

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**GENES DIFERENCIALMENTE EXPRESSOS E SPLICING
ALTERNATIVOS RELACIONADOS COM
CARACTERÍSTICAS DE CARÇAÇA E CARNE DE
BOVINOS DA RAÇA NELORE**

Danielly Beraldo dos Santos Silva
Biotecnologista

2019

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**GENES DIFERENCIALMENTE EXPRESSOS E SPLICING
ALTERNATIVOS RELACIONADOS COM
CARACTERÍSTICAS DE CARCAÇA E CARNE DE
BOVINOS DA RAÇA NELORE**

Discente: Danielly Beraldo dos Santos Silva

Orientadora: Profa. Dra. Lucia Galvão de Albuquerque

Co-Orientadores: Prof. Dr. Daniel Guariz Pinheiro

Profa. Dra. Maria Inês Tiraboschi Ferro

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.

S586g

Silva, Danielly Beraldo dos Santos

Genes diferencialmente expressos e splicing alternativos relacionados com características de carcaça e carne de bovinos da raça Nelore / Danielly

Beraldo dos Santos Silva. -- Jaboticabal, 2019

117 p. : il., tabs. + 1 CD-ROM

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

Orientadora: Lucia Galvão de Albuquerque

Coorientador: Daniel Guariz Pinheiro

1. Bovino. 2. Área de olho de Lombo (AOL). 3. Conteúdo de gordura intramuscular (GI). 4. Genes. 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.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: GENES DIFERENCIALMENTE EXPRESSOS E SPLICING ALTERNATIVOS RELACIONADOS COM CARACTERÍSTICAS DE CARÇAÇA E CARNE DE BOVINOS DA RAÇA NELORE

AUTORA: DANIELLY BERALDO DOS SANTOS SILVA
ORIENTADORA: LUCIA GALVÃO DE ALBUQUERQUE
COORIENTADOR: DANIEL GUARIZ PINHEIRO
COORIENTADORA: MARIA INES TIRABOSCHI FERRO

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



Profa. Dra. LUCIA GALVÃO DE ALBUQUERQUE
Departamento de Zootecnia / FCAV / Unesp - Jaboticabal



Prof. Dr. JESUS APARECIDO FERRO
Departamento de Tecnologia / FCAV - UNESP - Jaboticabal



Pós-doutoranda LARISSA FERNANDA SIMIELLI FONSECA
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal



Profa. Dra. ALEXÉIA BARUFATTI GRISÓLIA
UFGD/Dourados



Pesquisadora POLIANA FERNANDA GIACHETTO
Empresa Brasileira de Pesquisa Agropecuária / EMBRAPA - Campinas/SP

Jaboticabal, 25 de janeiro de 2019

DADOS CURRICULARES DA AUTORA

Danielly Beraldo dos Santos Silva nasceu em 13 de Fevereiro de 1990, na cidade de Dourados (MS), filha de Edson Beraldo da Silva e Iraci Pereira dos Santos Silva. Em Março de 2009 iniciou o curso de graduação em Biotecnologia na Universidade Federal da Grande Dourados, UFGD, Dourados-MS. No ano de 2010 foi voluntária na modalidade PIVIC-UFGD (Programa Institucional Voluntário de Iniciação Científica), sob a orientação do Prof. Dr. Leonardo de Oliveira Seno. De 2011 à 2013, foi bolsista da UFGD, na modalidade PIBIC (Programa Institucional de Bolsas de Iniciação Científica), sob orientação da Profa. Dra. Alexéia Barufatti Grisolia. Em Abril de 2013, obteve o título de Bacharel em Biotecnologia. Em Maio de 2013, ingressou no Programa de Pós-graduação em Biologia Geral/Bioprospecção, na UFGD, tendo como linha de pesquisa Biotecnologia e Bioensaios. Neste período foi bolsista CAPES, e obteve o título de Mestra em Fevereiro de 2015, sob orientação da Profa. Dra. Alexéia Barufatti Grisolia. Em março de 2015 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), foi bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), tendo como linha de pesquisa genética molecular e seleção genômica. De Fevereiro de 2017 a Fevereiro de 2018, realizou estagio na *University of Missouri* – MO, EUA, sob supervisão do Prof. Dr. Jeremy F. Taylor. Em Janeiro de 2019, obteve o título de doutora em Genética e Melhoramento Animal, sob orientação da Profa. Dra Lucia Galvão de Albuquerque.

"O Senhor não olha tanto a grandeza das nossas obras.

Olha mais o amor com que são feitas"

(Santa Teresa de Ávila)

Aos meus pais (Edson e Iraci), aos quais eu sou imensamente grata por todo apoio, por me ensinar a conduzir a vida com dignidade, sabedoria e humildade, por todo amor e carinho. À minha irmã Nadielly, por ser minha inspiração e o meu maior amor.

“Tua família te ama e te espera”

“Por toda minha vida”

Dedico e Ofereço.

AGRADECIMENTOS

À Deus. *“Porque aos teus anjos ele mandou para que te guarde pelos teus caminhos”*
(Salmos 90, 11).

Agradeço a toda minha família e amigos. Obrigada, pelo incentivo e carinho.

Agradeço imensamente a Prof^a. Dr^a. Lucia Galvão de Albuquerque pela orientação, incentivo e também pela confiança. Obrigada por me receber, ser um exemplo de profissionalismo e dedicação.

Agradeço ao Prof. Dr. Daniel Guariz Pinheiro e a Prof^a. Dr^a. Maria Inês Tiraboschi Ferro pela co-orientação, pelas contribuições, ensino e atenção. Agradeço ao Dr. Jerry Taylor pela oportunidade em me receber em seu grupo de pesquisa nos EUA, pelas contribuições e ensino.

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

A todos os amigos dos Departamentos de Zootecnia e Melhoramento Genético Animal, especialmente aos colegas de sala e aos meus queridos amigos: Ana Cristina, Ana Fabrícia, André, Ândrea, Bruna, Gabriela, Larissa, Lúcio Flávio, Malane, Rafael, Thaise e Willian. Obrigada não só pelo auxílio nas pesquisas, como também pela amizade, pelos momentos de descontração e alegria.

A todos os professores, especialmente àqueles do Programa de Pós-Graduação em Genética e Melhoramento Animal. Também aos técnicos administrativos e demais funcionários da Universidade.

A 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, #2015/16850-9 e #2016/22894-1).

SUMÁRIO

CAPÍTULO 1 - Considerações gerais.....	1
1. Introdução.....	2
2. Objetivos.....	2
3. Revisão de Literatura.....	3
4. Referências.....	9
CHAPTER 2 - Transcriptome profiling of muscle in Nelore cattle phenotypically divergent for the ribeye muscle area.....	15
Abstract.....	15
Background.....	15
Methods.....	17
Results.....	19
Discussion.....	23
Conclusions.....	28
List of abbreviations.....	28
Declarations.....	29
References.....	29
Additional Files.....	33
CHAPTER 3 - Prediction of hub genes associated with intramuscular fat content in Nelore cattle.....	45
Abstract.....	45
Background.....	45
Methods.....	47
Results.....	49
Discussion.....	56
Conclusions.....	62
List of abbreviations.....	62
Declarations.....	62
References.....	63
Additional Files.....	68
CHAPTER 4 - Differentially expressed alternatively spliced genes associated with carcass and meat traits in Nelore cattle.....	72
Abstract.....	72
Introduction.....	72
Results.....	73
Discussion.....	80

Conclusion.....	86
Methods.....	86
References.....	90
Acknowledgements.....	93
Additional Files.....	94
CAPÍTULO 5 - Considerações finais.....	115

GENES DIFERENCIALMENTE EXPRESSOS E SPLICING ALTERNATIVOS RELACIONADOS COM CARACTERÍSTICAS DE CARÇAÇA E CARNE DE BOVINOS DA RAÇA NELORE

RESUMO – O sequenciamento do RNA (RNA-Seq) é uma das abordagens utilizadas para identificar transcritos correspondentes a genes candidatos que provocam variações em vias metabólicas, resultando em diferentes fenótipos. Uma vez que, área de olho de lombo (AOL) e o conteúdo de gordura intramuscular (GI) são características poligênicas, muitos genes e mecanismos que induzem as diferenças no crescimento e desenvolvimento muscular, bem como nos índices de conteúdo de GI da raça Nelore, ainda são desconhecidos. Portanto, o objetivo foi estudar o transcriptoma do músculo de bovinos da raça Nelore, com o propósito de identificar genes e eventos de *splicing* alternativos diferencialmente expressos (DEGs e DAS) associados à característica de carcaça (AOL) e carne (conteúdo de GI), por meio do RNA-Seq. Para tanto, um total de 80 animais foram sequenciados e fenotipados para AOL e conteúdo de GI (mensurados pelo método químico - que avalia lipídios totais). Os resultados foram apresentados nos capítulos 2, 3 e 4. No capítulo 2, foram selecionados para análise de expressão diferencial, 15 animais com maiores e 15 animais com menores AOL. Após as análises, 288 genes foram diferencialmente expressos ($q\text{-value} \leq 0,05$), dos quais 182 foram superexpressos e 106 foram reprimidos no grupo de animais com maiores AOL em comparação com o grupo com menores AOL. Genes pertencentes a famílias da actina, miosina, colágeno, integrina, transportador de soluto, ubiquitina e *kelch-like* foram diferencialmente expressos. A análise de enriquecimento funcional mostrou que muitos dos *Gene Ontology terms* (GO) estavam associados com o desenvolvimento, crescimento e degradação muscular. Por meio da análise de rede de co-expressão, foram previstos três genes *hub* (*PPP3R1*, *FAM129B* e *UBE2G1*), os quais estavam envolvidos no crescimento e proteólise muscular. No capítulo 3, foram selecionados para análise de expressão diferencial, 15 animais com maiores e 15 animais com menores conteúdos de GI. Após as análises, 65 DEGs foram identificados, incluindo 21 genes superexpressos e 44 reprimidos em animais com maiores conteúdos de GI em comparação com animais com menores conteúdos de GI. A análise de enriquecimento funcional mostrou que os GO *terms* foram associados ao metabolismo lipídico e muscular, assim como ao sistema imune. A análise da rede de co-expressão mostrou que os três genes *hub* preditos (*PDE4D*, *KLHL30* e *IL1RAP*) podem desempenhar um papel na biologia lipídica celular e/ou sistêmica em bovinos da raça Nelore. No capítulo 4, foram identificados eventos de *splicing* alternativos diferencialmente expressos (DAS), e seus genes associados, nos mesmos grupos de animais selecionados no capítulo 2 (fenotipicamente divergentes para AOL) e no capítulo 3 (animais divergentes para conteúdo de GI). Os eventos DAS foram identificados a partir da abordagem *exon-centric*. Os resultados mostraram 166 e 269 eventos DAS ($p\text{-value} \leq 0,05$), respectivamente, para AOL e conteúdo de GI. O salto de éxon e sítios alternativos 3' foram os eventos mais frequentemente encontrados para ambas características. Os DAS encontrados foram transcritos para 125 e 219 genes, respectivamente para AOL e para GI. Os genes encontrados no capítulo 4, no geral, desempenham um papel no metabolismo muscular e lipídico. Para ambas as características, alguns DAS pertenciam às famílias de genes da miosina e miotilina. Os resultados dessa tese indicam possíveis genes candidatos e fornecem a base teórica para uma melhor compreensão dos mecanismos moleculares do desenvolvimento e crescimento da AOL, bem como da deposição de GI em bovinos da raça Nelore.

