNA-Seq quantification affect genes of relevance to human disease. **Genome biology**, v. 16, n. 1, 3 set. 2015.

ROBERT, C.; WATSON, M. The incredible complexity of RNA splicing. **Genome Biology**, v. 17, n. 1, 30 dez. 2016.

RUVOLO, P. P. et al. LGALS3 is connected to CD74 in a previously unknown protein network that is associated with poor survival in patients with AML. **EBioMedicine**, v. 44, p. 126–137, 1 jun. 2019.

RYU, J. H. et al. Myocardial SSAT induction via AMPK signaling and its implication for ischemic injury. **Biochemical and biophysical research communications**, v. 366, n. 2, p. 438–444, 8 fev. 2008.

SHANNON, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome research**, v. 13, n. 11, p. 2498–2504, nov.

2003.

SILVA, D. B. S. et al. Spliced genes in muscle from Nelore Cattle and their association with carcass and meat quality. **Scientific Reports 2020 10:1**, v. 10, n. 1, p. 1–13, 7 set. 2020.

SMYTH, G. K. et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. **F1000Research**, v. 5, 2016.

SONG, C.; SU, Z.; GUO, J. Research Article Thymosin β 10 is overexpressed and associated with unfavorable prognosis in hepatocellular carcinoma. **Bioscience Reports**, v. 39, n. 3, 2019.

STEVENS, A. et al. Network analysis: a new approach to study endocrine disorders. **Journal of molecular endocrinology**, v. 52, n. 1, 1 out. 2013.

SUBRAMANIAN, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. **Proceedings of the National Academy of Sciences**, v. 102, n. 43, p. 15545–15550, 25 out. 2005.

SUBRAMANIAN, M. et al. The nucleosome-binding protein HMGN2 modulates global genome repair. **The FEBS journal**, v. 276, n. 22, p. 6646, nov. 2009.

SUN, H. Z. et al. Gene co-expression and alternative splicing analysis of key metabolic tissues to unravel the regulatory signatures of fatty acid composition in cattle. **RNA biology**, v. 18, n. 6, p. 854–862, 2021.

SUN, J. et al. Deficiency of antigen-presenting cell invariant chain reduces atherosclerosis in mice. **Circulation**, v. 122, n. 8, p. 808–820, 24 ago. 2010.

SZENKER, E.; RAY-GALLET, D.; ALMOUZNI, G. The double face of the histone variant H3.3. **Cell research**, v. 21, n. 3, p. 421–434, mar. 2011.

TAPPIA, P. S. et al. Predictive value and clinical significance of increased SSAT-1 activity in healthy adults. **Future science OA**, v. 5, n. 7, 1 abr. 2019.

TAURINES, R. et al. Review: Proteomic research in psychiatry. **Journal of Psychopharmacology**, v. 25, n. 2, p. 151–196, fev. 2011.

TILGNER, H. et al. Defining a personal, allele-specific, and single-molecule long-read transcriptome. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 27, p. 9869–9874, 8 jul. 2014. D

TIZIOTO, P. C. et al. Gene expression differences in Longissimus muscle of Nelore steers genetically divergent for residual feed intake. **Scientific Reports 2016 6:1**, v. 6, n. 1, p. 1–12, 22 dez. 2016.

TRAPNELL, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. **Nature Protocols**, v. 7, n. 3, p. 562–578, 1 mar. 2012.

TUNG, L. H.; SHAO, M.; KINGSFORD, C. Quantifying the benefit offered by transcript assembly with Scallop-LR on single-molecule long reads. **Genome Biology**, v. 20, n. 1, 18 dez. 2019.

UEMURA, T. et al. Identification and characterization of a diamine exporter in colon epithelial cells. **The Journal of biological chemistry**, v. 283, n. 39, p. 26428–26435, 26 set. 2008.

WANG, E. T. et al. Alternative isoform regulation in human tissue transcriptomes. **Nature**, v. 456, n. 7221, p. 470–476, 27 nov. 2008.

WANG, X. et al. High-mobility group nucleosome-binding domain 2 protein inhibits the invasion of Klebsiella pneumoniae into mouse lungs in vivo. **Molecular medicine reports**, v. 12, n. 1, p. 1279–1285, 1 jul. 2015.

WINGETT, S. W.; ANDREWS, S. FastQ Screen: A tool for multi-genome mapping and quality control. **F1000Research**, v. 7, p. 1338, 17 set. 2018.

XIE, P. et al. Expression of HMGB1 and HMGN2 in gingival tissues, GCF and PICF of periodontitis patients and peri-implantitis. **Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]**, v. 42, n. 3, p. 1213–1219, 2011.

ZAHEDI, K. et al. Proximal tubule epithelial cell specific ablation of the spermidine/spermine N1-acetyltransferase gene reduces the severity of renal ischemia/reperfusion injury. **PloS one**, v. 9, n. 11, 12 nov. 2014.

ZHANG, D. et al. Transcriptome Analysis Identifies Candidate Genes and Pathways Associated With Feed Efficiency in Hu Sheep. **Frontiers in Genetics**, v. 10, p. 1183, 19 nov. 2019a.

ZHANG, F. et al. High-mobility group protein N2 induces autophagy by activating AMPK/ULK1 pathway and thereby boosts UPEC proliferation within bladder epithelial cells. **Biochemical and biophysical research communications**, v. 513, n. 1, p. 193–200, 21 maio 2019b.

ZHANG, K. et al. Characterization of H3.3 and HIRA expression and function in bovine early embryos. **Molecular reproduction and development**, v. 85, n. 2, p. 106–116, 1 fev. 2018.

ZUO, J. et al. Epstein-Barr Virus Evades CD4+ T Cell Responses in Lytic Cycle through BZLF1-mediated Downregulation of CD74 and the Cooperation of vBcl-2. **PLoS Pathogens**, v. 7, n. 12, dez. 2011.

Additional Files

Supplementary Tables

Table S1. Alternatively spliced and differentially expressed ($p < 0.05$) transcripts identified in the liver tissue of Nelore bulls with the highest and lowest Residual Feed Intake (LRFI and HRFI) in Genetic group 1.

Gene	Transcripts type	ID transcripts	Chr	fold	statistic	p-value	padj
CD74	complete, exact match of intron chain	rna23815_7_61747921_61755746	7	0.49639	3.18589	0.00363	0.53289
BOLA-DRA	complete, exact match of intron chain	rna64263_23_25838391_25843092	23	0.46380	2.97613	0.00610	0.53289
C1QC	exonic overlap on the opposite strand	rna6207_2_130181311_130185682	2	0.43038	2.20061	0.03653	0.53289
CD74	complete, exact match of intron chain	rna23816_7_61747923_61755685	7	0.42912	2.65424	0.01318	0.53289
CSTB	complete, exact match of intron chain	rna2688_1_143846935_143850386	1	0.39829	2.67264	0.01263	0.53289
IDH1	complete, exact match of intron chain	rna4826_2_96510381_96531400	2	0.35041	2.50172	0.01874	0.53289
TMSB10	complete, exact match of intron chain	rna33617_11_50094264_50095274	11	0.33944	2.38858	0.02419	0.53289
RPL12	complete, exact match of intron chain	rna34815_11_98179657_98183783	11	0.33019	2.56655	0.01616	0.53289
ACTG1	complete, exact match of intron chain	rna56837_19_51259442_51262291	19	0.32825	2.40373	0.02339	0.53289
RPS23	multi-exon with at least one junction match	rna24137_7_82217490_82219272	7	0.32462	2.27493	0.03110	0.53289
CCL5	complete, exact match of intron chain	rna53802_19_14509140_14515888	19	0.32433	2.12422	0.04299	0.53289
LOC518526	complete, exact match of intron chain	rna11649_4_40319923_40360802	4	0.31958	2.65090	0.01328	0.53289
NPC2	complete, exact match of intron chain	rna31803_10_85774970_85783554	10	0.31392	3.17915	0.00370	0.53289
AIF1	complete, exact match of intron chain	rna64484_23_27685447_27687723	23	0.30858	2.08910	0.04629	0.54973
LOC783680	complete, exact match of intron chain	rna32157_10_103116413_103122698	10	0.29662	3.60050	0.00126	0.53289
SDC4	complete, exact match of intron chain	rna39728_13_73669855_73691288	13	0.29200	2.32774	0.02769	0.53289
CXCL16	unclassified	rna54590_19_26615779_26619263	19	0.28188	2.30123	0.02936	0.53289
GNAI2	complete, exact match of intron chain	rna62383_22_50099967_50120142	22	0.26311	2.22611	0.03458	0.53289
MIF	complete, exact match of intron chain	rna48965_17_71292117_71292909	17	0.26185	2.38671	0.02430	0.53289

SCPEP1	unclassified	rna53406_19_7813226_7847070	19	0.25103	2.06218	0.04897	0.54973
OTC	complete, exact match of intron chain	rna77538_X_104841977_104918525	x	0.25102	2.43303	0.02190	0.53289
B2M	complete, exact match of intron chain	rna32156_10_103095506_103103791	10	0.24893	3.05302	0.00505	0.53289
SAT1	complete, exact match of intron chain	rna77669_X_119431931_119434979	x	0.24164	3.00657	0.00567	0.53289
CTSC	complete, exact match of intron chain	rna73162_29_7375413_7415746	29	0.22479	2.43857	0.02163	0.53289
PROS1	multi-exon with at least one junction match	rna383_1_38203318_38267163	1	0.21866	2.15114	0.04060	0.53289
SYNGR2	complete, exact match of intron chain	rna57050_19_53992203_53995976	19	0.21115	2.44264	0.02143	0.53289
RPL12	complete, exact match of intron chain	rna14028_5_25781131_25785642	5	0.19812	2.58634	0.01543	0.53289
RPS12	complete, exact match of intron chain	rna28227_9_70977043_70980384	9	0.19737	2.27548	0.03106	0.53289
PSME2	complete, exact match of intron chain	rna29684_10_21004530_21007790	10	0.19719	2.78887	0.00959	0.53289
H3F3B	complete, exact match of intron chain	rna57283_19_55821418_55823783	19	0.19461	2.57755	0.01575	0.53289
HMGN2	multi-exon with at least one junction match	rna5986_2_126660981_126664547	2	0.19224	2.31206	0.02867	0.53289
ARPC2	complete, exact match of intron chain	rna5041_2_106256774_106284391	2	0.17232	2.49919	0.01885	0.53289
PABPC1	complete, exact match of intron chain	rna41491_14_63625778_63643551	14	0.16525	2.18521	0.03776	0.53289
LAMP2	complete, exact match of intron chain	rna75237_X_4990445_5023582	x	0.14281	2.12742	0.04270	0.53289
IAH1	complete, exact match of intron chain	rna34477_11_87961967_87973214	11	-0.14707	-2.46702	0.02028	0.53289
HSD17B10	complete, exact match of intron chain	rna77269_X_90870246_90872569	x	-0.15234	-2.40857	0.02314	0.53289
COX5B	complete, exact match of intron chain	rna32319_11_2947406_2949454	11	-0.16232	-2.11929	0.04344	0.53289
ITIH3	multi-exon with at least one junction match	rna62148_22_48078022_48091969	22	-0.16953	-2.24029	0.03353	0.53289
LOC525415	complete, exact match of intron chain	rna53784_19_14243659_14279682	19	-0.17179	-2.27168	0.03132	0.53289
F2	unclassified	rna43996_15_76656867_76671235	15	-0.18064	-2.30439	0.02915	0.53289
EPHX1	complete, exact match of intron chain	rna45189_16_28833630_28870961	16	-0.18313	-2.24127	0.03346	0.53289
UBL5	complete, exact match of intron chain	rna21274_7_14420644_14422238	7	-0.18340	-2.06594	0.04859	0.54973
HPD	complete, exact match of intron chain	rna48000_17_53458851_53470907	17	-0.18496	-2.06855	0.04832	0.54973
ASL	complete, exact match of intron chain	rna68462_25_27989925_27998697	25	-0.18909	-2.11871	0.04349	0.53289

APOM	complete, exact match of intron chain	rna64456_23_27641578_27643685	23	-0.18952	-2.64862	0.01335	0.53289
SEC11C	complete, exact match of intron chain	rna66812_24_58152254_58163466	24	-0.19695	-2.15685	0.04011	0.53289
METTL7B	complete, exact match of intron chain	rna15385_5_57571968_57574965	5	-0.19903	-2.14343	0.04128	0.53289
QPRT	complete, exact match of intron chain	rna68282_25_26491479_26506183	25	-0.20507	-2.78937	0.00958	0.53289
ASS1	complete, exact match of intron chain	rna35155_11_100770227_100822223	11	-0.20547	-2.40913	0.02311	0.53289
APOE	complete, exact match of intron chain	rna51716_18_52637890_52640576	18	-0.20564	-2.18439	0.03782	0.53289
FAH	complete, exact match of intron chain	rna59374_21_26110597_26141035	21	-0.20710	-2.80884	0.00915	0.53289
HMGCL	complete, exact match of intron chain	rna6145_2_129091585_129109718	2	-0.21530	-2.13335	0.04217	0.53289
TCN2	complete, exact match of intron chain	rna48814_17_69579627_69595610	17	-0.22096	-2.34800	0.02648	0.53289
LRG1	unclassified	rna21952_7_19596348_19599306	7	-0.22609	-2.11631	0.04371	0.53289
HAO2	complete, exact match of intron chain	rna8167_3_23732793_23763956	3	-0.22763	-2.53549	0.01735	0.53289
FMO1	unclassified	rna45408_16_38810273_38849071	16	-0.24513	-2.62769	0.01402	0.53289
FGA	complete, exact match of intron chain	rna47012_17_2860755_2867515	17	-0.25390	-2.92171	0.00697	0.53289
SPON2	complete, exact match of intron chain	rna20168_6_117253243_117256854	6	-0.25515	-2.37980	0.02467	0.53289
CRP	complete, exact match of intron chain	rna6959_3_9983501_9987619	3	-0.25849	-2.24657	0.03308	0.53289
ITIH4	unclassified	rna62146_22_48059589_48074580	22	-0.26083	-2.05805	0.04939	0.54973
CPN2	complete, exact match of intron chain	rna1224_1_73178781_73188077	1	-0.26639	-2.11712	0.04364	0.53289
A1BG	unclassified	rna53260_18_65641101_65645263	18	-0.28368	-2.05837	0.04936	0.54973
DECR2	unclassified	rna66974_25_388144_396055	25	-0.28871	-2.27659	0.03098	0.53289
ACAA2	multi-exon with at least one junction match	rna66608_24_49442437_49477387	24	-0.30727	-2.42819	0.02214	0.53289
NIPSNAP1	complete, exact match of intron chain	rna48705_17_68732899_68748529	17	-0.32529	-2.26024	0.03211	0.53289
LOC515150	complete, exact match of intron chain	rna44885_16_5142272_5150340	16	-0.32974	-3.56135	0.00140	0.53289
SERPINA3-8	complete, exact match of intron chain	59317212_59326423	21	-0.35906	-2.75761	0.01033	0.53289
SERPINA3-7	complete, exact match of intron chain	59355554_59363576	21	-0.35917	-2.14510	0.04113	0.53289
PCBD1	complete, exact match of intron chain	27067757_27072686	28	-0.48801	-2.65911	0.01303	0.53289

Table S2. Alternatively spliced and differentially expressed ($p < 0.05$) transcripts identified in the liver tissue of Nelore bulls with the highest and lowest Residual Feed Intake (LRFI and HRFI) in Genetic group 2.

gene	Transcripts type	ID trancripts	Chr	fold	statistic	p-value	padj
FABP1	complete, exact match of intron chain	rna33500_11_47917547_47923246	11	0.62693	3.59964	0.00150	0.16464
C1QA	complete, exact match of intron chain	rna6209_2_130189866_130192754	2	0.51303	3.52348	0.00181	0.16464
CFD	complete, exact match of intron chain	rna22893_7_43416267_43418803	7	0.42186	2.61168	0.01555	0.29301
C1QC	exonic overlap on the opposite strand	rna6207_2_130181311_130185682	2	0.40576	3.25770	0.00345	0.17257
LECT2	unclassified	rna23202_7_47392297_47404084	7	0.39054	3.49311	0.00195	0.16464
NPC2	complete, exact match of intron chain	rna31803_10_85774970_85783554	10	0.37130	3.68332	0.00122	0.16464
LOC510860	complete, exact match of intron chain	rna44884_16_5087507_5098711	16	0.37069	3.19550	0.00400	0.17257
COX7C	complete, exact match of intron chain	rna24165_7_86288746_86290953	7	0.36121	2.83112	0.00943	0.23089
COX6A1	complete, exact match of intron chain	rna48418_17_62744136_62746010	17	0.35995	3.58989	0.00154	0.16464
RPL27	complete, exact match of intron chain	rna56113_19_43026105_43028407	19	0.34522	3.10124	0.00501	0.17257
BOLA-DRA	complete, exact match of intron chain	rna64263_23_25838391_25843092	23	0.34182	2.46625	0.02149	0.30646
TMSB10	complete, exact match of intron chain	rna33617_11_50094264_50095274	11	0.33486	3.23566	0.00363	0.17257
RPSA	unclassified	rna61249_22_12685736_12697678	22	0.33478	4.00798	0.00055	0.16464
CST3	multi-exon with at least one junction match	rna38230_13_42185282_42189207	13	0.33366	2.84082	0.00923	0.23089
COX6B1	complete, exact match of intron chain	rna50914_18_46378496_46387003	18	0.32596	3.09400	0.00510	0.17257
TMSB4X	exonic overlap on the opposite strand	rna77928_X_130698922_130700970	x	0.32545	3.20555	0.00391	0.17257
CD74	complete, exact match of intron chain	rna23816_7_61747923_61755685	7	0.30881	2.16851	0.04065	0.39821
YBX1	complete, exact match of intron chain	rna10182_3_103520069_103539604	3	0.29856	2.39883	0.02490	0.33658
ACTB	complete, exact match of intron chain	rna69064_25_38799230_38802643	25	0.28946	3.36606	0.00265	0.16551
RPS21	complete, exact match of intron chain	rna38841_13_54866479_54867739	13	0.28391	2.81002	0.00990	0.23089
UQCR10	complete, exact match of intron chain	rna48713_17_68887320_68889516	17	0.27628	3.04373	0.00574	0.17257