Palavras-Chave: AOL, bovino, genes, gordura intramuscular, *splicing* alternativo

DIFFERENTIALLY EXPRESSED GENES AND ALTERNATIVE SPLICING RELATED TO CARCASS AND MEAT TRAITS OF THE NELORE CATTLE

ABSTRACT - RNA sequencing (RNA-Seq) is an approach for screening the expression of functional candidate genes and identifying important molecular mechanisms creating variation in pathways that result in different tissue phenotypes. Since ribeye muscle area (REA) and intramuscular fat content (IF) are polygenic traits, many genes and mechanisms that induce differences in muscle growth and development, as well as IF content of the Nelore cattle still unknown. The aim was to study the muscle transcriptome of the Nelore cattle with the purpose of identifying differentially expressed genes (DEGs) and differentially expressed alternative splicing events (DAS) associated with the carcass (REA) and meat (IF) traits, through of the RNA-seq approach. Therefore, a total of 80 animals were sequenced and phenotyped for REA and IF content (measured by the chemical method - which evaluate total lipids). The results were presented in chapters 2, 3 and 4. In chapter 2, 15 animals with highest REA and 15 animals with lowest REA were selected for differential expression analysis. Upon analysis, 289 DEGs were identified (q -value ≤ 0.05): 183 upregulated and 106 downregulated in the highest REA group. Genes belonging to important families, such as actin, myosin, collagen, integrin, solute carrier, ubiquitin, and kelch-like, were among the DEGs. Gene set enrichment analysis showed that many of the significantly enriched gene ontology (GO) terms are closely associated with muscle development, growth, and degradation. Through co-expression network analysis, we predicted three hub genes (PPP3R1, FAM129B and UBE2G1). These genes are involved in muscle growth and proteolysis. In chapter 3, 15 animals with highest and 15 animals with lowest IF content were selected for differential expression analysis. Upon analysis, 65 DEGs were identified, including 21 upregulated and 44 downregulated genes in highest IF content animals. Gene set enrichment analysis was performed and showed that the GO terms were closely associated with lipid and muscle metabolism, as well as genes related to the immune system. Gene co-expression network analysis showed three hub genes (*PDE4D*, *KLHL30*, and *IL1RAP*) that may play a role in cellular and/or systemic lipid biology in Nelore cattle. In Chapter 4, differentially expressed alternative splicing events (DAS), and their associated genes, were identified in the same groups of animals selected in Chapter 2 (phenotypically divergent for REA) and Chapter 3 (phenotypically divergent for IF content). The DAS events were identified from the exon-centric approach. The results showed 166 and 269 DAS events (p -value $\leq 0,05$), respectively, for REA and IF. The exon cassette and alternative 3' splice site were the most frequently found events for both traits. The DAS found were transcribed for 125 and 219 genes, respectively for REA and for IF. The genes found in chapter 4, in general, play a role in muscle metabolism and systemic lipid biology. The results reported in this tese show possible candidate genes and provide a theoretical basis for a better understanding of the molecular mechanism of REA development and growth as well as the deposition of IF in Nelore cattle.

Keywords: Alternative splicing, bovine, genes, intramuscular fat, REA

CERTIFICADO DA COMISSÃO DE ÉTICA



CERTIFICADO

Certificamos que o Projeto intitulado **“Ferramentas genômicas no melhoramento genético de características de importância econômica direta em bovinos da raça Nelore”**, protocolo nº 18.340/16, sob a responsabilidade do Prof.^a Dr.^a Lucia Galvão de Albuquerque, que envolve a produção, manutenção e/ou utilização de animais pertencentes ao Filo Chordata, subfilo Vertebrata (exceto o homem), para fins de pesquisa científica (ou ensino) - encontra-se de acordo com os preceitos da lei nº 11.794, de 08 de outubro de 2008, no decreto 6.899, de 15 de junho de 2009, e com as normas editadas pelo Conselho Nacional de Controle da Experimentação Animal (CONCEA), e foi aprovado pela COMISSÃO DE ÉTICA NO USO DE ANIMAIS (CEUA), da FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS, UNESP - CÂMPUS DE JABOTICABAL-SP, em reunião ordinária de 14 de dezembro de 2016.

Vigência do Projeto	01/01/2017 a 01/01/2020
Espécie / Linhagem	<i>Bos indicus</i> (Bovinos)
Nº de animais	10.037
Peso / Idade	480 kg / 730 dias
Sexo	Macho e fêmea
Origem	Programas de Melhoramento Nelore Qualitas, Delta Gen, Cia do Melhoramento, Paint e Instituto de Zootecnia

Jaboticabal, 14 de dezembro de 2016.


Prof. Dr. Lizandra Amoroso
 Coordenadora – CEUA

CAPÍTULO 1 - Considerações gerais

1. Introdução

A identificação de genes e o conhecimento dos mecanismos moleculares envolvidos no crescimento muscular e metabolismo lipídico pode auxiliar na seleção de animais da raça Nelore com valor genético elevado para características como a área de olho de lombo (AOL) e conteúdo de gordura intramuscular (GI), contribuindo assim, para a melhoria da carne brasileira. A AOL é uma característica de carcaça e está relacionada com a musculosidade e ganho de peso do animal (Caetano et al., 2013; Scholz et al., 2015). O conteúdo de GI é um dos principais fatores que afetam positivamente a qualidade da carne, devido a sua influência na palatabilidade, maciez e valor nutricional (Wood et al., 2008; Costa et al., 2013; Pickworth et al., 2010).

A maioria das características fenotípicas é controlada por múltiplos genes e/ou isoformas, portanto a análise global de transcriptoma tem sido muito utilizada na identificação de genes diferencialmente expressos associados a diversas características de interesse (Ponsuksili et al., 2008; Hamill et al., 2012). A sequência completa dos RNAs de um organismo (transcriptoma) pode ser obtida por meio do sequenciamento de nova geração, técnica conhecida como RNA-Seq. (Ramayo-Caldas et al., 2012). Por meio do RNA-Seq, é possível quantificar todos os genes expressos e suas isoformas (*splicing* alternativo) de maneira mais acurada, diferentemente de outros métodos tradicionais de análise de expressão gênica, como microarranjos e análise da expressão gênica em série (SAGE, *Serial analysis of gene expression*) (Tang et al., 2011).

O RNA-Seq permitiu descobertas importantes em relação os mecanismos moleculares de características de interesse econômico de animais da raça Nelore. Estas descobertas incluíram a identificação de genes diferencialmente expressos em tecido muscular relacionados com características de carcaça: AOL e gordura subcutânea (Silva-Vignato et al., 2017), bem como características da carne: maciez (Fonseca et al., 2017); composição de ácidos graxos (Berton et al., 2016); gordura intramuscular (mensurado por scores visuais de marmoreio) (Cesar et al., 2015) e conteúdo de ferro (Diniz et al., 2016). Por meio do RNA-Seq também foi possível a

identificação de eventos de *splicing* alternativos associados à mastite (Wang et al., 2016), desenvolvimento embrionário (Huang e Khatib, 2010) e muscular (He e Liu, 2013) em bovinos de origem taurina. Em animais de origem indiana, especialmente os da raça Nelore, o conhecimento sobre os mecanismos moleculares relacionados a características de carcaça e carne, regulados pelos eventos *splicing* alternativo, ainda é limitado. Deste modo, RNA-Seq é uma importante ferramenta para desvendar os mecanismos subjacentes de características complexas, promovendo uma melhor compreensão da regulação genética do fenótipo de interesse (Tizioto et al., 2015).

Uma vez que, AOL e o conteúdo de GI são características poligênicas e são influenciados por diversos fatores ambientais, muitos genes e mecanismos que induzem as diferenças no crescimento e desenvolvimento muscular, bem como nos índices de conteúdo GI da raça Nelore, ainda são desconhecidos. A compreensão destes mecanismos moleculares podem fornecer *insights* para ajudar no desenvolvimento de estratégias de melhoramento e, como consequência, melhorar a qualidade da carne (Ouali et al., 2013; PicarD et al., 2015; D'alessandro e Zolla, 2013).

2. Objetivos

2.1. Geral

- Estudar o transcriptoma do músculo *Longissimus thoracis* de bovinos da raça Nelore, com o propósito de identificar genes diferencialmente expressos e eventos de *splicing* alternativo associados à característica de carcaça (área de olho de lombo) e carne (conteúdo de gordura intramuscular), fornecendo assim, informações sobre os mecanismos genéticos e moleculares que regulam essas características.

2.2. Específicos

- Identificar genes diferencialmente expressos em tecido muscular de animais com fenótipos divergentes para área de olho de lombo, utilizando dados de expressão gênica obtidos pelo sequenciamento de RNA.
- Identificar genes diferencialmente expressos em tecido muscular de animais com fenótipos divergentes para conteúdo de gordura intramuscular

(mensurados pelo método químico - que avaliam lipídios totais), utilizando dados de expressão gênica obtidos pelo sequenciamento de RNA.

- Identificar os genes que codificam transcritos processados (*spliced*) alternativamente expressos em tecido muscular de animais com fenótipos divergentes para área de olho de lombo e conteúdo de gordura intramuscular (lipídios totais), utilizando dados de expressão gênica obtidos pelo sequenciamento de RNA.

3. Revisão de Literatura

O RNA (Ácido ribonucleico, do inglês *Ribonucleic acid*), considerado a molécula da vida, desempenha funções no armazenamento e na transmissão da informação genética, assim como funções catalíticas. O DNA (Ácido desoxirribonucleico, do inglês *Desoxybonucleic acid*) é o repositório da informação genética que, pelo processo de transcrição, pode ser transmitida para uma molécula intermediária de RNA (transcrito), que, por sua vez, pode ser precursora na síntese proteica (tradução) (Alberts et al., 2017; Nelson e Cox, 2017; Maia, 2017).

Brenner et al. (1961) mostrou, experimentalmente, a importância desse intermediário de RNA como molde para síntese proteica e, devido a sua instabilidade e curta meia-vida, o denominaram de RNA mensageiro (mRNA). Além do mRNA, existem outras classes de RNAs que não são traduzidas para aminoácidos, como por exemplo: ncRNAs (RNAs não codificadores), lncRNAs (RNAs não codificadores longos), small-ncRNAs (RNAs não codificadores pequenos) e circRNAs (RNAs circulares). Todas essas classes podem desempenhar inúmeras funções celulares, dentre elas, funções estruturais, catalíticas e reguladoras (Pereira 2017).

Nos eucariotos, de maneira resumida, no processo de transcrição o RNA é catalizado pela polimerase II e o produto da transcrição é um transcrito primário (ou pré-RNA) que contém éxons (possíveis codificadores), íntrons (não codificadores), 7-metilguanossina (cap) na extremidade 5' e calda de adenosinas (poli A) na extremidade 3'. A partir do pré-RNA, os íntrons são removidos por um complexo ribonucleoproteico chamado spliceossomo, o qual reconhece as junções éxons-íntrons de maneira específica e promove as reações de

processamento (*splicing*, do inglês) (Alberts et al., 2017; Nelson e Cox, 2017; Maia, 2017).

Padrões alternativos de processamento podem gerar diferentes combinações de éxons, e, portanto, mRNA com diferentes capacidades codificadoras, aumentando assim, a plasticidade do transcriptoma celular. Esse mecanismo pode ser chamado de processamento alternativo do mRNA (*alternative splicing*, do inglês) (Alberts et al., 2017; Nelson e Cox, 2017). O processamento alternativo do mRNA pode ser diferencialmente regulado de acordo com o tipo celular, estágio do desenvolvimento ou diferentes padrões de sinalização (Alberts et al., 1994). Esse processo é um dos principais mecanismos que proporcionam a diversidade proteica (Alberts et al., 2017).

As diferentes clivagens dos limites éxons-íntrons realizadas pelos spliceossomos, na maioria das vezes, se enquadram em subgrupos de eventos de processamento alternativos do mRNA (Nakka et al., 2018): saltos de éxons (*cassette exon*); éxons alternativos mutualmente excludentes (*mutually exclusive exon*); saltos de éxons coordenados (*coordinate cassette exon*); sítios alternativos de processamento 5' e 3' (*alternative 5' and 3' splice sites*); retenção de íntrons (*intron retention*); primeiro e último éxons alternativos (*alternative first e last exon*) (Blencowe, 2006; Nilsen e Graveley, 2010) (Figura 1). Os eventos de *splicing* alternativos mais comuns em mamíferos são os saltos de éxons e os sítios alternativos de processamento 5' e 3', e em contraste, retenção de íntron é o evento de *splicing* alternativo mais raro encontrado em mamíferos (Nakka et al., 2018; Sammeth et al., 2008).