H3F3A	complete, exact match of intron chain	rna45200_16_29079663_29087863	16	0.26409	2.77904	0.01064	0.23967
RPS14	complete, exact match of intron chain	rna23820_7_61774364_61778425	7	0.26403	2.61753	0.01535	0.29301
TCN2	complete, exact match of intron chain	rna48814_17_69579627_69595610	17	0.25390	3.13095	0.00467	0.17257
ATP5PB	multi-exon with at least one junction match	rna8414_3_31846422_31863330	3	0.25188	2.24930	0.03431	0.39307
RPL39	complete, exact match of intron chain	rna75179_X_4150544_4153940	x	0.24801	2.68356	0.01323	0.27938
ATP5F1D	complete, exact match of intron chain	rna22967_7_43683704_43685876	7	0.24509	2.87764	0.00847	0.22940
SAT1	complete, exact match of intron chain	rna77669_X_119431931_119434979	x	0.24426	2.57718	0.01680	0.29301
ABHD14B	complete, exact match of intron chain	rna62250_22_48961511_48965813	22	0.24294	2.25881	0.03362	0.39307
GAMT	complete, exact match of intron chain	rna22993_7_43814686_43817695	7	0.24175	2.25675	0.03377	0.39307
SPON2	complete, exact match of intron chain	rna20168_6_117253243_117256854	6	0.24172	3.38928	0.00251	0.16551
GRHPR	complete, exact match of intron chain	rna25648_8_61555137_61564033	8	0.23970	2.47465	0.02109	0.30646
RPS29	complete, exact match of intron chain	rna30500_10_42627498_42629350	10	0.23841	3.16649	0.00429	0.17257
B2M	complete, exact match of intron chain	rna32156_10_103095506_103103791	10	0.23654	2.69779	0.01280	0.27923
RPL18A	complete, exact match of intron chain	rna20625_7_5279456_5282849	7	0.22984	3.35999	0.00269	0.16551
RPL28	complete, exact match of intron chain	rna52834_18_62099120_62101876	18	0.22797	3.53000	0.00178	0.16464
SLC27A5	complete, exact match of intron chain	rna53285_18_65732879_65741718	18	0.22367	2.26097	0.03347	0.39307
ATP5MPL	complete, exact match of intron chain	rna60739_21_68460844_68466200	21	0.22140	3.03422	0.00587	0.17257
HINT2	complete, exact match of intron chain	rna25555_8_59979706_59981974	8	0.22138	2.46108	0.02173	0.30646
CTSZ	complete, exact match of intron chain	rna38905_13_57381213_57390707	13	0.22078	2.48650	0.02055	0.30646
RACK1	complete, exact match of intron chain	rna22702_7_40371008_40375831	7	0.21949	2.81078	0.00989	0.23089
MYL6	complete, exact match of intron chain	rna15309_5_57161868_57165070	5	0.21615	2.17018	0.04050	0.39821
RPS9	complete, exact match of intron chain	rna53009_18_63196984_63204297	18	0.21238	2.87698	0.00848	0.22940
RPS28	complete, exact match of intron chain	rna21646_7_16968061_16969193	7	0.20776	3.06279	0.00549	0.17257
AKR1C4	complete, exact match of intron chain	rna38331_13_43616923_43626630	13	0.20603	2.08210	0.04858	0.42293
PPDPF	unclassified	rna38750_13_54136186_54137833	13	0.20480	2.31007	0.03015	0.37052

TPI1	complete, exact match of intron chain	rna16748_5_103580104_103583461	5	0.19874	2.31035	0.03013	0.37052
POLR2L	complete, exact match of intron chain	rna74965_29_50530776_50535563	29	0.19869	2.59619	0.01610	0.29301
VKORC1	complete, exact match of intron chain	rna68384_25_27201070_27203315	25	0.19646	2.53719	0.01837	0.30646
H3F3B	complete, exact match of intron chain	rna57283_19_55821418_55823783	19	0.19033	2.38520	0.02564	0.33990
RXRA	complete, exact match of intron chain	rna35578_11_105086960_105114880	11	0.18413	2.07990	0.04880	0.42293
HMGN2	multi-exon with at least one junction match	rna5986_2_126660981_126664547	2	0.18354	2.50300	0.01981	0.30646
AHCY	multi-exon with at least one junction match	rna39293_13_63686719_63702437	13	0.18189	2.09210	0.04760	0.42293
RPL34	complete, exact match of intron chain	rna18127_6_16671257_16675919	6	0.18138	2.22099	0.03642	0.39379
HMGN2	multi-exon with at least one junction match	rna70082_26_23982412_23987702	26	0.18116	2.50815	0.01959	0.30646
RPS27A	multi-exon with at least one junction match	rna33242_11_37970413_37972642	11	0.17719	2.10136	0.04670	0.42293
NDUFA7	complete, exact match of intron chain	rna21645_7_16960703_16967936	7	0.17546	2.14865	0.04236	0.40120
RPS17	complete, exact match of intron chain	rna59249_21_22839050_22842498	21	0.17400	2.57450	0.01690	0.29301
GAPDH	complete, exact match of intron chain	rna16824_5_103870384_103874667	5	0.17271	2.19515	0.03845	0.39379
ITM2B	complete, exact match of intron chain	rna36000_12_18036781_18061776	12	0.16924	2.20353	0.03778	0.39379
RPL29	unclassified	rna62244_22_48945147_48947322	22	0.16888	2.50291	0.01982	0.30646
RPS13	complete, exact match of intron chain	rna42631_15_35408334_35411055	15	0.16826	2.18141	0.03957	0.39821
SH3BGRL3	complete, exact match of intron chain	rna6011_2_126837551_126839005	2	0.16741	2.13290	0.04376	0.40525
UQCRH	complete, exact match of intron chain	rna9844_3_99790148_99799423	3	0.16398	2.35494	0.02738	0.34918
LOC515042	other sample strand overlap with reference exons	rna6233_2_130801971_130806189	2	0.16301	2.23833	0.03511	0.39379
RPS25	complete, exact match of intron chain	rna42437_15_29588693_29590414	15	0.16019	2.14441	0.04273	0.40120
COX4I1	complete, exact match of intron chain	rna49602_18_11759822_11767079	18	0.15580	2.57814	0.01677	0.29301
CALM1	complete, exact match of intron chain	rna32117_10_102005688_102013428	10	0.15479	2.15934	0.04143	0.40008
RPS16	complete, exact match of intron chain	rna51284_18_49122734_49125200	18	0.14995	2.57769	0.01678	0.29301
DAZAP2	complete, exact match of intron chain	rna14309_5_28547801_28552009	5	0.14907	2.43361	0.02308	0.31839

PARK7	complete, exact match of intron chain	rna45713_16_45400217_45416974	16	0.14891	2.08556	0.04824	0.42293
ATP5ME	complete, exact match of intron chain	rna20216_6_117650797_117652333	6	0.14666	2.21625	0.03678	0.39379
AMY2B	complete, exact match of intron chain	rna8639_3_39784084_39806732	3	-0.20270	-2.21886	0.03658	0.39379
SLC38A4	complete, exact match of intron chain	rna14664_5_33435201_33486055	5	-0.27078	-2.20074	0.03800	0.39379
CHPT1	complete, exact match of intron chain	rna15627_5_65511775_65541515	5	-0.27955	-2.37532	0.02620	0.34057
STT3A	complete, exact match of intron chain	rna73614_29_29006101_29027284	29	-0.29378	-3.57112	0.00161	0.16464
CPS1	complete, exact match of intron chain	rna4929_2_98327892_98467680	2	-0.36465	-3.05396	0.00560	0.17257
ACSL1	complete, exact match of intron chain	rna71205_27_15178952_15243834	27	-0.49234	-2.46046	0.02176	0.30646

Table S3. Biological process GO terms containing genes with differentially expressed alternatively spliced transcripts in liver tissue of Nelore bulls selected to be divergent for RFI in G1 and G2. Gene enrichment analysis (padj<0.05) used GSEA software.

G1					
GO term	**N**	**LOR***	**p-value**	**padj**	**name**
GO:0030041	10	1.407	6.90E-08	6.72E-04	actin filament polymerization
GO:0008154	12	1.252	1.24E-06	6.04E-03	actin polymerization or depolymerization
GO:0045321	80	0.605	2.56E-06	8.31E-03	leukocyte activation
GO:0007015	16	1.109	5.37E-06	1.15E-02	actin filament organization
GO:0009057	143	0.439	5.90E-06	1.15E-02	macromolecule catabolic process
GO:0019222	308	0.349	8.10E-06	1.16E-02	regulation of metabolic process
GO:0010605	195	0.388	9.16E-06	1.16E-02	negative regulation of macromolecule metabolic process
GO:0060255	274	0.349	1.21E-05	1.16E-02	regulation of macromolecule metabolic process
GO:0044265	132	0.432	1.43E-05	1.16E-02	cellular macromolecule catabolic process
GO:0010608	51	0.648	1.49E-05	1.16E-02	posttranscriptional regulation of gene expression
GO:0006401	93	0.489	1.53E-05	1.16E-02	RNA catabolic process

GO:0002274	67	0.597	1.62E-05	1.16E-02	myeloid leukocyte activation
GO:0046651	11	1.331	1.71E-05	1.16E-02	lymphocyte proliferation
GO:0032943	11	1.331	1.71E-05	1.16E-02	mononuclear cell proliferation
GO:0006402	91	0.490	1.78E-05	1.16E-02	mRNA catabolic process
GO:0038093	22	0.924	2.07E-05	1.26E-02	Fc receptor signaling pathway
GO:0070661	12	1.236	2.25E-05	1.29E-02	leukocyte proliferation
GO:0034655	100	0.462	2.57E-05	1.33E-02	nucleobase-containing compound catabolic process
GO:0044260	339	0.323	2.59E-05	1.33E-02	cellular macromolecule metabolic process
GO:0009059	208	0.354	3.10E-05	1.47E-02	macromolecule biosynthetic process
GO:0001775	90	0.505	3.18E-05	1.47E-02	cell activation
GO:0006955	144	0.407	3.59E-05	1.57E-02	immune response
GO:0034645	206	0.351	3.70E-05	1.57E-02	cellular macromolecule biosynthetic process
GO:0065007	459	0.321	4.20E-05	1.68E-02	biological regulation
GO:0050789	430	0.316	4.32E-05	1.68E-02	regulation of biological process
GO:0046700	110	0.428	5.51E-05	2.06E-02	heterocycle catabolic process
GO:0016070	168	0.365	5.83E-05	2.10E-02	RNA metabolic process
GO:0044419	190	0.352	6.28E-05	2.14E-02	interspecies interaction between organisms
GO:0016071	101	0.437	6.49E-05	2.14E-02	mRNA metabolic process
GO:0002694	22	0.925	6.59E-05	2.14E-02	regulation of leukocyte activation
GO:0010468	199	0.341	7.28E-05	2.14E-02	regulation of gene expression
GO:0002768	25	0.812	7.31E-05	2.14E-02	immune response-regulating cell surface receptor signaling pathway
GO:0002764	25	0.812	7.31E-05	2.14E-02	immune response-regulating signaling pathway
GO:0046649	22	0.940	7.48E-05	2.14E-02	lymphocyte activation
GO:0016032	124	0.399	8.24E-05	2.29E-02	viral process
GO:0002252	102	0.444	8.74E-05	2.36E-02	immune effector process
GO:0019884	25	0.810	9.42E-05	2.40E-02	antigen processing and presentation of exogenous antigen
GO:0002478	25	0.810	9.42E-05	2.40E-02	antigen processing and presentation of exogenous peptide antigen
GO:0090304	175	0.349	9.64E-05	2.40E-02	nucleic acid metabolic process
GO:0009892	209	0.329	1.00E-04	2.44E-02	negative regulation of metabolic process

GO:0044267	295	0.299	1.19E-04	2.66E-02	cellular protein metabolic process
GO:0002376	177	0.346	1.20E-04	2.66E-02	immune system process
GO:0110053	12	1.094	1.21E-04	2.66E-02	regulation of actin filament organization
GO:0030029	25	0.817	1.22E-04	2.66E-02	actin filament-based process
GO:0030036	24	0.840	1.23E-04	2.66E-02	actin cytoskeleton organization
GO:0080090	222	0.319	1.29E-04	2.74E-02	regulation of primary metabolic process
GO:0002429	24	0.792	1.49E-04	3.01E-02	immune response-activating cell surface receptor signaling pathway
GO:0002757	24	0.792	1.49E-04	3.01E-02	immune response-activating signal transduction
GO:2000112	110	0.406	1.51E-04	3.01E-02	regulation of cellular macromolecule biosynthetic process
GO:0006412	109	0.401	1.57E-04	3.06E-02	translation
GO:0032956	15	0.973	1.66E-04	3.10E-02	regulation of actin cytoskeleton organization
GO:0032970	15	0.973	1.66E-04	3.10E-02	regulation of actin filament-based process
GO:0044270	113	0.391	1.89E-04	3.47E-02	cellular nitrogen compound catabolic process
GO:0006413	85	0.436	2.00E-04	3.58E-02	translational initiation
GO:1903706	25	0.811	2.02E-04	3.58E-02	regulation of hemopoiesis
GO:0043043	110	0.391	2.12E-04	3.68E-02	peptide biosynthetic process
GO:0036230	59	0.529	2.25E-04	3.74E-02	granulocyte activation
GO:0042119	59	0.529	2.25E-04	3.74E-02	neutrophil activation
GO:0048534	37	0.646	2.34E-04	3.74E-02	hematopoietic or lymphoid organ development
GO:0030097	37	0.646	2.34E-04	3.74E-02	hemopoiesis
GO:0002520	37	0.646	2.34E-04	3.74E-02	immune system development
GO:0002263	62	0.512	2.48E-04	3.82E-02	cell activation involved in immune response
GO:0002366	62	0.512	2.48E-04	3.82E-02	leukocyte activation involved in immune response
GO:0044271	236	0.298	2.53E-04	3.82E-02	cellular nitrogen compound biosynthetic process
GO:0050794	344	0.278	2.58E-04	3.82E-02	regulation of cellular process
GO:0010556	111	0.389	2.59E-04	3.82E-02	regulation of macromolecule biosynthetic process
GO:0045727	11	1.008	2.86E-04	4.06E-02	positive regulation of translation
GO:0048519	290	0.281	2.91E-04	4.06E-02	negative regulation of biological process
GO:0061013	22	0.763	3.00E-04	4.06E-02	regulation of mRNA catabolic process

GO:0043488	22	0.763	3.00E-04	4.06E-02	regulation of mRNA stability
GO:0043487	22	0.763	3.00E-04	4.06E-02	regulation of RNA stability
GO:0043299	59	0.517	3.01E-04	4.06E-02	leukocyte degranulation
GO:0019882	26	0.738	3.20E-04	4.21E-02	antigen processing and presentation
GO:0048002	26	0.738	3.20E-04	4.21E-02	antigen processing and presentation of peptide antigen
GO:0050867	14	1.012	3.34E-04	4.28E-02	positive regulation of cell activation
GO:0002696	14	1.012	3.34E-04	4.28E-02	positive regulation of leukocyte activation
G2					
GO term	**N**	**LOR**	**p-value**	**padj**	**name**
GO:0019080	70	0.868	2.58E-10	1.08E-06	viral gene expression
GO:0019083	70	0.868	2.58E-10	1.08E-06	viral transcription
GO:0000956	70	0.801	5.19E-09	1.09E-05	nuclear-transcribed mRNA catabolic process
GO:0000184	70	0.801	5.19E-09	1.09E-05	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay
GO:0006413	76	0.753	1.48E-08	2.09E-05	translational initiation
GO:0006401	81	0.729	1.50E-08	2.09E-05	RNA catabolic process
GO:0006402	80	0.717	2.76E-08	3.30E-05	mRNA catabolic process
GO:0006614	71	0.752	3.69E-08	3.86E-05	SRP-dependent cotranslational protein targeting to membrane
GO:0006518	111	0.596	1.27E-07	1.02E-04	peptide metabolic process
GO:0006613	73	0.706	1.38E-07	1.02E-04	cotranslational protein targeting to membrane
GO:0006612	73	0.706	1.38E-07	1.02E-04	protein targeting to membrane
GO:0016071	91	0.647	1.45E-07	1.02E-04	mRNA metabolic process
GO:0006412	90	0.645	1.74E-07	1.11E-04	translation
GO:0072599	74	0.691	2.11E-07	1.11E-04	establishment of protein localization to endoplasmic reticulum
GO:0045047	74	0.691	2.11E-07	1.11E-04	protein targeting to ER
GO:0090150	79	0.677	2.12E-07	1.11E-04	establishment of protein localization to membrane
GO:0070972	78	0.662	3.09E-07	1.43E-04	protein localization to endoplasmic reticulum
GO:0043043	91	0.626	3.15E-07	1.43E-04	peptide biosynthetic process
GO:0034655	91	0.626	3.23E-07	1.43E-04	nucleobase-containing compound catabolic process
GO:0044265	114	0.555	5.91E-07	2.47E-04	cellular macromolecule catabolic process

GO:0072657	90	0.585	1.77E-06	7.07E-04	protein localization to membrane
GO:0046700	100	0.533	4.27E-06	1.63E-03	heterocycle catabolic process
GO:1902600	29	0.881	5.59E-06	2.04E-03	proton transmembrane transport
GO:0006605	102	0.505	8.77E-06	3.06E-03	protein targeting
GO:0002181	27	0.891	1.10E-05	3.70E-03	cytoplasmic translation
GO:0019439	112	0.476	1.40E-05	4.52E-03	aromatic compound catabolic process
GO:0044270	102	0.493	1.56E-05	4.85E-03	cellular nitrogen compound catabolic process
GO:0072594	101	0.490	1.71E-05	5.13E-03	establishment of protein localization to organelle
GO:0016032	115	0.467	1.82E-05	5.25E-03	viral process
GO:0042274	16	1.154	1.93E-05	5.40E-03	ribosomal small subunit biogenesis
GO:0010605	174	0.397	2.02E-05	5.42E-03	negative regulation of macromolecule metabolic process
GO:0016070	149	0.418	2.07E-05	5.42E-03	RNA metabolic process
GO:0043604	100	0.485	2.91E-05	7.39E-03	amide biosynthetic process
GO:0044403	120	0.446	3.06E-05	7.55E-03	symbiotic process
GO:0009892	189	0.373	3.72E-05	8.90E-03	negative regulation of metabolic process
GO:1901361	117	0.433	5.08E-05	1.18E-02	organic cyclic compound catabolic process
GO:0016072	23	0.855	6.76E-05	1.48E-02	rRNA metabolic process
GO:0006364	23	0.855	6.76E-05	1.48E-02	rRNA processing
GO:0033365	113	0.430	6.88E-05	1.48E-02	protein localization to organelle
GO:0042254	35	0.715	8.26E-05	1.73E-02	ribosome biogenesis
GO:0046907	164	0.367	9.75E-05	1.99E-02	intracellular transport
GO:0009057	129	0.399	1.03E-04	2.06E-02	macromolecule catabolic process
GO:0043603	133	0.394	1.09E-04	2.09E-02	cellular amide metabolic process
GO:0090304	156	0.371	1.10E-04	2.09E-02	nucleic acid metabolic process
GO:0098655	49	0.585	1.52E-04	2.73E-02	cation transmembrane transport
GO:0098662	49	0.585	1.52E-04	2.73E-02	inorganic cation transmembrane transport
GO:0010468	177	0.347	1.53E-04	2.73E-02	regulation of gene expression
GO:0044419	172	0.346	1.88E-04	3.29E-02	interspecies interaction between organisms
GO:0034470	27	0.733	2.10E-04	3.52E-02	ncRNA processing

GO:0071840	248	0.310	2.10E-04	3.52E-02	cellular component organization or biogenesis
GO:0010629	128	0.378	2.32E-04	3.81E-02	negative regulation of gene expression
GO:0044271	198	0.325	2.37E-04	3.82E-02	cellular nitrogen compound biosynthetic process
GO:0006396	40	0.616	2.59E-04	4.09E-02	RNA processing
GO:0010467	199	0.320	2.83E-04	4.40E-02	gene expression

*The LOR (Logarithm of Ratio) refer to LRFI group, this value categorizes each biological process in a classification in relation to the rest (total), allowing to identify biological processes, molecular mechanisms and significant cellular components and their regulation within the groups.

Table S4A. Biological process GO terms and pathways containing genes with differentially expressed alternatively spliced transcripts (DAS) in liver tissue of Nelore bulls selected to be divergent for RFI in G1 and G2 used for analysis in GSEA software.