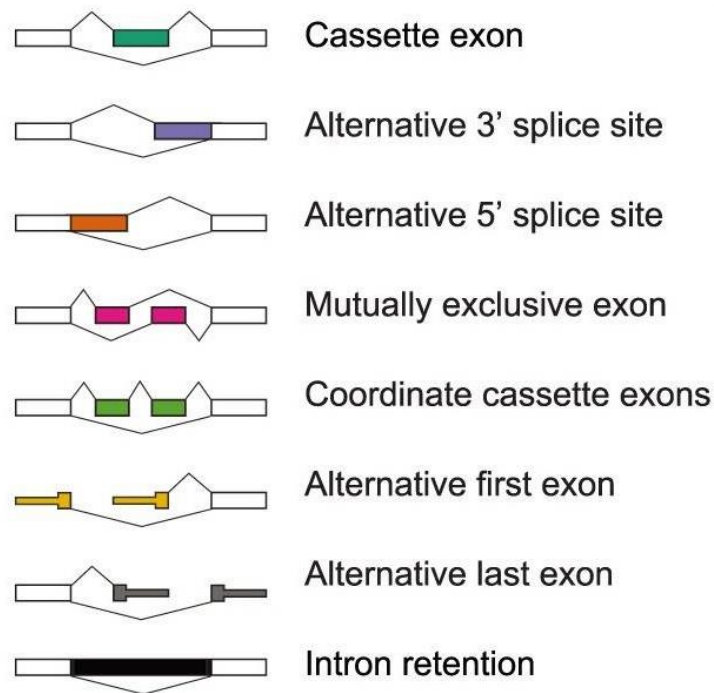


Figura 1. Representação da classificação dos tipos de processamento alternativos do mRNA (Figura adaptada de Brooks et al., 2014). Os exons constitutivos são mostrados em branco enquanto que os éxons alternativos mostrados em coloridos. *Cassette exon* (salto de éxon): um éxon pode ser retido no transcrito. *Alternative 3' splice sites* (sítios alternativos de processamento 3'): um sítio alternativo 3' é usado, alterando o limite 5' do éxon a jusante. *Alternative 5' splice sites* (sítios alternativos de processamento 5'): um sítio alternativo 5' é usado, alterando o limite 3' do éxon a montante. *Mutually exclusive exon* (éxons alternativos mutuamente excludentes): um dos dois éxons é retido no transcrito, mas não ambos. *Coordinate cassette exon* (saltos de éxons coordenados): ambos éxons são retidos no transcrito. *Alternative first exon* (primeiro éxon alternativo): diferentes locais de início da transcrição. *Alternative last exon* (último éxon alternativo): diferentes locais de parada de transcrição. *Intron retention* (retenção de intron): o intron é transcrito.

A maioria das características fenotípicas é controlada por múltiplos genes e/ou isoformas, portanto a análise global de transcriptoma tem sido muito utilizada na identificação de genes e/ou isoformas diferencialmente expressos para diversas características (Ponsuksili et al., 2009; Hamill et al., 2012). Existem várias metodologias utilizadas para análise do transcriptoma, como por exemplo, microarranjos, análise da expressão gênica em série (SAGE, *Serial analysis of gene expression*) (Tang et al., 2011) e RNA-Seq (*High-throughput RNA sequencing*). O RNA-Seq é a técnica que atualmente vem sendo muito utilizada para investigação de transcriptomas. Essa metodologia é um das abordagens utilizadas para a triagem da expressão de genes e seus transcritos processados alternativamente, e identificação de importantes mecanismos

moleculares que geram variação nas vias resultando em diferentes fenótipos teciduais (Trapnell, et al., 2010; He e Liu, 2013).

Com a crescente popularidade do RNA-Seq, muitas abordagens e softwares foram implementados, e tem sido amplamente utilizados para análise de expressão diferencial, conforme descrito por Griffith et al. (2015) e Conessa et al. (2016). Segundo esses mesmos autores, basicamente o roteiro geral para identificação de genes ou *splicing* alternativos diferencialmente expressos, por meio de RNA-Seq, em duas ou mais condições experimentais envolve: 1) obtenção de dados brutos (sequenciamento do RNA mensageiro); 2) realização do controle de qualidade das leituras; 3) alinhamento das leituras com o genoma ou transcriptoma de referência; 4) quantificação e normalização do nível de expressão; 5) teste estatístico para a expressão diferencial de genes e/ou *splicing* alternativo; 6) análise de enriquecimento funcional.

Existem várias tecnologias de sequenciamento de segunda geração disponíveis e, dentre elas, uma das mais utilizadas é a Plataforma Solexa Illumina. Em resumo, o mRNA é convertido em bibliotecas de DNA complementar (cDNA). O cDNA é fragmentado, posteriormente, os adaptadores são ligados às extremidades 5' dos fragmentos de cDNA, clonados *in vitro* (flow cell), e então são sequenciadas. No caso do sequenciamento na Plataforma Solexa Illumina, o sequenciamento é realizado por síntese, para tanto é utilizado a DNA polimerase e nucleotídeos terminadores marcados com diferentes fluoróforos. Como resultado, são geradas milhões de leituras (*reads*) brutas. (Carvalho e Silva, 2010). Para o controle de qualidade das leituras, tipicamente são avaliados o conteúdo de GC, a presença de adaptadores, *K-mer* super-representados, bases ambíguas, sequências ribossomais, leituras duplicadas, fragmentos mais curtos do que o comprimento das leituras, entre outros fatores (Griffith et al., 2015; Conessa et al.; 2016).

Os programas de alinhamento de leituras de transcrição, na sua maioria, foram desenvolvidos para identificar *spliced reads* geradas pelos eventos de *splicing* e para determinar corretamente os limites éxon-íntron (Engström et al., 2013). Os alinhadores do tipo *spliced*, normalmente realizam o mapeamento em duas etapas. Inicialmente, as leituras são mapeadas no genoma ou transcriptoma e usadas para construir todas as possíveis junções de éxon-éxon. Essa informação é utilizada em uma segunda etapa de mapeamento com as

leituras que não foram mapeadas na primeira etapa (Trapnell et al., 2010; Engström et al., 2013; Kim et al., 2016).

A aplicação mais comum do RNA-seq é a estimação da expressão gênica. Esta aplicação é baseada principalmente no número de leituras que mapeiam cada sequência de transcrição. Os métodos de normalização controlam as diferenças das amostras, e assim facilita a precisão das comparações entre as diferentes condições estudadas (Conessa et al.; 2016). Vários métodos foram desenvolvidos para a normalização dos dados, como por exemplo, RPKM ou sua variação, FPKM (*reads or fragments per kilobase of exon per million reads*). A normalização por RPKM ou FPKM consiste em dividir o número de leituras ou fragmentos mapeados no transcrito pelo número total de leituras ou fragmentos mapeados em cada amostra, vezes o tamanho do transcrito (Mortazavi et al., 2008). Desse modo, uma amostra com uma cobertura de sequenciamento maior não atrapalha a comparação entre amostras.

Uma vez que as leituras foram normalizadas, então é possível testar hipóteses a respeito da existência de expressão diferencial de transcritos entre as diferentes condições avaliadas. Para tanto, os métodos propostos para tal avaliação se baseiam em distribuições discretas de probabilidade, tais como binomial, poisson e binomial negativa. O programa Cuffdiff pertencente ao conjunto de ferramentas do Cufflinks (Trapnell et al., 2010). O Cuffdiff assume que os dados apresentam distribuição normal, ou seja, o número de leituras produzidas por cada transcrito é proporcional à sua abundância. O Cuffdiff calcula o *p-value* (ajustado, *q-value*) para os genes diferencialmente expressos, baseado na contagem dos valores de FPKM entre dois ou mais tratamentos por meio do teste T (Rapaport et al., 2013).

Para análise de expressão diferencial de *splicing* alternativo entre as diferentes condições experimentais também é necessário testar hipóteses. Para tanto, podem ser utilizadas as abordagens baseada na isoforma (*isoform-centric*) ou baseada no *exon* (*exon-centric*) (Chen, 2013). A abordagem baseada na isoforma, inicialmente quantifica cada isoforma e a mesma é comparada entre si para inferir o nível de *splicing*. Os métodos baseados em *exon* considera a razão da inclusão de cada *exon* individual (Chen, 2013). Esse método é apropriado se o objetivo do estudo não estiver nas isoformas inteiras, mas na inclusão e

exclusão de *exons* específicos e os domínios das proteínas funcionais (Ver Figura 1)

A análise de enriquecimento funcional de genes requer a disponibilidade de dados suficientes de anotação funcional para o transcriptoma em estudo. Recursos como, por exemplo, Gene Ontology (GO) (Ashburner et al., 2000) contêm dados de anotação para a maioria das espécies modelo (Conessa et al., 2016). Outra forma de buscar associações entre os genes é por meio de estudos de redes (*networks*) gênicas. As redes gênicas mostram a interação entre os genes e outras moléculas celulares, os quais regulam os níveis de expressão gênica de RNA mensageiro e proteínas. Abordagens de rede têm sido usadas para identificar a regulação transcricional complexa, isto é, identificação de genes centrais (*hub* ou *keyplayers*). Genes *hub* são altamente correlacionados com um grande número de genes e podem desempenhar importantes funções como os principais reguladores em redes de co-expressão (Filteau et al., 2013; Lim et al., 2013; Kogelman et al., 2014), ou seja, são propostos para desempenhar um papel importante na biologia global dos organismos (Langfelder et al., 2013). Assim, a análise de redes de co-expressão pode facilitar a detecção de importantes vias biológicas envolvidas em fenótipos direcionados.

De acordo com Costa-Silva et al. (2017), para análises de expressão diferencial ainda não existe um consenso sobre qual metodologia é mais apropriada ou qual abordagem garante a validade dos resultados em termos de robustez, acurácia e reprodutibilidade. Cada metodologia é inerente à amostra (conjunto de dados) estudada. Como muitas diferenças podem ocorrer devido a variantes de cada abordagem, esta discussão no âmbito da bioinformática ainda está em desenvolvimento (Garber et al., 2011; Rapaport et al., 2013; Conessa et al., 2016; Zhang et al., 2018; Hrdlickova et al., 2017).

RNA-Seq é uma importante ferramenta para desvendar os mecanismos subjacentes de características complexas, promovendo uma nova compreensão da regulação genética do fenótipo de interesse (Tizioto et al., 2015). Um dos primeiros estudos com a aplicação de RNA-Seq em bovinos foi realizada por Medrano et al. (2010). Esses autores realizaram uma análise comparativa entre o transcriptoma de células somáticas e tecido mamário em bovinos da raça Holandesa, demonstrando a eficiência dessa tecnologia para análise de

expressão gênica diferencial em animais de produção. Desde então, o RNA-Seq tem sido amplamente utilizado para estudos de transcriptomas de animais de produção, como bovinos (Fortes et al., 2016; Casey et al., 2011); abelhas (Badaoui et al., 2017); frangos (Cui et al., 2017); cavalos (Correia et al., 2018); suínos (Cardoso et al., 2017) e peixes (Sudhagar et al., 2018).

O RNA-Seq foi utilizado com o objetivo de identificar genes diferencialmente expressos em tecido muscular de bovinos da Raça Nelore, associados à: maciez da carne (Fonseca et al., 2017) e composição de ácidos graxos (Berton et al., 2016) de uma população de animais comerciais não castrados; área de olho de lombo (AOL) e gordura subcutânea (Silva-Vignato et al., 2017); gordura intramuscular (mensurado por meio dos scores visuais de marmoreio) (Cesar et al., 2015); consumo alimentar residual (Tizioto et al., 2016); e conteúdo de ferro da carne (Diniz et al., 2016) de novilhos castrados pertencentes à um rebanho experimental. Todos esses autores mostraram que as características estudadas são influenciadas por diferentes genes e uma complexa rede de interações genéticas no músculo de animais da raça Nelore.

Muitos estudos têm identificado e classificado eventos de *splicing* alternativos em diferentes organismos como, por exemplo: bactérias (Liang et al., 2016), plantas (Potenza et al., 2015), insetos (Brooks et al., 2011) e humanos (Guan et al., 2017). Em bovinos de origem taurina, eventos de *splicing* alternativos têm sido associados à mastite (Wang et al., 2016), desenvolvimento embrionário (Huang e Khatib, 2010) e muscular (He e Liu, 2013). Em animais de origem indiana, especialmente os da raça Nelore, o conhecimento sobre os mecanismos moleculares regulados pelos eventos *splicing* alternativo, principalmente no contexto de diferenças individuais nas características de carcaça e carne, ainda é limitado.

4. Referências

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (Ed.) (1994). **Biologia molecular da célula**. Porto Alegre: Artes Médicas, 1294 p.

Ashburner M, Ball CA, et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. **Nat Genet** 25(1):25-9.

Badaoui B, Fougeroux A, et al. (2017). RNA-sequence analysis of gene expression from honeybees (*Apis mellifera*) infected with *Nosema ceranae*. **PLoS One**. 12(3):e0173438.

Berton MP, Fonseca LFS, et al. (2016) Gene expression profile of intramuscular muscle in Nellore cattle with extreme values of fatty acid. **BMC Genomics** 17:972.