G1				
Go term	**N**	**LOR**	**p-value**	**biological Processes**
GO:0045321	13	1.025	0.006712	leukocyte activation
GO:0002274	12	1.027	0.008079	myeloid leukocyte activation
GO:0044282	10	-0.994	0.015600	small molecule catabolic process
GO:0036230	10	0.964	0.016459	granulocyte activation
GO:0042119	10	0.964	0.016459	neutrophil activation
GO:0065007	40	0.640	0.037370	biological regulation
GO:0002446	10	0.825	0.037387	neutrophil mediated immunity
GO:0048513	11	0.770	0.040254	animal organ development
GO:0007275	16	0.656	0.045203	multicellular organism development
GO:0048731	16	0.656	0.045203	system development
GO:0002252	18	0.620	0.049619	immune effector process
G2				
Go term	**N**	**LOR**	**p-value**	**biological Processes**
GO:0019637	12	-1.039	0.009403	organophosphate metabolic process
GO:1901135	14	-0.894	0.012560	carbohydrate derivative metabolic process
GO:0055086	12	-0.933	0.014009	nucleobase-containing small molecule metabolic process
GO:0006082	12	-1.009	0.015563	organic acid metabolic process

GO:0043436	12	-1.009	0.015563	oxoacid metabolic process
GO:0019752	11	-1.008	0.020698	carboxylic acid metabolic process
GO:0051179	44	0.618	0.022510	localization
GO:1901362	17	-0.697	0.026245	organic cyclic compound biosynthetic process
GO:0044281	22	-0.666	0.028773	small molecule metabolic process
GO:0002252	10	0.850	0.032386	immune effector process
GO:0006810	42	0.563	0.032994	transport
GO:0051234	42	0.563	0.032994	establishment of localization
GO:0018130	15	-0.695	0.033486	heterocycle biosynthetic process
GO:0051641	32	0.562	0.034009	cellular localization
GO:0072657	19	0.612	0.040637	protein localization to membrane
GO:0044419	29	0.539	0.045135	interspecies interaction between organisms
GO:0019080	17	0.610	0.047187	viral gene expression
GO:0019083	17	0.610	0.047187	viral transcription
GO:0000184	16	0.622	0.048237	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay
GO:0000956	16	0.622	0.048237	nuclear-transcribed mRNA catabolic process
GO:0006605	16	0.622	0.048237	protein targeting
GO:0006612	16	0.622	0.048237	protein targeting to membrane
GO:0006613	16	0.622	0.048237	cotranslational protein targeting to membrane
GO:0006614	16	0.622	0.048237	SRP-dependent cotranslational protein targeting to membrane
GO:0045047	16	0.622	0.048237	protein targeting to ER
GO:0070972	16	0.622	0.048237	protein localization to endoplasmic reticulum
GO:0072599	16	0.622	0.048237	establishment of protein localization to endoplasmic reticulum

Table S4B. Biological process GO terms and pathways containing genes with differentially expressed alternatively spliced transcripts (DAS) in liver tissue of Nelore bulls selected to be divergent for RFI in G1 and G2 used for analysis in DAVID software.

G1				
Term	**N**	**p-value**	**Genes**	**FDR**
GO:0070062~extracellular exosome	36	1.36E-15	CRP, ITIH4, SPON2, ITIH3, ACAA2, SDC4, PROS1, RPL12, LOC783680, A1BG, CPN2, ACTG1, GNAI2, SCPEP1, SYNGR2, APOM, LAMP2, PCBD1, BOLA-DRA, HNRNPA1, B2M, HAO2, CTSC, CD74, IAH1, MIF, F2, FAH, ASS1, LRG1, TCN2, ARPC2, NPC2, PSME2, PABPC1, C1QC	1.17E-13
Acetylation	26	2.96E-11	CSTB, ACAA2, RPL12, COX5B, HSD17B10, AIF1, ACTG1, HMGCL, SYNGR2, ASL, PCBD1, HNRNPA1, HMGN2, TMSB10, H3F3B, IDH1, MIF, FAH, ARPC2, NPC2, PSME2, PABPC1, HPD, RPS23, OTC	3.52E-09
GO:0072562~blood microparticle	8	1.49E-08	ITIH4, PROS1, APOE, A1BG, F2, CPN2, C1QC, ACTG1	6.42E-07
Secreted	16	2.51E-07	ITIH4, FGA, ITIH3, PROS1, LOC783680, A1BG, MIF, F2, TCN2, NPC2, CCL5, APOE, SERPINA3-8, B2M, SERPINA3-7, C1QC	1.49E-05
Protease inhibitor	6	2.22E-06	CSTB, ITIH4, ITIH3, SERPINA3-8, SERPINA3-7	8.79E-05
Signal	26	4.38E-06	CRP, ITIH4, SPON2, ITIH3, SDC4, PROS1, LOC783680, A1BG, CPN2, CXCL16, SCPEP1, APOM, CCL5, LAMP2, BOLA-DRA, APOE, B2M, CTSC, FGA, F2, LRG1, TCN2, NPC2, LOC515150, SERPINA3-8, SERPINA3-7	1.30E-04
GO:0010951~negative regulation of endopeptidase activity	6	4.59E-06	CSTB, ITIH4, ITIH3, SERPINA3-8, SERPINA3-7	1.52E-03
binding site:Substrate	9	9.27E-06	HMGCL, IAH1, IDH1, QPRT, MIF, FAH, HSD17B10, SAT1, HAO2	1.41E-03
Arginine biosynthesis	3	3.69E-05	ASL, ASS1, OTC	8.77E-04
Urea cycle	3	1.22E-04	ASL, ASS1, OTC	2.43E-03
Glycoprotein	14	1.71E-04	ITIH4, FGA, ITIH3, SDC4, PROS1, A1BG, F2, CXCL16, TCN2, NPC2, APOE, SERPINA3-8, SERPINA3-7, CTSC	2.90E-03
Serine protease inhibitor	4	2.33E-04	ITIH4, ITIH3, SERPINA3-8, SERPINA3-7	3.47E-03
GO:0000050~urea cycle	3	3.58E-04	ASL, ASS1, OTC	5.91E-02
bta04612:Antigen processing and presentation	5	7.56E-04	CD74, PSME2, LOC783680, BOLA-DRA, B2M	5.05E-02
bta01130:Biosynthesis of antibiotics	7	8.86E-04	ACAA2, IDH1, ASL, HSD17B10, HAO2, ASS1, OTC	5.05E-02
GO:0005615~extracellular space	12	9.49E-04	CRP, SPON2, LRG1, CCL5, LAMP2, LOC783680, MIF, SERPINA3-8, B2M, SERPINA3-7, CTSC, CXCL16	2.72E-02
Disulfide bond	15	1.52E-03	ITIH4, FGA, PROS1, LOC783680, A1BG, F2, CXCL16, HMGCL, TCN2, NPC2, CCL5, LOC515150, BOLA-DRA, B2M, CTSC	2.01E-02
GO:0001895~retina homeostasis	3	2.00E-03	LOC783680, B2M, ACTG1	2.20E-01
Immunity	5	2.48E-03	FGA, LOC783680, BOLA-DRA, MIF, B2M	2.94E-02
Amino-acid biosynthesis	3	2.72E-03	ASL, ASS1, OTC	2.94E-02
GO:0071222~cellular response to	4	3.14E-03	SPON2, LOC783680, B2M, CXCL16	2.59E-01

lipopolysaccharide				
GO:0033344~cholesterol efflux	3	4.50E-03	APOM, NPC2, APOE	2.97E-01
bta00220:Arginine biosynthesis	3	4.87E-03	ASL, ASS1, OTC	1.80E-01
Hemostasis	3	5.71E-03	FGA, PROS1, F2	5.22E-02
Blood coagulation	3	5.71E-03	FGA, PROS1, F2	5.22E-02
GO:0004867~serine-type endopeptidase inhibitor activity	4	5.71E-03	ITIH4, ITIH3, SERPINA3-8, SERPINA3-7	4.52E-01
bta01230:Biosynthesis of amino acids	4	7.04E-03	IDH1, ASL, ASS1, OTC	1.80E-01
signal peptide	15	7.22E-03	ITIH4, FGA, ITIH3, PROS1, LOC783680, A1BG, F2, CXCL16, TCN2, NPC2, CCL5, BOLA-DRA, APOE, B2M, CTSC	5.49E-01
IPR015707:Beta-2-Microglobulin	2	7.37E-03	LOC783680, B2M	6.46E-01
GO:0051262~protein tetramerization	3	7.44E-03	HMGCL, CCL5, ASL	3.45E-01
GO:0005925~focal adhesion	6	7.74E-03	ARPC2, RPL12, LOC783680, PABPC1, B2M, ACTG1	1.66E-01
bta04610:Complement and coagulation cascades	4	7.89E-03	FGA, PROS1, F2, C1QC	1.80E-01
GO:0005102~receptor binding	4	8.22E-03	FGA, HMGCL, F2, HAO2	4.52E-01
GO:0042450~arginine biosynthetic process via ornithine	2	8.37E-03	ASL, OTC	3.45E-01
IPR013785:Aldolase-type TIM barrel	3	9.55E-03	HMGCL, QPRT, HAO2	6.46E-01
Tyrosine catabolism	2	1.06E-02	FAH, HPD	8.99E-02
GO:0006572~tyrosine catabolic process	2	1.25E-02	FAH, HPD	3.45E-01
GO:1904437~positive regulation of transferrin receptor binding	2	1.25E-02	LOC783680, B2M	3.45E-01
GO:1904434~positive regulation of ferrous iron binding	2	1.25E-02	LOC783680, B2M	3.45E-01
GO:2000343~positive regulation of chemokine (C-X-C motif) ligand 2 production	2	1.25E-02	CD74, MIF	3.45E-01
GO:1903991~positive regulation of ferrous iron import into cell	2	1.25E-02	LOC783680, B2M	3.45E-01
GO:0070053~thrombospondin receptor activity	2	1.28E-02	SDC4, F2	4.68E-01
chain:Beta-2-microglobulin	2	1.40E-02	LOC783680, B2M	6.04E-01
Phenylalanine catabolism	2	1.41E-02	FAH, HPD	1.12E-01

IPR001713:Proteinase inhibitor I25A, stefin A	2	1.47E-02	CSTB	6.46E-01
GO:0045087~innate immune response	5	1.50E-02	FGA, SPON2, LOC783680, MIF, B2M	3.82E-01
IPR003006:Immunoglobulin/major histocompatibility complex, conserved site	3	1.52E-02	LOC783680, BOLA-DRA, B2M	6.46E-01
GO:0002906~negative regulation of mature B cell apoptotic process	2	1.67E-02	CD74, MIF	3.93E-01
Cytoplasm	13	1.68E-02	CSTB, ARPC2, IDH1, PCBD1, PABPC1, MIF, HNRNPA1, HMGN2, TMSB10, AIF1, SAT1, ACTG1	1.19E-01
Phosphoprotein	17	1.70E-02	ACAA2, H3F3B, RPL12, IDH1, QPRT, FAH, AIF1, ASS1, SYNGR2, PSME2, APOE, PABPC1, HNRNPA1, HMGN2, TMSB10, HPD, OTC	1.19E-01
IPR010600:Inter-alpha-trypsin inhibitor heavy chain, C-terminal	2	1.83E-02	ITIH4, ITIH3	6.46E-01
GO:0042802~identical protein binding	4	2.00E-02	PCBD1, B2M, ASS1, ACTG1	5.49E-01
bta00350:Tyrosine metabolism	3	2.06E-02	MIF, FAH, HPD	3.92E-01
GO:0006526~arginine biosynthetic process	2	2.08E-02	ASL, ASS1	4.20E-01
SM00407:IGc1	3	2.20E-02	LOC783680, BOLA-DRA, B2M	4.42E-01
GO:0008283~cell proliferation	4	2.29E-02	CD74, H3F3B, MIF, GNAI2	4.20E-01
SM00609:VIT	2	2.46E-02	ITIH4, ITIH3	4.42E-01
IPR003597:Immunoglobulin C1-set	3	2.47E-02	LOC783680, BOLA-DRA, B2M	6.46E-01
GO:0030212~hyaluronan metabolic process	2	2.49E-02	ITIH4, ITIH3	4.20E-01
GO:0060907~positive regulation of macrophage cytokine production	2	2.49E-02	CD74, SPON2	4.20E-01
GO:0006559~L-phenylalanine catabolic process	2	2.49E-02	FAH, HPD	4.20E-01
bta05142:Chagas disease (American trypanosomiasis)	4	2.51E-02	CCL5, LOC525415, C1QC, GNAI2	4.09E-01
disulfide bond	12	2.52E-02	FGA, ITIH4, TCN2, NPC2, PROS1, CCL5, LOC783680, A1BG, F2, B2M, CTSC, CXCL16	6.04E-01
IPR018073:Proteinase inhibitor I25, cystatin, conserved site	2	2.56E-02	CSTB	6.46E-01
GO:1990712~HFE-transferrin receptor complex	2	2.69E-02	LOC783680, B2M	4.62E-01
domain:VIT	2	2.78E-02	ITIH4, ITIH3	6.04E-01
short sequence motif:Secondary area of contact	2	2.78E-02	CSTB	6.04E-01

site:Reactive site	2	2.78E-02	CSTB	6.04E-01
GO:0003254~regulation of membrane depolarization	2	2.90E-02	LOC783680, B2M	4.20E-01
GO:0071281~cellular response to iron ion	2	2.90E-02	LOC783680, B2M	4.20E-01
bta00280:Valine, leucine and isoleucine degradation	3	3.13E-02	HMGCL, ACAA2, HSD17B10	4.14E-01
bta01100:Metabolic pathways	13	3.27E-02	ACAA2, IDH1, QPRT, COX5B, FAH, HSD17B10, SAT1, ASS1, HMGCL, ASL, HAO2, HPD, OTC	4.14E-01
IPR013694:VIT domain	2	3.27E-02	ITIH4, ITIH3	7.24E-01
GO:0034380~high-density lipoprotein particle assembly	2	3.31E-02	APOM, APOE	4.20E-01
GO:1900121~negative regulation of receptor binding	2	3.31E-02	LOC783680, B2M	4.20E-01
GO:0001916~positive regulation of T cell mediated cytotoxicity	2	3.31E-02	LOC783680, B2M	4.20E-01
GO:0019885~antigen processing and presentation of endogenous peptide antigen via MHC class I	2	3.31E-02	LOC783680, B2M	4.20E-01
GO:0002481~antigen processing and presentation of exogenous protein antigen via MHC class Ib, TAP-dependent	2	3.31E-02	LOC783680, B2M	4.20E-01
site:Cleavage; by thrombin	2	3.46E-02	PROS1, F2	6.58E-01
Zymogen	3	3.64E-02	PROS1, F2, CTSC	2.40E-01
GO:1902166~negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	2	3.71E-02	CD74, MIF	4.38E-01
GO:0002726~positive regulation of T cell cytokine production	2	3.71E-02	LOC783680, B2M	4.38E-01
GO:0042583~chromaffin granule	2	3.81E-02	SERPINA3-8, SERPINA3-7	5.30E-01
Gamma-carboxyglutamic acid	2	3.82E-02	PROS1, F2	2.40E-01
GO:0051258~protein polymerization	2	4.12E-02	FGA, F2	4.69E-01
domain:Ig-like C1-type	2	4.14E-02	LOC783680, B2M	6.99E-01
GO:0016597~amino acid binding	2	4.20E-02	ASS1, OTC	9.23E-01
IPR017857:Coagulation factor, subgroup, Gla domain	2	4.34E-02	PROS1, F2	8.54E-01

Oxidation	2	4.50E-02	APOE, ACTG1	2.44E-01
MHC I	2	4.50E-02	LOC783680, B2M	2.44E-01
Acute phase	2	4.50E-02	ITIH4, F2	2.44E-01
GO:0042026~protein refolding	2	4.52E-02	LOC783680, B2M	4.81E-01
GO:0043691~reverse cholesterol transport	2	4.52E-02	APOM, APOE	4.81E-01
SM00069:GLA	2	4.86E-02	PROS1, F2	5.58E-01
GO:0034361~very-low-density lipoprotein particle	2	4.93E-02	APOM, APOE	5.30E-01
GO:0042612~MHC class I protein complex	2	4.93E-02	LOC783680, B2M	5.30E-01
G2				
Term	**N**	**p-value**	**Genes**	**FDR**
Acetylation	33	1.97E-16	AHCY, COX4I1, PARK7, YBX1, COX7C, UQCRH, ACTB, HINT2, RACK1, SH3BGRL3, HMGN2, TMSB10, RPS13, GAMT, NDUFA7, RPS9, TPI1, H3F3B, H3F3A, RPSA, COX6B1, FABP1, RPS25, MYL6, RPS29, NPC2, RPL27, CHPT1, RPL29, CALM1, RPS21, RPL28, GAPDH	2.32E-14
Ribosomal protein	15	2.55E-15	RPS9, RPSA, RPS25, RPS14, RPS17, RPS28, RPS16, RPL18A, RPS29, RPL27, RPL29, RPS21, RPL28, RPL39, RPS13	1.51E-13
bta03010:Ribosome	16	2.16E-14	RPS9, RPL34, RPSA, RPS25, RPS14, RPS17, RPS28, RPS16, RPL18A, RPS29, RPL27, RPL29, RPS21, RPL28, RPL39, RPS13	2.53E-12
GO:0003735~structural constituent of ribosome	16	8.51E-14	NDUFA7, RPS9, RPL34, RPSA, RPS14, RPS17, RPS28, RPS16, RPL18A, RPS29, RPL27, RPL29, RPS21, RPL28, RPL39, RPS13	1.07E-11
Ribonucleoprotein	15	1.39E-13	RPS9, RPSA, RPS25, RPS14, RPS17, RPS28, RPS16, RPL18A, RPS29, RPL27, RPL29, RPS21, RPL28, RPL39, RPS13	5.45E-12
GO:0006412~translation	14	8.21E-13	RPS9, RPL34, RPSA, RPS14, RPS17, RPS28, RPS16, RPL18A, RPS29, RPL27, RPS21, RPL28, RPL39, RPS13	2.71E-10
GO:0022627~cytosolic small ribosomal subunit	9	4.32E-12	RPS25, RPS17, RPS28, RPS9, RPS16, RPS29, RPSA, RPS21, RPS13	4.06E-10
GO:0070062~extracellular exosome	33	1.19E-11	SPON2, C1QA, AHCY, RPL34, COX4I1, CTSZ, PARK7, UQCR10, YBX1, ACTB, RPS16, RACK1, BOLA-DRA, B2M, RPS13, GAMT, CD74, RPS9, TPI1, ABHD14B, FABP1, RPS25, GRHPR, MYL6, RPS28, TCN2, RPS29, NPC2, CALM1, RPL28, GAPDH, ITM2B, C1QC	5.58E-10
GO:1902600~hydrogen ion transmembrane transport	6	3.90E-07	COX4I1, UQCR10, COX6A1, COX7C, UQCRH, COX6B1	6.43E-05
Ubl conjugation	13	8.69E-07	DAZAP2, TPI1, H3F3B, H3F3A, PARK7, YBX1, ACTB, RPS17, RPL18A, HMGN2, CALM1, RPS21, GAPDH	2.56E-05
GO:0022625~cytosolic large ribosomal subunit	6	6.54E-06	RPL18A, RPL34, RPL27, RPL29, RPL28, RPL39	2.05E-04
Isopeptide bond	10	8.88E-06	FABP1, RPS17, TPI1, RPL18A, PARK7, YBX1, RPL29, HMGN2, CALM1, RPS21	2.10E-04