Blencowe BJ. (2006). Alternative splicing: new insights from global analyses. **Cell** 126:37-47.

Brenner S, Jacob F, Meselson M. (1961). An unstable intermediate carrying information from genes to ribosomes for protein synthesis. **Nature** 100, 676.

Brooks AN, Yang L, Duff MO, Hansen KD, Park JW, Dudoit S, Brenner SE, Graveley BR (2011) Conservation of an RNA regulatory map between drosophila and mammals. **Genome Research** 21:193-202.

Brooks AN, Choi PS, et al. (2014). A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. **PLoS One** 9(1):e87361.

Caetano SL, Savegnago RP, Boligon AA, Ramos SB, Chud TCS, Lôbo RB, Munari DP (2013) Estimates of genetic parameters for carcass growth and reproductive traits in Nellore cattle. **Livestock Science** 155(1):1–7.

Cardoso TF, Cánovas A, Canela-Xandri O, González-Prendes R, Amills M, Quintanilla R.(2017). RNA-seq based detection of differentially expressed genes in the skeletal muscle of Duroc pigs with distinct lipid profiles. **Sci Rep** 7:40005.

Carvalho MCCG, Silva DCG. (2010). Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. *Ciência Rural* 40(3).

Casey T, Dover H, et al. (2011). Transcriptome analysis of epithelial and stromal contributions to mammaryogenesis in three week prepartum cows. **PLoS One** 6(7):e22541.

Cesar AS, Regitano LC, et al. (2015). Putative regulatory factors associated with intramuscular fat content. **PLoS One** 10 (6):1-21.

Chen L. (2013). Statistical and Computational Methods for High-Throughput Sequencing Data Analysis of Alternative Splicing. **Stat Biosci** 5(1):138-155.

Conesa A, Madrigal P, et al. (2016). A survey of best practices for RNA-seq data analysis. **Genome Biology** 17:13.

Correia CN, McLoughlin KE, et al. (2018). RNA Sequencing (RNA-Seq) Reveals Extremely Low Levels of Reticulocyte-Derived Globin Gene Transcripts in Peripheral Blood From Horses (*Equus caballus*) and Cattle (*Bos taurus*). **Front Genet**. 2018;9:278.

Costa ASH, Pires VMR, Fontes CMGA, Prates JAM (2013). Expression of genes controlling fat deposition in two genetically diverse beef cattle breeds fed high or low silage diets. **BMC Veterinary Research** 9(118):1-16.

Costa-Silva J, Domingues D, Lopes FM (2017). RNA-Seq differential expression analysis: An extended review and a software tool. **PLoS One** 12(12):e0190152

Cui X, Marshall B, Shi N, Chen SY, Rekaya R, Liu HX. (2017). RNA-Seq analysis on chicken taste sensory organs: An ideal system to study organogenesis. **Sci Rep** 7(1):9131.

D'Alessandro A, Zolla L (2013) Meat science: From proteomics to integrated omics towards system biology. **J Proteomics** 78:558–77.

Diniz WJ, Coutinho LL, Tizioto PC, Cesar AS, Gromboni CF, Nogueira AR, Oliveira PSN, Souza MM, Regitano LCA (2016) Iron content affects Lipogenic Gene expression in the muscle of Nelore beef cattle. **PLoS One** 11(8):e0161160.

Engström PG, Steijger T, et al. (2013). Systematic evaluation of spliced alignment programs for RNA-seq data. **Nature methods**, 10(12):1185-91.

Filteau M, Pavey SA, St-Cyr J, Bernatchez L. (2013). Gene Coexpression networks reveal key drivers of phenotypic divergence in Lake whitefish. **Mol Biol Evol.** 30(6):1384–96.

Fonseca LFS, Gimenez DFJ, Silva DBS, Barthelson R, Baldi F, Ferro JA, Albuquerque LG (2017) Differences in global gene expression in muscle tissue of Nelore cattle with divergent meat tenderness. **BMC Genomics** 18(945):1-12.

Fortes MR, Nguyen LT, et al. (2016). Transcriptome analyses identify five transcription factors differentially expressed in the hypothalamus of post-versus prepubertal Brahman heifers. **J Ani Sci.** 94(9): 3693-3702.

Garber M, Grabherr MG, Guttman M, Trapnell C (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. **Nature methods** 8(6):469–477.

Griffith M, Walker JR, Spies NC, Ainscough BJ, Griffith OL (2015). Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud. **PLOS one Computational Biology** 6:1-20.

Guan Y, Liang G, Martin GB, Guan LL (2017). Functional changes in mRNA expression and alternative pre-mRNA splicing associated with the effects of nutrition on apoptosis and spermatogenesis in the adult testis. **BMC Genomics** 18:64.

Hamill RM, McBryan J, Mcgee C, Mullen AM, Sweeney T, Talbot A, Cairns MT, Davey GC. (2012). Functional analysis of muscle gene expression profiles

associated with tenderness and intramuscular fat content in pork. **Meat Science** 92(4):440-50.

He H, Liu X (2013). Characterization of Transcriptional Complexity during Longissimus Muscle Development in Bovines Using High-Throughput Sequencing. **PLoS One** 8(6):e64356.

Hrdlickova R, Toloue M, Tian B (2017). RNA-Seq methods for transcriptome analysis. **Wiley Interdiscip Rev RNA** 8(1):1-24.

Huang W, Khatib H (2010) Comparison of transcriptomic landscapes of bovine embryos using RNA-Seq. **BMC Genomics** 11:1-10.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. **Genome Biology**. 14:R36.

Kogelman LJA, Cirera S, Zhernakova DV, Fredholm M, Franke L, Kadarmideen HN. (2014). Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA sequencing in a porcine model. **BMC Med Genet** 7(1):57.

Langfelder P, Mischel PS, Horvath S. (2013). When is hub gene selection better than standard meta-analysis? **PLoS One** 8:e61505.

Liang G, Malmuthuge N, Guan Y, Ren Y, Griebel PJ, Guan LL (2016). Altered microRNA expression and pre-mRNA splicing events reveal new mechanisms associated with early stage Mycobacterium avium subspecies paratuberculosis infection. **Scientific Reports** 6:24964.

Lim D, Lee S-H, et al. (2013). Gene co-expression analysis to characterize genes related to marbling trait in Hanwoo (Korean) cattle. **Asian Australas J Anim Sci** 26(1):19–29.

Maia IG (2017). O RNA mensageiro (mRNA). In.: Pereira TC (Ed.) **Introdução ao universo dos non-coding RNAs**. Ribeirão Preto: Sociendade Brasileira de Genética, p.24-25.

Medrano JF, Rincon G, Islas-Trejo A (2010) Comparative analysis of bovine milk and mammary gland transcriptome using RNA-Seq. In: 9th World congress on genetics applied to livestock production, Leipzig, Germany, August 1-6, 2010, Paper no 0852.

Mortazavi A, Williams BA, Mccue K, Schaeffer L, Wold B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nature Methods** 5(7):621-628.

Nakka K, Ghigna C, Gabellini D, Dilworth FJ. (2018). Diversification of the muscle proteome through alternative splicing. **Skelet Muscle** 8(1):8.

Nelson DL, Cox MM. (2017). RNA metabolism. In: Lehninger, Albert L, David L. Nelson, Michael M. Cox. **Lehninger Principles of Biochemistry**. New York: W. H. Freeman, 1021-1063.

Nilsen TW, Graveley BR (2010). Expansion of the eukaryotic proteome by alternative splicing. **Nature** 463:457-63.

Ouali A, Gagaoua M, Boudida Y, Becila S, Boudjellal A, Herrera-Mendez CH, Sentandreu MA (2013). Biomarkers of meat tenderness: present knowledge and perspectives. **Meat Science** 95(4):854-70.

Pereira TC (2017). Visão geral dos ncRNAs na célula. In.: Pereira TC (Ed.) **Introdução ao universo dos non-coding RNAs**. Ribeirão Preto: Sociedade Brasileira de Genética, p.26-27.

Picard B, Lebret B, Cassar-Malek I, Liaubet L, Berri C, Le Bihan-Duval E, Hocquette JF, Renand G (2015). Recent advances in omic technologies for meat quality management. **Meat Science** 109:18–26.

Pickworth CL, Loerch SC, Velleman SG, Pate JL, Poole DH, Fluharty FL. (2010) Adipogenic differentiation state-specific gene expression as related to bovine carcass adiposity. **Journal Animal Science** 89(2):355-366.

Ponsuksili S, Murani E, Phatsara C, Schwerin M, Schellander K, Wimmers K. (2009). Porcine muscle sensory attributes associate with major changes in gene networks involving CAPZB, ANKRD1, and CTBP2. **Funct Integr Genomics** 9(4):455–71.

Potenza E, Racchi ML, Sterck L, Coller E, Asquini E, Tosatto SCE, Velasco R, Van de Pee Y, Cestaro A (2015). Exploration of alternative splicing events in ten different grapevine cultivars. **BMC Genomics** 16(1):706.

Ramayo-Caldas Y, Mach N, et al. (2012) Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. **BMC Genomics** 13(547):2-18.

Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, Mason CE, Socci ND, Betel D (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. **Genome Biology** 14(9):R95.

Sammeth, M., Foissac, S. & Guigó, R. A general definition and nomenclature for alternative splicing events. *Plos Comp Biol*. 4:e1000147 (2008).

Scholz AM, Bünger L, Kongsro J, Baulain U, Mitchell AD (2015) Non-invasive methods for the determination of body and carcass composition in livestock: dual-energy X-ray absorptiometry, computed tomography, magnetic resonance imaging and ultrasound: invited review. **Animal** 9(7):1250–1264.

Silva-Vignato B, Coutinho LL, Cesar ASM, Poleti MD, Regitano LCA, Balieiro JCC (2017). Comparative muscle transcriptome associated with carcass traits of Nelore cattle. **BMC Genomics** 18:506.

Sudhagar A, Kumar G, El-Matbouli M. (2018). Transcriptome Analysis Based on RNA-Seq in Understanding Pathogenic Mechanisms of Diseases and the Immune System of Fish: A Comprehensive Review. **Int J Mol Sci** 19(1):245.

Tang F, Lao K, Surani MA. (2011). Development and applications of single-cell transcriptome analysis. **Nature methods** 8(4 Suppl):S6-11.

Tizioto PC, Coutinho LL, et al. (2015) Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC Genomics** 16(242):1-14.

Tizioto PC, Coutinho LL, et al. (2016) Gene expression differences in Longissimus muscle of Nelore steers genetically divergent for residual feed intake. **Scientific Reports** 6:39493.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. **Nature Biotechnology** 28(5):511-515.

Wang XG, Ju ZH, et al. (2016). Deciphering transcriptome and complex alternative splicing transcripts in mammary gland tissues from cows naturally infected with *Staphylococcus aureus* mastitis. **PLoS One** 11:e0159719.

Wood JD, Enser M, Fisher AV, Nute GR, Sheard PR, Richardson RI, Hughes SI, Whittington FM (2008). Fat deposition, fatty acid composition and meat quality: A review. **Meat Science** 78(4):343-358.

Zhang W, Tong H, Zhang Z, Shao S, Liu D, Li S, Yan Y (2018) Transcription factor EGR1 promotes differentiation of bovine skeletal muscle satellite cells by regulating MyoG gene expression. **J Cell Physiol** 233(1):350-362.