Phosphoprotein	25	2.33E-05	AHCY, COX4I1, PARK7, YBX1, RPS17, RPL18A, RACK1, HMGN2, TMSB10, RPS13, RPS9, TPI1, H3F3B, H3F3A, ABHD14B, RPSA, FABP1, MYL6, RPS28, RPS29, PPDPF, RPL29, CALM1, RPL28, GAPDH	4.59E-04
bta05010:Alzheimer's disease	9	4.41E-05	NDUFA7, COX4I1, UQCR10, COX6A1, CALM1, COX7C, GAPDH, UQCRH, COX6B1	2.58E-03
bta05012:Parkinson's disease	8	1.18E-04	NDUFA7, COX4I1, PARK7, UQCR10, COX6A1, COX7C, UQCRH, COX6B1	4.62E-03
bta04932:Non-alcoholic fatty liver disease (NAFLD)	8	1.63E-04	NDUFA7, RXRA, COX4I1, UQCR10, COX6A1, COX7C, UQCRH, COX6B1	4.77E-03
Methylation	8	2.94E-04	TPI1, H3F3B, RPS29, H3F3A, RPL29, CALM1, GAPDH, ACTB	4.96E-03
bta04260:Cardiac muscle contraction	6	2.95E-04	COX4I1, UQCR10, COX6A1, COX7C, UQCRH, COX6B1	6.90E-03
GO:0004129~cytochrome-c oxidase activity	4	3.50E-04	COX4I1, COX6A1, COX7C, COX6B1	2.20E-02
bta00190:Oxidative phosphorylation	7	5.53E-04	NDUFA7, COX4I1, UQCR10, COX6A1, COX7C, UQCRH, COX6B1	1.03E-02
bta05016:Huntington's disease	8	6.17E-04	NDUFA7, COX4I1, UQCR10, COX6A1, COX7C, UQCRH, COX6B1, POLR2L	1.03E-02
Mitochondrion inner membrane	6	9.10E-04	NDUFA7, COX4I1, UQCR10, COX6A1, COX7C, UQCRH	1.34E-02
bta01100:Metabolic pathways	20	9.66E-04	GAMT, NDUFA7, TPI1, AHCY, ACSL1, AMY2B, COX4I1, UQCR10, COX6A1, COX7C, SAT1, UQCRH, COX6B1, GRHPR, CPS1, STT3A, CHPT1, GAPDH, SLC27A5, POLR2L	1.41E-02
GO:0006123~mitochondrial electron transport, cytochrome c to oxygen	3	1.35E-03	COX4I1, COX6A1, COX7C	1.49E-01
GO:0005751~mitochondrial respiratory chain complex IV	3	1.53E-03	COX4I1, COX6A1, COX7C	3.60E-02
GO:0045277~respiratory chain complex IV	3	2.01E-03	COX4I1, COX7C, COX6B1	3.78E-02
GO:0000028~ribosomal small subunit assembly	3	2.13E-03	RPS17, RPS28, RPSA	1.76E-01
Cytoplasm	16	2.19E-03	DAZAP2, RPS9, AHCY, ABHD14B, RPSA, PARK7, YBX1, SAT1, ACTB, FABP1, RACK1, SH3BGRL3, HMGN2, CALM1, TMSB10, GAPDH	2.88E-02
GO:0005840~ribosome	4	2.69E-03	RPS25, RPS14, RPS17, RPL34	4.21E-02
Mitochondrion	9	3.65E-03	NDUFA7, HINT2, COX4I1, PARK7, UQCR10, COX6A1, COX7C, UQCRH, COX6B1	4.31E-02
GO:0007566~embryo implantation	3	5.91E-03	RXRA, H3F3B, H3F3A	3.05E-01
Secreted	10	6.12E-03	CFD, CST3, C1QA, TCN2, NPC2, LECT2, YBX1, B2M, ITM2B, C1QC	6.57E-02
GO:0009060~aerobic respiration	3	6.38E-03	UQCR10, COX6A1, UQCRH	3.05E-01
GO:0031492~nucleosomal DNA binding	3	7.86E-03	H3F3B, H3F3A, HMGN2	3.30E-01
GO:0090230~regulation of centromere complex assembly	2	9.24E-03	H3F3B, H3F3A	3.05E-01
GO:0031509~telomeric heterochromatin assembly	2	9.24E-03	H3F3B, H3F3A	3.05E-01

GO:1902340~negative regulation of chromosome condensation	2	9.24E-03	H3F3B, H3F3A	3.05E-01
GO:0000461~endonucleolytic cleavage to generate mature 3'-end of SSU-rRNA from (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	2	9.24E-03	RPSA, RPS21	3.05E-01
bta05150:Staphylococcus aureus infection	4	9.37E-03	CFD, C1QA, BOLA-DRA, C1QC	1.22E-01
Disulfide bond	14	9.67E-03	CFD, VKORC1, C1QA, LECT2, CTSZ, UQCRH, COX6B1, CST3, TCN2, NPC2, LOC510860, BOLA-DRA, B2M, ITM2B	9.50E-02
GO:0005829~cytosol	11	1.01E-02	GRHPR, MYL6, TPI1, AHCY, RACK1, ABHD14B, PARK7, CALM1, GAPDH, SAT1, ACTB	1.27E-01
GO:0007338~single fertilization	3	1.08E-02	H3F3B, H3F3A, PARK7	3.23E-01
GO:0005925~focal adhesion	6	1.08E-02	RPS9, RPS16, RPS29, B2M, ACTB, RPS13	1.27E-01
GO:0051287~NAD binding	3	1.16E-02	GRHPR, AHCY, GAPDH	3.65E-01
SM00152:THY	2	1.37E-02	TMSB4X, TMSB10	4.65E-01
bta03320:PPAR signaling pathway	4	1.49E-02	FABP1, RXRA, ACSL1, SLC27A5	1.74E-01
IPR001152:Thymosin beta-4	2	1.66E-02	TMSB4X, TMSB10	1.00E+00
GO:0031508~pericentric heterochromatin assembly	2	1.84E-02	H3F3B, H3F3A	4.67E-01
GO:0051343~positive regulation of cyclic-nucleotide phosphodiesterase activity	2	1.84E-02	RACK1, CALM1	4.67E-01
GO:0001740~Barr body	2	2.08E-02	H3F3B, H3F3A	2.18E-01
GO:0044822~poly(A) RNA binding	10	2.22E-02	RPS25, RPS28, RPS9, RPS16, RPL18A, RACK1, YBX1, RPL29, HMGN2, RPL28	5.61E-01
Respiratory chain	3	2.27E-02	NDUFA7, UQCR10, UQCRH	2.04E-01
Immunity	4	2.42E-02	CFD, C1QA, BOLA-DRA, B2M	2.04E-01
PIRSF001828:thymosin beta	2	2.74E-02	TMSB4X, TMSB10	3.83E-01
GO:0060907~positive regulation of macrophage cytokine production	2	2.75E-02	CD74, SPON2	6.04E-01
GO:0000447~endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	2	2.75E-02	RPSA, RPS21	6.04E-01
GO:0030529~intracellular ribonucleoprotein complex	3	2.80E-02	RPS9, YBX1, GAPDH	2.63E-01

GO:0008121~ubiquinol-cytochrome-c reductase activity	2	3.42E-02	UQCR10, UQCRH	7.19E-01
GO:0016020~membrane	9	3.46E-02	MYL6, RPS9, RPS16, RPL18A, ACSL1, RPSA, RPL29, RPL28, RPS13	2.96E-01
GO:0098609~cell-cell adhesion	3	3.85E-02	RACK1, PARK7, RPL29	7.94E-01
Transit peptide	5	4.02E-02	HINT2, COX4I1, COX6A1, COX7C, UQCRH	3.16E-01
GO:0005750~mitochondrial respiratory chain complex III	2	4.13E-02	UQCR10, UQCRH	3.23E-01
GO:0098641~cadherin binding involved in cell-cell adhesion	3	4.13E-02	RACK1, PARK7, RPL29	7.44E-01
bta05322:Systemic lupus erythematosus	5	4.25E-02	C1QA, H3F3B, H3F3A, BOLA-DRA, C1QC	4.52E-01
Zymogen	3	4.30E-02	CFD, CTSZ, PARK7	3.17E-01
chain:Histone H3.3	2	4.39E-02	H3F3B, H3F3A	1.00E+00
Electron transport	3	4.67E-02	NDUFA7, UQCR10, UQCRH	3.23E-01
Oxidation	2	4.92E-02	PARK7, ACTB	3.23E-01

Table S5. Biological process GO terms and pathways containing for nine common genes differentially expressed alternatively spliced transcripts (DAS) in liver tissue of Nelore bulls selected to be divergent for RFI in G1 and G2 used for analysis in DAVID software.

Term	**N**	**p-value**	**Genes**	**Bonferroni**	**Benjamini**	**FDR**
bta04612:Antigen processing and presentation	3	0.00142	CD74, BOLA-DRA, B2M	0.04047	0.04128	0.04128
bta05322:Systemic lupus erythematosus	3	0.00779	H3F3B, BOLA-DRA, C1QC	0.20290	0.11295	0.11295
GO:0042613~MHC class II protein complex	2	0.00979	CD74, BOLA-DRA	0.24081	0.23369	0.23369
GO:0031492~nucleosomal DNA binding	2	0.01435	H3F3B, HMGN2	0.18316	0.20086	0.20086
GO:0070062~extracellular exosome	5	0.01669	CD74, NPC2, BOLA-DRA, B2M, C1QC	0.37583	0.23369	0.23369
IPR003006:Immunoglobulin/major histocompatibility complex, conserved site	2	0.02403	BOLA-DRA, B2M	0.42849	0.35653	0.35653
Acetylation	4	0.02704	H3F3B, NPC2, HMGN2, TMSB10	0.59533	0.76638	0.76638

Ethics approval and consent to participate

All experimental procedures were approved by Ethics Committee of the School of Agricultural and Veterinarian Sciences, São Paulo State University (UNESP), Jaboticabal, SP, Brazil (protocol number 18.340/16).

Funding

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and São Paulo Research Foundation – FAPESP (grant #2009/16118-5, #2017/10630-2 e #2018/20026-8).

Acknowledgements

The authors thank the Qualitas Nelore breeding program company for providing the tissue samples and database used in this study.

CAPÍTULO 5 – Considerações Finais

De acordo com os resultados anteriormente apresentados nessa tese, ao comparar o transcriptoma do fígado de bovinos Nelore avaliados para consumo alimentar residual, pertencentes à dois grupos genéticos (G1 e G2), foram encontrados transcritos e *splicing* alternativos relacionados à: metabolismo energético, turnover proteico, homeostase redox e sistema imunológico.

Ao comprar os resultados dos capítulos 2, 3 e 4, (Figura 1) foram observados:

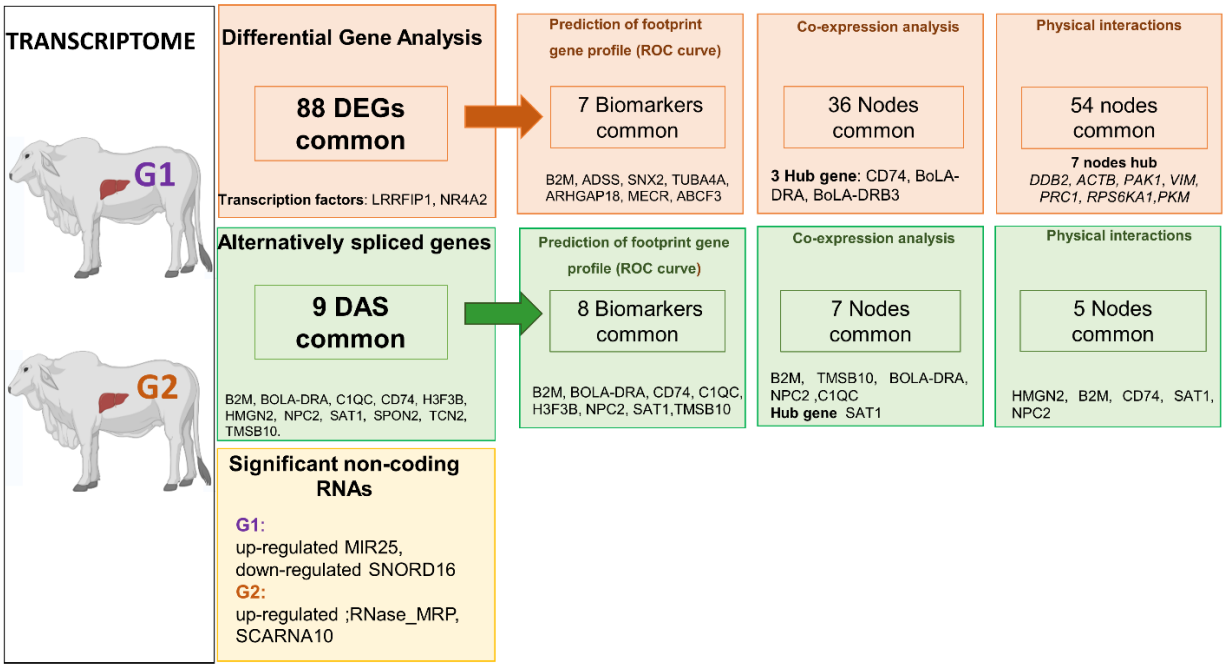


Figura 1. Síntese dos resultados obtidos nos capítulos 2 e 3 (em vermelho e amarello), 4 (em verde) dessa tese para os dois grupos genéticos (G1 e G2) associados à característica CAR em fígado para a raça Nelore.

❖ **Capítulo 2-3**: 88 DEGs comuns foram identificados nos dois grupos genéticos com a mesma orientação, destacando:

- Os fatores de transcrição *LRRFIP1* e *NR4A2*.
- Foram identificados 7 genes preditos como biomarcadores (curva ROC): *B2M, ADSS, SNX2, TUBA4A, ARHGAP18, MECR, ABCF3*.
- O gene *B2M* foi superexpresso no grupo mais eficiente. Esse gene participa da maioria dos processos celulares que diferenciam grupos com maior ou menor eficiência alimentar.
- Por meio da análise de co-expressão foram detectados 36 nós (*nodes*), sendo que 3 deles foram genes *hub*; *CD74, BoLA-DRA* and *BoLA-DRB3*. Também destacamos pela sua importância na característica seguintes os genes: *FYB1, TUBA4A, LCP1, LETMD1, SNX1, ANXA1, GRN, COTL1* e *CYBB*.

- Por meio da análise de interação física foram detectados 54 nós (com BioGrid), dos quais, destacamos 7, devido a sua ação regulatória sobre outros 15 genes cada um: *DDB2,* ACTB, PAK1, VIM, PRC1, RPS6KA1 *e PKM*.
- Os genes *TUBA4A, B2M* e *SNX2* foram comuns em 3 análises diferentes: foram considerados nós nas análises de interação física e co-expressão, além de terem sido detectados como biomarcadores.

❖ **Capítulo 4**: 9 genes que codificam transcritos alternativamente *spliced* (DAS) foram diferencialmente expressos e associados com RFI: *B2M, BOLA-DRA, C1QC, CD74, H3F3B, HMGN2, NPC2, SAT1, SPON2, TCN2 e TMSB10,* dentre eles:

- 8 genes foram preditos como biomarcadores (curva ROC): *B2M, BOLA-DRA, CD74, C1QC, H3F3B, NPC2, SAT1* e *TMSB10*.
- 6 genes foram considerados nós nas análises de co-expressão: *B2M, TMSB10, BOLA-DRA, NPC2, C1QC* e *SAT1*, sendo o último considerado *hub* gene.
- 5 genes foram considerados nós na análise de interações físicas (BioGrid): *HMGN2, B2M, CD74, SAT1* e *NPC2*.
- Os genes *SAT1, NPC2* e *B2M* foram comuns nas análises de interações físicas, co-expressão e biomarcador.

❖ Três elementos únicos (genes e *splincing* alternativos comuns para os dois grupos genéticos) foram diferencialmente expressos; *B2M, CD74* e *NPC2* (Figura 2).

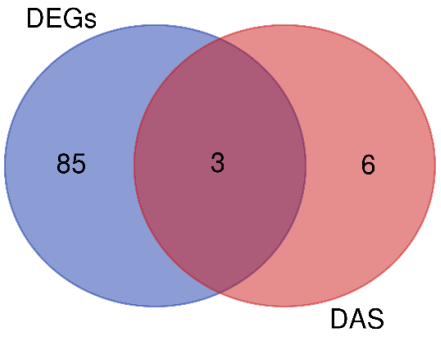


Figura 2. *Venn diagram*. Genes diferencialmente expressos (DEGs) e genes que codificam transcritos *spliced* alternativamente que foram diferencialmente expressos (DAS) em dois grupos genéticos de animais da raça Nelore com fenótipos divergentes para RFI.

❖ RNAs não codificantes foram significativos: MIR25 regulado positivamente e SNORD16 regulados negativamente em RLFI para G1. Para G2, RNase_MRP e SCARNA10 regulados positivamente. Destacamos o MIR25 sendo capaz de atuar bloqueando a citotoxicidade e o estresse oxidativo e o RMRP como bloqueador do dano mitocondrial.

As interações genéticas mostraram que os DEGs (Capítulo 2 e 3) e DAS (Capítulo 4) encontrados para CAR participam de processos biológicos e vias que estão intimamente relacionados ao metabolismo energético. Além disso, em todos os capítulos, foram encontrados genes que regulam o sistema imunológico, essencial para que os animais mantenham sua saúde.

A ativação do sistema imunológico utiliza muita energia, uma vez que a produção metabólica de calor e o consumo de oxigênio aumentam em 20-30% (JOHNSON et al., 2013). O metabolismo energético está fortemente relacionado com a diminuição da produtividade, e a resposta imunológica poderia regular o metabolismo energético (VERHAGEN, 1987). No gado bovino, vários estudos têm associado o sistema imune ao CAR (PARADIS et al., 2015; ZAREK et al., 2017).

O presente estudo contribui para a base de conhecimento sobre a regulação transcriptômica da variação da eficiência alimentar. Embora outros estudos tenham fornecido informações sobre a regulação do CAR através da fisiologia do fígado, os processos biológicos que governam as diferenças na eficiência alimentar não estão totalmente elucidados e mais pesquisas são necessárias. Nesse contexto, nosso trabalho se destaca pelo ineditismo da comparação de dois grupos genéticos e pelos genes comuns encontrados para ambos, por meio de diferentes análises, indicando a descoberta de novos biomarcadores que podem ser usados com o intuito de melhorar a eficiência alimentar do rebanho brasileiro.

Referencias

JOHNSON, K. A. et al. Differential expression of innate immune system genes in liver of beef cattle with divergent phenotypes for RFI. **Energy and protein metabolism and nutrition in sustainable animal production**, p. 371–372, 2013.

PARADIS, F. et al. Transcriptomic analysis by RNA sequencing reveals that hepatic interferon-induced genes may be associated with feed efficiency in beef heifers. **Journal of animal science**, v. 93, n. 7, p. 3331–3341, 2015.

VERHAGEN, J. M. F. Energy Metabolism and Immune Function. **Energy Metabolism in Farm Animals**, p. 291–303, 1987.

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.