CHAPETER 2 - Transcriptome profiling of muscle in Nelore cattle phenotypically divergent for the ribeye muscle area^a

ABSTRACT – Background: During growth, animals undergo changes in tissue proportions, especially in the muscle tissue. Ribeye muscle area (REA) is an indicator of carcass composition and is related to weight gain and animal musculature. Identifying genes and understanding the molecular mechanisms involved in muscle growth may aid the selection of animals with favorable genetics for REA, thus contributing to the improvement in meat quality. RNA-Seq is approach to identify transcripts from candidate genes that can cause variations in metabolic pathways, resulting in different phenotypes. Therefore, this study aimed to use RNA-Seq to identify differentially expressed genes (DEGs) in longissimus thoracis of uncastrated Nelore males phenotypically divergent for REA. A total of 80 animals were phenotyped for REA, and 15 animals each with the highest REA (HREA) and the lowest REA (LREA) were selected for analyses. **Results:** The transcriptomics dataset was aligned to a reference genome, and 288 DEGs were identified (q-value ≤ 0.05): 182 up-regulated and 106 down-regulated in the HREA group. Genes belonging to important families related to muscle cell growth, development, motility, and proteolysis, such as actin, myosin, collagen, integrin, solute carrier, ubiquitin, and kelch-like, were among the DEGs. Gene set enrichment analysis was performed, and it showed that many of the significantly enriched gene ontology (GO) terms are closely associated with muscle development, growth, and degradation. Through co-expression network analysis, we predicted three hub genes (PPP3R1, FAM129B and UBE2G1). These genes are involved in muscle growth and proteolysis. Functional enrichment analysis showed that PPP3R1 (down-regulated in the HREA group) functions in cGMP-PKG signaling (bta04022), glucagon signaling (bta04922), and oxytocin signaling (bta04921) pathways. **Conclusions:** DEGs and molecular processes involved in muscle growth and degradation were identified. Therefore, the genes expression levels and its biological process found this study may result in differences in muscle deposition, and therefore, Nelore animals with different REA proportions. **Keywords:** Bovine, carcass trait, co-expression, genes, RNA-Seq

Background

During growth, animals increase in weight and size and undergo changes in tissue proportions, especially the proportion of muscle tissue. REA is an indicator of carcass composition and is related to weight gain, animal musculature and meat quality traits [1]. Boinin et al. [2] identified great variability in the genetic potential of Nelore bulls to produce offspring with large REA. The identification and selection of genetically superior animals for REA could result in improved carcass quality, and consequently, an increased yield of cuts with high commercial value [3].

^aThis chapter corresponds to the manuscript submitted to "Genetics Selection Evolution" at November 13, 2018.

In beef cattle, muscle development and growth, which involve hyperplasia (increasing the number of muscle fibers) and hypertrophy (increasing the size of existing fibers), are influenced by castration; castrated and uncastrated males differ in muscle deposition, and consequently, REA size. The most of the meat consumed by European countries, produced and exported by Brazil, is composed of uncastrated Nelore cattle slaughtered less than two years old and grass-fed based. Production systems based on uncastrated cattle are interesting due to the better performance in relation to the castrated. Uncastrated cattle have a higher weight gain, lower fat content and higher feed conversion ratio [4]. According to Morgan et al [5], uncastrated animals had higher rates of protein synthesis and lower rates of protein degradation compared to castrated animals, which explains the larger REA in uncastrated animals. In addition, muscle deposition is also controlled by nutritional factors and biological processes regulated by multiple genes and a variety of metabolic pathways, as previously reported by Silva-Vignato et al. [6] which may result in animals with different REA proportions.

Tools such as next generation sequencing (NGS) and high-density panels of single nucleotide polymorphisms (SNPs), which can be used to identify SNPs, genes, and regions associated with carcass traits, may improve the accuracy of animal selection and thereby improve important traits, for example REA, in beef cattle. NGS applied to tissue transcriptomes (RNA-Seq) is an approach to screen for functional candidate genes, and to identify differentially expressed genes (DEGs) and important molecular mechanisms underlying the variation in pathways that result in different tissue phenotypes [7]. This technology has been used to identify DEGs associated with: tenderness [8] in a commercial population of uncastrated Nelore males; and ribeye area and backfat thickness [6] in Nelore steers (castrated males) in a breeding herd.

REA development and growth are influenced by complex gene network interactions in the muscle, but there are still genes and molecular mechanisms involved that remain to be elucidated, especially hub genes that are proposed to play an important role in the overall biology of organisms [9]. Therefore, the aim of this study was to use an RNA-Seq approach to identify DEGs and their enriched biological processes associated with REA of the uncastrated Nelore

cattle. These animals represent a commercial population of bovines that are exported to different countries. In addition, gene co-expression network analysis was performed to identify hub genes associated with the REA phenotype.

Methods

Animals and phenotypes

This study involved 80 uncastrated Nelore males from the Capivara Farm (São Paulo state, Brazil), which participates in the Nelore Qualitas Breeding Program. The animals remained together from birth until slaughter and were reared on grazing systems (foraging of *Brachiaria* sp. and *Panicum* sp. with free access to mineral salts). The animals were finished in confinement for approximately 90 days. The diet was based on whole-plant silage and mix of sorghum grain, soybean meal or sunflower seeds were used as concentrate, with a concentrate / roughage ratio from 50/50 to 70/30. Animals were slaughtered at an average age of 24 months in a commercial slaughterhouse, in accordance with the Brazilian Federal Inspection Service procedures.

Longissimus thoracis muscle samples were collected from between the 12th and 13th ribs on the left side of each carcass at two time points: at slaughter for RNA-sequencing analysis (as described by Fonseca et al. [8]), and 24 h after slaughter for the evaluation of REA. REA was measured using the grid method in units of squared centimeters [10]. Out of the 80 animals used in this study, samples from 15 animals with the highest and 15 animals with the lowest REA (HREA and LREA groups) were selected for analyses. A Student's t-test was used to evaluate whether a significant difference between the selected groups existed (HREA and LREA) (Table 1).

Table 1. Descriptive statistics for ribeye muscle area phenotypes in Nelore cattle

Phenotype	N	Mean±SD	Minimum	Maximum
HREA (cm ²)	15	83.66±2.55	81	88
LREA (cm ²)	15	65.73±2.05	61	68

There is a significant difference between means (Student's t-test $p \leq 0.01$). SD = standard deviation. HREA=Highest ribeye muscle area; LREA=Lowest ribeye muscle area; N = Number of observations.

RNA-Seq library preparation

Total RNA was extracted from approximately 50 mg of muscle tissue using the RNeasy Lipid Tissue Mini Kit (Qiagen, Valencia, 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 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.

Analysis of RNA sequencing data

The quality of RNA-Seq reads (quality scores, GC content, N content, length distributions, duplication levels, overrepresented sequences and K-mer content) was checked using FastQC (v.0.11.4) software [11]. The reads that had low quality were trimmed using Trimmomatic (v.0.36) [12], following parameters: cut of adapter, other illumina-specific sequences from the read, bases off the start of a read, bases off the end of a read; read to a specified length by removing bases from the end, sliding window trimming, drop the read if the average quality is below the specified level; and the read if it is below a specified length. After filtering, HISAT2 (v.2.0.5) [13] was used to map paired-end reads. The HISAT2 index for mapping was built from a multi-sequence FASTA file containing the bovine genome (UMD3.1.1 *Bos Taurus*) and Y chromosome (*Btau* 4.6.1), both deposited in the National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/>).

The Cufflinks2 v.2.1.1 suite of tools [14] was used for transcriptome assembly and differential expression analysis. Cufflinks2 was used to assemble the aligned reads for each sample and to estimate the number of transcripts,

expressed as fragments per kilobase of transcript per million reads mapped (FPKM). After initial assembly, the Cufflinks2 output for each sample was merged using Cuffmerge2 [14]. We obtained a set of uniform transcripts for every sample and then ran Cuffdiff2 [14] to identify differently expressed genes (DEGs) between REA groups. False discovery rates (FDR) were controlled using the Benjamini-Hochberg [15] method, in which we considered a gene to be differentially expressed if it had a q-value ≤ 0.05 .

Co-expression network analysis and hub gene identification

Co-expression analyses were performed using applications from Cytoscape v.3.4 [16]. The co-expression network was generated using the ExpressionCorrelation plugin [17] and the similarity matrix (normalized expression counts of DEGs) was computed using the Pearson Correlation. A histogram tool was used for the screening criteria with a node score cut-off ≥ 0.9 and ≤ -0.9). The metrics of network: degree distribution; betweenness centrality; network density and clustering coefficient, were determined using the Network Analyzer plugin [18]. The networks that have density and clustering values close to one contain many edges and no isolated nodes [19,20]. Genes with the most interactions (highest maximal clique centrality - MCC), were identified as hub genes using the Cytohubba plugin [21].

Gene set enrichment analysis

Gene set enrichment analysis was performed on all the DEGs using DAVID 6.8 (Database for Annotation, Visualization and Integrated Discovery) [22]. The Benjamini-Hochberg method [15] was used to determine p values. The DAVID Functional Annotation Tool was used to find the most relevant Gene Ontology (GO) terms in the biological process category. DAVID Pathway was used to map the enriched pathways using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [23].

Results

Alignment statistics

The mean number of paired-end reads per sample after filtering was approximately 22.5 million reads. Of the trimmed reads, 88% uniquely mapped to

either the bovine reference genome UMD3.1.1 *Bos Taurus* or Y chromosome of *Btau* 4.6.1. The overall alignment rate was 96.5% (approximately 21.6 million reads mapped in pairs) with sequencing coverage of 45x (coverage for all transcripts of all samples), for more details, see Additional File: Table S1. Additional file: Figure S2 shows a box plot containing the transformed FPKM values (log₁₀) for each group and the plot of principal component analysis (PCA). These plots were constructed using the cummerbund package [24]. 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 and close to -1, indicating that the level of sequencing coverage permitted the identification of low-expressed genes [25, 26]. PCA showed the formation of different groups (highest and lowest REA), indicating differences in the expression of genes between the HREA and LREA groups.

Differentially expressed genes (DEGs)

This study identified 288 DEGs ($q \leq 0.05$) in HREA animals; 182 up-regulated and 106 down-regulated genes (Additional File: Table S2). Only four of these genes were unannotated in cattle. Transcription factors, non-coding RNAs, pseudogenes, transporters and several protein coding genes related to the function, development, growth, proteolysis and motility of muscle cells were among the DEGs. The chromosomal locations of the DEGs were analyzed and found to be distributed over all the bovine chromosomes except the mitochondrial (MT) and Y chromosomes, albeit in different proportions (Figure 1).

The DEGs belonged to important gene families, such as actin (*ACT*), myosin (*MYH*), collagen (*COL*), integrin (*ITG*), solute carrier (*SLC*), ubiquitin (*UBE*) and kelch-like (*KLHL*). All DEGs belonging to the actin, collagen and integrin families were up-regulated in HREA animals. All DEGs belonging to the ubiquitin family were down-regulated in HREA animals. The other families (myosin, solute carrier and kelch-like) contained DEGs that were both up- and down-regulated in HREA animals. The DEGs included *MSTN* (myostatin) that down-regulated in HREA animals, has been identified as the cause of the double-muscling phenotype in cattle [27]; therefore, this gene is considered a molecular marker for muscle growth.

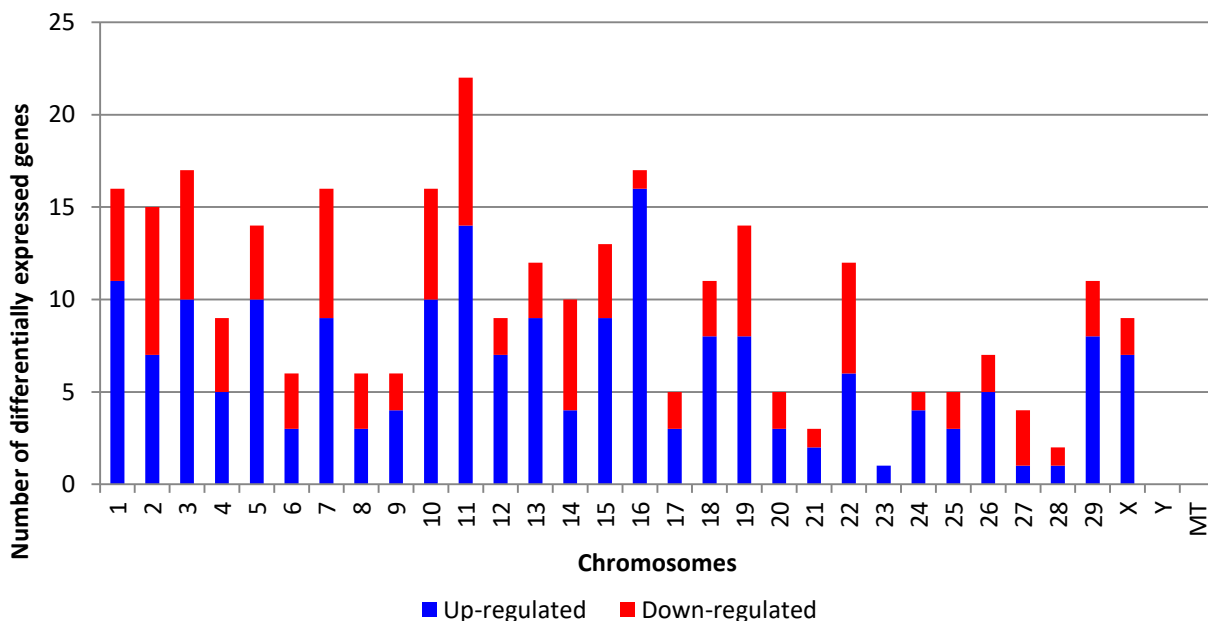


Fig. 1 Distribution of the differentially expression genes across the chromosomes. Blue bars represent up-regulated genes and red bars down-regulated genes in the highest ribeye muscle area group.

Co-expression network analysis and hub gene identification

The co-expression network (Figure 2) had: gene pairs (edges) = 400; genes (nodes) = 150; density = 0.036; clustering coefficient = 0.327. When network maximal clique centrality (MCC) was examined, the top three genes with the highest MCC scores were considered as the hub genes: *PPP3R1* (Protein Phosphatase 3 Regulatory Subunit B, Alpha) *UBE2G1* (Ubiquitin Conjugating Enzyme E2 G1) and *FAM129B* (Family with Sequence Similarity 129 Member B). According to Langfelder et al. [9], hub genes are expected to play an important role in the biology of the organism studied. *PPP3R1* and *UBE2G1* genes were down-regulated in HREA group. *PPP3R1* encodes a phosphatase protein; a regulatory subunit of Calcineurin that plays a role in neuronal calcium signaling [28]. E2 ubiquitin-conjugating enzymes, such as *UBE2G1* (down-regulated in HREA animals), play a role in the second step of the ubiquitination reaction that targets a protein for degradation via the proteasome [29]. *FAM129B* was up-regulated in the HREA group; it functions as a target of the MAPK (Erk1/2) signaling cascade [30]. The hub genes found may have a putative effect on the muscle metabolism of Nelore cattle.

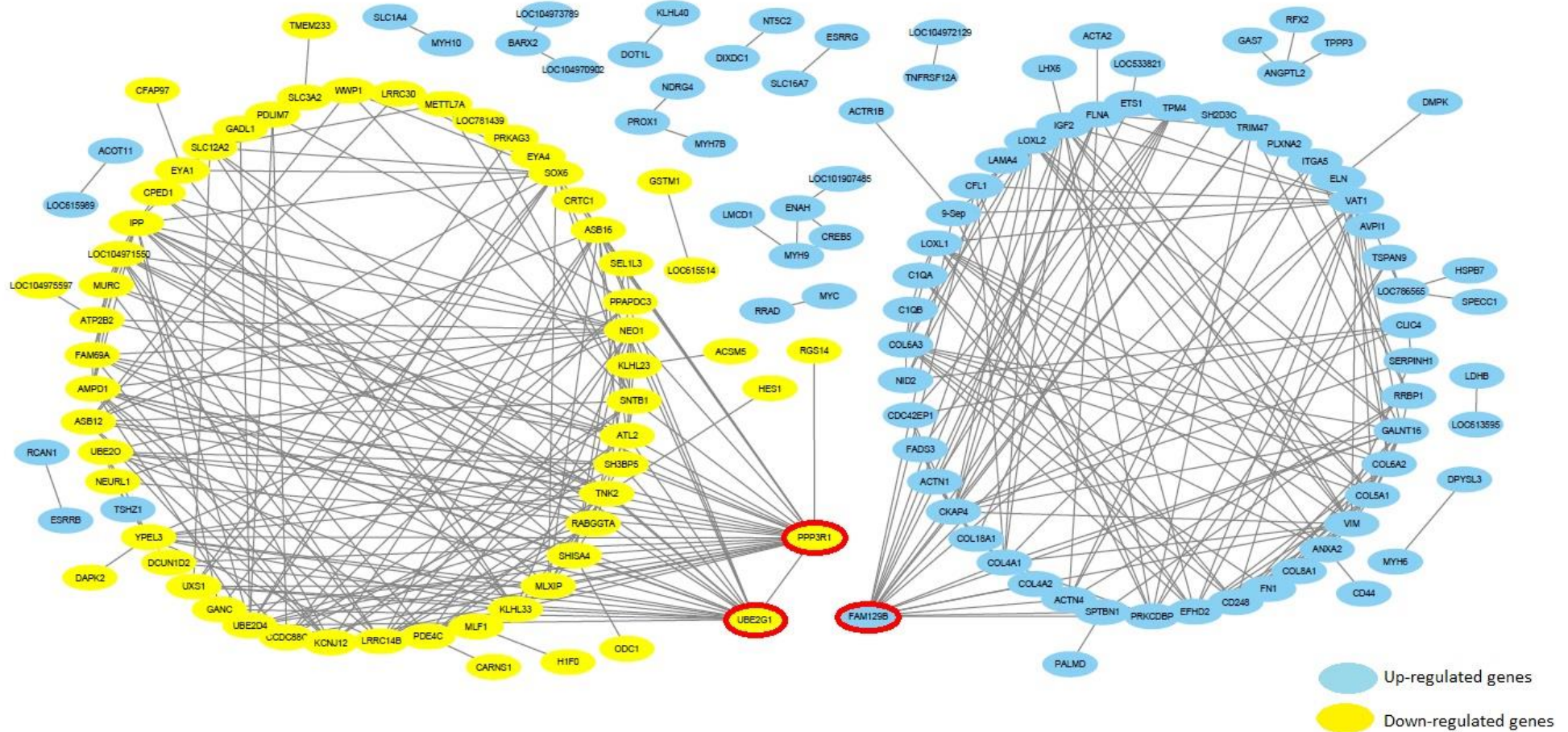


Fig. 2 Gene co-expression network for the differentially expressed genes in muscle of Nelore cattle divergent for ribeye muscle area. Genes circled in red were considered hubs.

Gene set enrichment analysis

The gene set enrichment analysis performed using DAVID software showed that 41 GO terms were significantly enriched in the biological process category (Additional File: Table S3). Many of the significantly enriched GO terms are closely associated with muscle development, growth and degradation. Genes belonging to the actin (*ACTA2*, *ACTN1*, and *ACTN4*), myosin (*MYH10*), collagen (*COL18A1*), and integrin (*ITGA5*) families, in addition to genes that encode extracellular matrix proteins (*LAMA4*, *TNC*, and *FN1*) and transcription factors (*MYOG*, *PROX1*, and *SOX6*), contributed to the enrichment of GO terms related to muscle development and growth, for example: sarcomere organization (GO:0045214); skeletal muscle fiber development (GO:0048741); extracellular matrix organization (GO:0030198); actin crosslink formation (GO:0051764); cell-cell adhesion mediated by integrin (GO:0033631); cell adhesion (GO:0007155); actin filament bundle assembly (GO:0051017); muscle organ development (GO:0007517); skeletal muscle thin filament assembly (GO:0030240); and cartilage development (GO:0051216) (Additional File: Table S3).

The canonical Wnt signaling pathway (GO:0060070), which has been associated with myogenesis and adipogenesis in fetal ruminant animals [31], was also enriched in the biological process category. In addition, significantly enriched biological process GO terms related to muscle degradation, protein ubiquitination (GO:0016567) and protein K63-linked ubiquitination (GO:0070534) were found. *UBE* and *KLHL* gene families contributed to the enrichment of these terms. DEGs were involved in 18 over-represented pathways (Additional File: Table S3), identified using DAVID software (KEGG pathway). Actin, myosin and collagen families contributed to the over-representation of focal adhesion (bta04510), PI3K-Akt signaling (bta04151) and extracellular matrix (ECM)-receptor (bta04512) pathways. The hub gene *PPP3R1* was present in cGMP-PKG (bta04022), glucagon (bta04922) and oxytocin (bta04921) signaling. These GO terms and pathways indicate that DEGs might play important roles in muscle metabolism and deposition.

Discussion

Several of the DEGs belonged to the actin, myosin, collagen, integrin, solute carrier, ubiquitin and kelch-like families. Actin and Myosin proteins play a

role in muscle contraction and are the main components of myofilaments (organelles that constitute the muscle cytoskeleton) [32]. It has previously been shown that genes belonging to the actin and collagen families are differentially expressed in muscle tissue of Nelore cattle phenotypically divergent for the residual feed intake [26]. Cesar et al. [33] found the expression of myosin genes in muscle tissue was associated with intramuscular fat content in Nelore cattle. All DEGs belonging to the actin, myosin (with the exception of a single gene), and collagen families were up-regulated in HREA animals (see Additional File: Table S2).

Integrins are adhesion proteins are present on cell membranes [34]. Wang et al. [35] identified integrin family genes associated with carcass traits in a broiler chicken population. Genes belonging to the solute carrier family play a role as transporters [36]. Solute carrier genes have been associated (in a GWAS, genome-wide association study) with REA in Nelore cattle [37]. Ubiquitin proteins play an important role in muscle loss [29] and genes belonging to this family were found previously differentially expressed in muscle tissue and associated with meat tenderness in Nelore cattle [8]. Kelch-like proteins are known to act as substrate adaptors for Cullin 3 ubiquitin ligases [38], both of which play a role in ubiquitination and have a putative effect on muscle loss [29,38].

In this study, *MSTN* was differentially expressed this study. *MSTN* is considered a molecular marker for muscle growth in animals, and high concentrations of myostatin cause a decrease in normal muscle development [39]. Since our results showed that *MSTN* was down-regulated in the HREA group, this gene may contribute to the increased muscle growth in these animals. *MSTN* has previously been associated with different carcass and meat traits in cattle [40].

The functional enrichment analysis highlighted several biological processes related to muscle development, contraction and motility (Additional File: Table S3), for example: sarcomere organization (GO:0045214), skeletal muscle fiber development (GO:0048741), actin crosslink formation (GO:0051764), and actin filament bundle assembly (GO:0051017). The *ACTN1* gene (belonging to the actin family of proteins) appears under all of these GO terms. Other DEGs belonging to the actin family (*ACTA2* and *ACTN4*) also

appear under some of these biological processes. *MYH10* (a member of the myosin superfamily) contributed to the enrichment of cardiac myofibril assembly (GO:0055003), cell adhesion (GO:0007155), regulation of cell shape (GO:0008360), and tight junctions (bta04530). Actin, together with myosin and ATP molecules, generates cellular and muscular movements [32]. Therefore, *ACTA1*, *ACTN1*, and *MYH10* expression may indicate the considerable capacity of Nelore cattle to regulate cytokinesis, motility and polarity of muscle cells.

COL4A1, *COL4A2*, *COL6A2*, *COL6A3*, and *COL5A1* (collagen family) genes were involved in the regulation of protein kinase B (PI3k-Akt) signaling (bta04151). Akt is mediated by PI3k and is a key signaling protein in cell pathways that cause skeletal muscle hypertrophy and tissue growth in general [41]. Collagen gene family members, together with the integrin gene (*ITGA5*) and other interactive molecules, such as *LAMA4*, *TNC*, and *FN1* (all up-regulated in HREA group), were present under GO terms related to cell adhesion (GO:0007155) and the ECM-receptor interaction (bta04512) and focal adhesion (bta04510) pathways.

The ECM comprises connective tissue layers that surround muscle fibers and is composed of fibrous and non-fibrous proteins, including collagen and proteoglycans [42]. ECM-receptor interactions were associated with the rapid neonatal gain of skeletal muscle in piglets [43]. Focal adhesions, which are large and dynamic proteins, serve as mechanical links to the ECM and other molecules, playing key roles in important biological processes [43]. Focal adhesion kinase is involved in mammalian myoblast fusion both *in vitro* and *in vivo* [44], 2009). Our results further highlight the biological importance of focal adhesions and ECM-receptors in determining the composition and organization of muscle tissue in animals, in addition to their putative effect on different REA proportions in Nelore cattle.

Another important biological process identified was the canonical Wnt signaling pathway (GO:0060070). Experimental evidence indicates that Wnt signaling regulates mesenchymal stem cells differentiation. Skeletal muscle development mainly involves myogenesis, but it also involves adipogenesis and fibrogenesis. In fetal ruminant animals, muscle cell, adipocyte, and fibroblast development are mainly derived from mesenchymal stem cells. Up-regulation of

Wnt/ β -catenin promotes myogenesis, and down-regulation enhances adipogenesis [31].

Genes encoding the transcription factors MYOG (Myogenin), PROX1 and SOX6 were differentially expressed. *MYOG* contributed to the enrichment of biological processes related to muscle development (skeletal muscle fiber development GO:0048741, muscle organ development GO:0007517, skeletal muscle thin filament assembly GO:0030240, and cartilage development GO:0051216) and is necessary for embryonic myoblast differentiation. However, its expression may be continuous in the mature muscle tissue of animals, especially during oxidative metabolism in muscle, suggesting that *MYOG* may play a role throughout the animal's life [45]. *PROX1* was up-regulated in HREA animals and has been shown to play a significant role in skeletal muscle development. *PROX1* is an important transcription factor that functions in embryonic development and as a key-regulatory protein in neurogenesis and cardiac muscle development [46]. *PROX1* is expressed in zebrafish muscle [47] and in the *longissimus* of bulls of a Charolais x Holstein F2-cross with divergent intramuscular fat content [48].

SOX6 was up-regulated in the LREA group. *SOX6* suppresses the transcription of specific slow-fiber genes during development, thus playing a role in muscle differentiation, in addition to regulating the expression of transcriptional factors important for muscle development, such as *PROX1* [49]. This suggests that the overexpression of this gene may have decreased the expression of, for example, *MYOG* and *PROX1* in the LREA group. Our data suggest that *SOX6* plays the role of a specific negative regulator of skeletal muscle differentiation, possibly interfering with the function of myogenic proteins and transcription factors that regulate other genes related to muscle development and growth.

Functional enrichment analysis showed that *PPP3R1* (down-regulated in the HREA group) was a hub gene and that it participates in cGMP-PKG signaling (bta04022), glucagon signaling (bta04922) and oxytocin signaling (bta04921) pathways. cGMP-dependent protein kinase, or Protein Kinase G (PKG), is a serine/threonine-specific protein kinase that is activated by cGMP. PKG is known to modulate the nitrous oxide (NO) signaling pathway involved in smooth muscle relaxation [50]. The glucagon signaling pathway regulates energy and glucose homeostasis, and is the main regulatory mechanism opposing the effects of

Insulin [51]. Elabd et al. [52] reported that oxytocin is essential for the regeneration of skeletal muscle tissue and maintenance of homeostasis in mice. In addition to *PPP3R1*, the *OXT* gene was also found to play a role in the oxytocin signaling pathway. This gene encodes a precursor protein that is processed to produce Oxytocin and Neurophysin 1, and has been implicated in the regeneration of skeletal muscle tissue and maintenance of homeostasis in mice [53,54].

FAM129B (up-regulated in the HREA group), which was indicated as a hub gene in this study, may play a role in apoptosis suppression and is also a target of the mitogen-activated protein kinase (MAPK) signaling cascade [30]. MAPK signaling is responsible for cell proliferation, differentiation and apoptosis, and plays an important role in hyperplastic growth, which has a positive effect on meat tenderness [53, 54].

Functional enrichment analyses showed that three genes belonging to the Ubiquitin-protein ligase family (*UBE2D4*, *UBE2G1* and *UBE2O*); three genes belonging to the Kelch-like family (*KLHL23*, *KLHL34*, *KLHL33*); and four other genes: *IPP*, *KBTBD13*, *SOCS3*, and *ASB12* were involved in ubiquitination processes (protein ubiquitination: GO:0016567 and protein K63-linked ubiquitination: GO:0070534). The various conditions that lead to loss of muscle mass involve distinct cascades of intracellular signaling that may induce apoptosis, increase protein degradation, or decreased activation of satellite cells (responsible for muscle regeneration). Some proteolytic systems are involved in muscle degradation [55], including the ubiquitin-proteasome system, which is highly conserved and degrades the major proteins of contractile skeletal muscle and plays an important role in muscle loss [29].

UBE2G1 (down-regulated in the HREA group and consequently up-regulated in LREA group) was indicated as a hub gene in this study. This gene encodes an E2 ubiquitin-conjugating enzyme and has effects in ubiquitin-proteasome systems. Ubiquitin can be conjugated to specific protein substrates in a process that requires three enzymatic components: E1, a ubiquitin activating enzyme; E2, a ubiquitin conjugating enzyme; and E3, a ubiquitin-binding enzyme [6]. The process of degradation of polyubiquitinated proteins occurs in the proteasome (20S or 26S), which is a complex composed of one or three large enzymes that function to degrade unnecessary or damaged proteins in the cell

[57]. Fernandes Júnior et al. [37], in a GWAS (genome-wide association study) with Nelore cattle populations, showed that some genes associated with REA and back fat contributed to the over-representation of GO terms related to protein turnover. These results showed that ubiquitin proteins may affect REA in Nelore via effects on ubiquitination processes. All these processes may result in differences in muscle deposition, and therefore, animals with different REA proportions.

Conclusion

The differences in gene expression highlight the complexity of the *longissimus thoracis* muscle transcriptome in Nelore cattle with divergent REA phenotypes. Our results showed that differences in muscle deposition (different REA proportions) in Nelore cattle may be controlled by genes (acting as positive or negative regulators) that function in various biological processes that regulate muscle growth, development, and degradation. The results of this study, together with previous reports, may help clarify the molecular processes involved in muscle deposition in Nelore cattle.

List of abbreviations

DEGs: Differentially expressed genes; REA: Ribeye muscle area; HREA: Highest ribeye muscle area; LIF: Lowest ribeye muscle area; MCC: Maximal clique centrality; FDR: False discovery rates; FPKM: Fragments per kilobase of transcript per million reads mapped; GO: Gene ontology; GWAS: Genome-wide association study. SNPs: Single nucleotide polymorphisms; NGS: Next generation sequencing; *PPP3R1*: Protein Phosphatase 3 Regulatory Subunit B, Alpha; *UBE2G1*: Ubiquitin Conjugating Enzyme E2 G1); *FAM129B*: Family With Sequence Similarity 129 Member B; *ACTA2*, *ACTN1*, and *ACTN4*: Actin family members; *COL4A1*, *COL4A2*, *COL6A2*, *COL6A3*, and *COL5A*: collagen family members; *KLHL23*, *KLHL34* and *KLHL33*: Kelch-like family members; *UBE2D4*, *UBE2G1* and *UBE2O*: Ubiquitin-protein ligase family members; *MSTN*: Myostatin; *MYH10*: Myosin Heavy Chain 10; *ITGA5*: Integrin Subunit Alpha 5; *LAMA4*: Laminin Subunit Alpha 4; *TNC*: Tenascin C; *FN1*: Fibronectin 1; *MYOG*: Myogenin; *PROX*: Prospero Homeobox 1; *SOX6*: SRY-Box 6; *IPP*: Intracisternal A Particle-Promoted Polypeptide; *KBTBD13*: Kelch Repeat And

BTB Domain Containing 13; SOCS3: Suppressor Of Cytokine Signaling 3; and ASB12: Ankyrin Repeat And SOCS Box Containing 12.

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 and grant #2015/16850-9).

Acknowledgements

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

Additional Files

Additional File: Figure S1

Additional File: Table S1

Additional File: Table S2

Additional File: Table S3

References

1. Miar Y, Plastow GS, Bruce HL, Moore SS, Durunna ON, Nkrumah JD, et al. Estimation of genetic and phenotypic parameters for ultrasound and carcass merit traits in crossbred beef cattle. *Can J Anim Sci.* 2014;94(2):273-80.
2. Bonin MN, Ferraz JB, Eler JP, Rezende FM, Cucco DC, Carvalho ME, et al. Sire effects on carcass and meat quality traits of young Nelore bulls. *Genet Mol Res.* 2014;13(2):3250-64.
3. Goodall JJ, Schmutz SM. IGF2 gene characterization and association with rib eye area in beef cattle. *Animal genetics.* 2007;38(2):154-161.
4. Restle J, Vaz FN, Alves Filho DC. Machos não-castrados para a produção de carne. In: RESTLE, J. (Ed.) Confinamento, pastagens e suplementação para produção de bovinos de corte. Santa Maria: Imprensa Universitária. p.215-231. 1999.

5. Morgan JB, Wheeler TL, Koohmaraie M, Crouse JD, Savell JW. Effect of castration on myofibrillar protein turnover, endogenous proteinase activities, and muscle growth in bovine skeletal muscle. *J Anim Sci*. 1993;71(2):408-14.
6. Silva-Vignato B, Coutinho LL, Cesar ASM, Poleti MD, Regitano LCA, Balieiro JCC. Comparative muscle transcriptome associated with carcass traits of Nellore cattle. *BMC Genomics*. 2017;18:506.
7. He H, Liu X. Characterization of transcriptional complexity during longissimus muscle development in bovines using high-throughput sequencing. *PLoS One*. 2013;8:e64356.
8. Fonseca LFS, Gimenez DFJ, Silva DBS, Barthelson R, Baldi F, Ferro JA, et al. Differences in global gene expression in muscle tissue of Nellore cattle with divergent meat tenderness. *BMC Genomics*. 2017;18(945):1-12.
9. Langfelder P, Mischel PS, Horvath S. When is hub gene selection better than standard meta-analysis? *PLoS One*. 2013;8:e61505.
10. United States Department of Agriculture – USDA. Official United States standards for grades of carcass beef. Washington, D.C.: Agricultural Marketing Service. (1997).
11. Andrews, S. FastQC: a quality control tool for high throughput sequence data. 2010. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Accessed 15 Jan 2017.
12. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*. 2014;30:2114-20.
13. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12:357–360.
14. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks. *Nat Protocols*. 2012;(3):562-578.
15. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc*. 1995;57(1):1-12.
16. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-504.
17. Hui S, Sander C, Potylitsine E, Whitaker W, Bader G, Morrison L, et al. ExpressionCorrelation. 2015. Available online at: <http://apps.cytoscape.org/apps/expressioncorrelation>
18. Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. *Bioinformatics*. 2008;24(2):282-4.
19. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet*. 2004;5:101-113.
20. Dong J, Horvath S. Understanding network concepts in modules. *BMC Syst Biol*. 2007;1:24.
21. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8(4):11.
22. Huang W, Sherman BT, Lempick RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocols*. 2009;4:44 – 57.
23. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000;28(1):27-30.
24. Goff L, Trapnell C, Kelley D. cummeRbund: Analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data. 2018. R package version 2.24.0.
25. Chapple RH, Tizioto PC, Wells KD. Characterization of the rat developmental liver transcriptome. *Physiol Genomics*. 2013;45:301–311.

26. Tizioto PC, Coutinho LL, Oliveira PSN, Cesar AS, Diniz WJ, Lima AO, et al. Gene expression differences in Longissimus muscle of Nelore steers genetically divergent for residual feed intake. *Sci Rep.* 2016; 6:39493.
27. Grobet L, Martin LJ, Poncelet D, Pirottin D, Brouwers B, Riquet J, et al. A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nat Genet.* 1997;17(1):71-4.
28. Baumgärtel K, Mansuy IM. Neural functions of calcineurin in synaptic plasticity and memory. *Learning & Memory.* 2012;19:375-384.
29. Attaix D, Ventadour S, Codran A, Béchet D, Taillandier D, Combaret L. The ubiquitin-proteasome system and skeletal muscle wasting. *Essays Biochem.* 2005;41:173-86.
30. Chen S, Evans HG, Evans DR. FAM129B/MINERVA, a novel adherens junction-associated protein, suppresses apoptosis in HeLa cells. *J Biol Chem.* 2011;286(12):10201-9.
31. Du M, Tong J, Zhao J, Underwood KR, Zhu M, Ford SP, Nathanielsz PW. Fetal programming of skeletal muscle development in ruminant animals. *J Ani Sci.* 2010;88:E51–E60.
32. Sgarbieri VC. Proteínas em alimentos protéicos. São Paulo: Varela; 1996.
33. Cesar AS, Regitano LC, Koltjes JE, Fritz-Waters ER, Lanna DP, Gasparin G, et al. Putative regulatory factors associated with intramuscular fat content. *PLoS One.* 2015;10 (6):1-21.
34. Giancotti FG, Ruoslahti E. Integrin signaling. *Science.* 1999;285(5430):1028–1032.
35. Wang W, Zhang T, Wang J, Zhang G, Wang Y, Zhang Y, et al. Genome-wide association study of 8 carcass traits in Jinghai Yellow chickens using specific-locus amplified fragment sequencing technology. *Poultry Science.* 2016;95(3):500–506.
36. Pérez de Heredia F, Wood IS, Trayhurn P. Hypoxia stimulates lactate release and modulates monocarboxylate transporter (MCT1, MCT2, and MCT4) expression in human adipocytes. *Pflugers Arch.* 2010;459:509–518.
37. Fernandes Júnior GA, Costa RB, de Camargo GM, Carneiro R, Rosa GJ, Baldi F, et al. Genome scan for postmortem carcass traits in Nelore cattle. *J Anim Sci.* 2016;94(10):4087-4095.
38. Dhanoa BS, Cogliati T, Satish AG, Bruford EA, Friedman JS. Update on the Kelch-like (KLHL) gene family. *Hum Genomics.* 2013;7:13.
39. Garikipati DK, Gahr SA, Roalson EH, Rodgers BD. Characterization of rainbow trout myostatin-2 genes (rtMSTN-2a and -2b): genomic organization, differential expression, and pseudogenization. *Endocrinology;* 2007;148(5):2106-15.
40. Fiems LO. Double Muscling in Cattle: Genes, Husbandry, Carcasses and Meat. *Animals.* 2012;2(3):472-506.
41. Fujita S, Dreyer HC, Drummond MJ, Glynn EL, Cadenas JG, Yoshizawa F, et al. Nutrient signalling in the regulation of human muscle protein synthesis. *The Journal of Physiology.* 2007;582(Pt 2):813-823.
42. Mariman ECM, Ping W. Adipocyte extracellular matrix composition, dynamics and role in obesity. *CMLS.* 2010;67:1277–1292.
43. Jiang C, Shi P, Li S, Dong R, Tian J, Wei J, Luo S. Gene Expression Profiling of Skeletal Muscle of Nursing Piglets. *Int J Biol Sci.* 2010; 6(7):627–638.
44. Quach NL, Biressi S, Reichardt LF, Keller C, Rando TA. Focal Adhesion Kinase Signaling Regulates the Expression of Caveolin 3 and β 1 Integrin, Genes Essential for Normal Myoblast Fusion. *Mol Biol Cell.* 2009; 20(14): 3422–3435.
45. Zhu LN, Ren Y, Chen JQ, Wang YZ. Effects of myogenin on muscle fiber types and key metabolic enzymes in gene transfer mice and C2C12 myoblasts. *Gene.* 2013;532(2):246-52.

46. Petchey LK, Risebro CA, Vieira JM, Roberts T, Bryson JB, Greensmith L. Loss of Prox1 in striated muscle causes slow to fast skeletal muscle fiber conversion and dilated cardiomyopathy. *Proc Natl Acad Sci.* 2014;111(26):9515-20.
47. Roy S, Wolff C, Ingham PW. The u-boot mutation identifies a Hedgehog-regulated myogenic switch for fiber-type diversification in the zebrafish embryo. *Genes Dev.* 2001;15:1563–76.
48. Komolka K, Ponsuksili S, Albrecht E, Kühn C, Wimmers K, Maak S. Gene expression profile of Musculus longissimus dorsi in bulls of a Charolais x Holstein F₂-cross with divergent intramuscular fat content. *Genomics Data.* 2016;7:131–133.
49. An CI, Dong Y, Hagiwara N. Genome-wide mapping of Sox6 binding sites in skeletal muscle reveals both direct and indirect regulation of muscle terminal differentiation by Sox6. *BMC Dev Biol.* 2011;11:59.
50. Lincoln TM, Dey N, Sellak H. Invited review: cGMP-dependent protein kinase signaling mechanisms in smooth muscle: from the regulation of tone to gene expression. *J Appl Physiol (1985).* 2001;91(3):1421-30
51. Jiang G, Zhang BB. Glucagon and regulation of glucose metabolism. *Am J Physiol Endocrinol Metab.* 2003;284:E671–E678.
52. Elabd C, Cousin W, Upadhyayula P, Chen RY, Chooljian MS, Li J, et al. Oxytocin is an age-specific circulating hormone that is necessary for muscle maintenance and regeneration. *Nature communications.* 2014;5:4082.
53. Ponsuksili S, Murani E, Phatsara C, Schwerin M, Schellander K, Wimmers K. Porcine muscle sensory attributes associate with major changes in gene networks involving CAPZB, ANKRD1, and CTBP2. *Funct Integr Genomics.* 2009; 9(4):455–71.
54. Taye M, Kim J, Yoon SH, Lee W, Hanotte O, Dessie T, et al. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. *BMC Genet.* 2017;18(1):11.
55. Purintrapiban J, Wang MC, Forsberg NE. Degradation of sarcomeric and cytoskeletal proteins in cultured skeletal muscle cells. *Comp Biochem Physiol B Biochem Mol Biol.* 2003;136(3):393-401.
56. Li YP, Chen YL, Li AS, Reid MB. Hydrogen peroxide stimulates ubiquitin-conjugating activity and expression of genes for specific E2 and E3 proteins in skeletal muscle myotubes. *Am J Physiol Cell Physiol.* 2003;285(4):C806-C12.
57. Hasselgren PO, Wray C, Mammen J. Molecular regulation of muscle cachexia: It may be more than the proteasome. *Biochem Biophys Res Commun.* 2002;290(1):1-10.

Additional Files

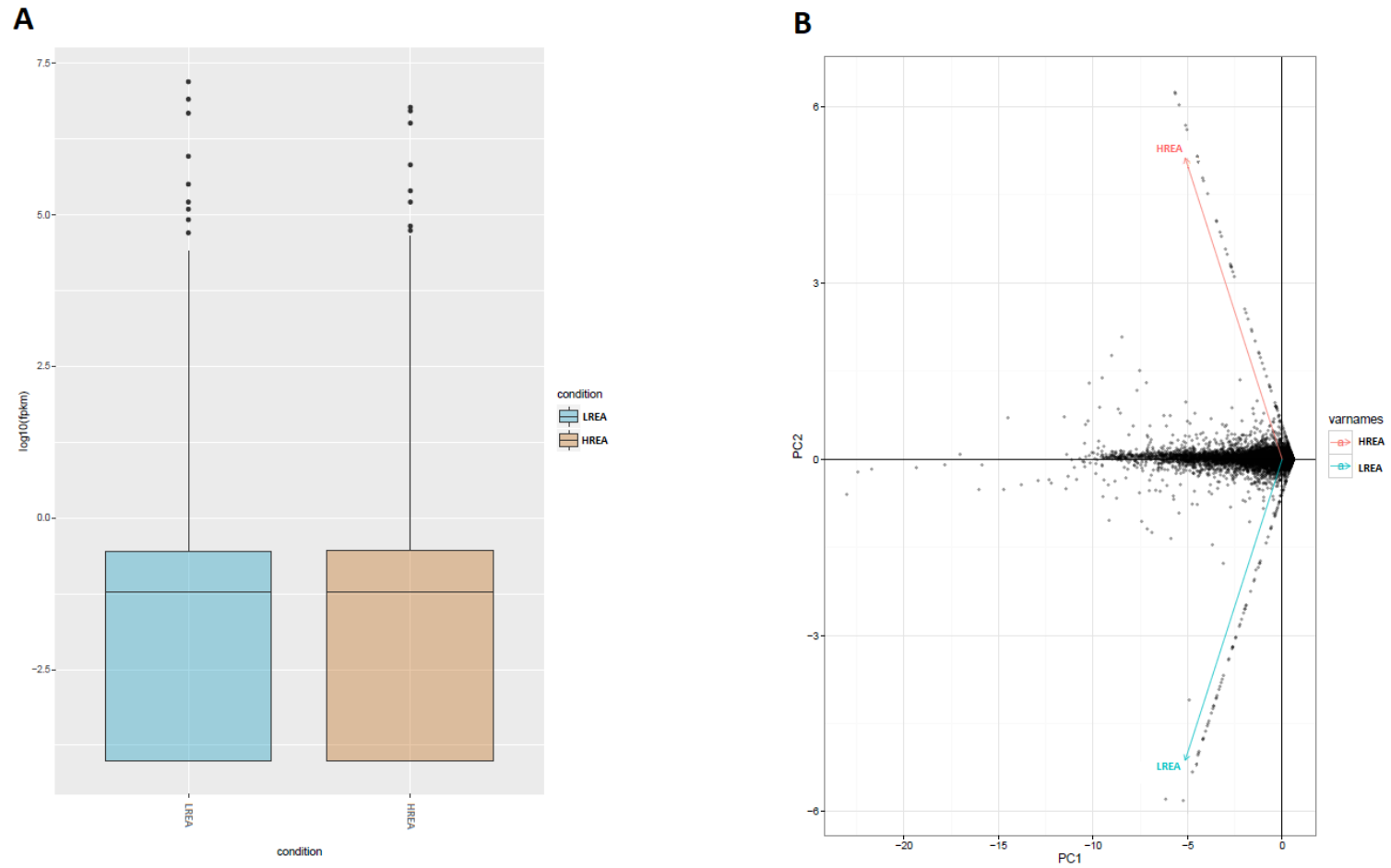


Figure S1. Global statistics and quality control. (A) Box plot of FPKM distributions for individual conditions. (B) PCA plot for gene-level features.

Table S1. Descriptive statistics for alignment parameters

Highest ribeye area (HREA)				Lowest ribeye area (LREA)			
Sequence ID	Reads raw	Reads cleaned (Mb)	Overall alignment rate (%)	Sequence ID	Reads raw	Reads cleaned (Mb)	Overall alignment rate (%)
N1	17,525,193	14,121,641	96.56	N16	37,389,485	34,511,728	96.53
N2	27,719,519	25,817,447	96.23	N17	12,069,555	9,706,541	96.18
N3	12,604,475	10,209,149	96.53	N18	37,716,840	30,583,246	96.40
N4	32,645,772	30,,068,141	96.22	N19	36,961,023	34,576,999	96.13
N5	35,983,786	33,405,266	96.95	N20	28,723,235	26,921,059	96.23
N6	26,222,229	21,813,841	96.10	N21	15,096,673	12,227,626	96.72
N7	13,897,839	11,316,252	96.63	N22	16,105,988	13,342,281	96.74
N8	31,387,621	26,298,812	96.31	N23	14,024,452	11,582,059	96.48
N9	12,301,661	9,933,368	96.63	N24	28,723,067	26,468,090	96.37
N10	11,097,629	9,005,236	96.63	N25	13,133,847	10,580,378	96.45
N11	37,606,988	34,953,931	96.57	N26	49,565,198	45,685,957	97.04
N12	13,898,951	11,215,972	96.34	N27	30,560,144	25,173,073	96.43
N13	15,074,201	12,984,323	96.44	N28	51,165,141	41,041,717	96.68
N14	32,897,105	30,731,330	94.46	N29	14,024,835	11,2961,32	97.12
N15	33,068,852	30,751,981	96.29	N30	34,107,958	27,737,678	96.82
Mean	23,595,455	20,841,779	96.33	Mean	27,957,829	24,095,638	96.55

Table S2. Differentially expressed genes identified in the *Longissimus thoracis* muscle of Nelore cattle divergent to ribeye muscle area (REA) phenotype (HREA = highest REA and LREA = lowest REA)

Gene	Locus	HREA	LREA	Fold change (log2)*	q value
LOC104973789	13:42894772-42926195	0.14	3.03	4.39	0.0040
BARX2	29:33240583-33256128	0.45	3.37	2.90	0.0040
LOC104970902	1:74137173-74154125	0.72	2.99	2.05	0.0040
FABP7	9:28834076-28837863	0.99	3.70	1.91	0.0040
SLC16A7	5:53987488-54214813	7.22	26.30	1.86	0.0040
MYH10	19:28619740-28801211	2.95	9.91	1.75	0.0040
ESRRG	16:20592131-21286840	2.03	6.73	1.73	0.0040
TNC	8:105965489-106067896	0.67	2.10	1.65	0.0040
MYH9	5:75094865-75179909	12.41	37.17	1.58	0.0040
ANKRD2	26:18626182-18636291	86.73	254.70	1.55	0.0040
ACSS1	13:42963396-43076852	8.29	23.13	1.48	0.0040
GEM	14:72385635-72399661	2.10	5.65	1.43	0.0040
BDH1	1:72571510-72608859	6.13	16.28	1.41	0.0040
SCN3B	15:34768897-34791667	2.86	7.53	1.40	0.0040
RCAN1	1:242245-362910	66.89	170.98	1.35	0.0040
LOC104969184	7:50893535-50936855	1.14	2.90	1.34	0.0070
LOC104972129	4:68199114-68201255	1.03	2.53	1.30	0.0268
LOC520016	12:74293958-74375569	0.97	2.37	1.29	0.0395
APLN	X:13555989-13565261	2.14	5.17	1.27	0.0040
CYP26B1	11:12371179-12387041	1.06	2.39	1.17	0.0040
ESRRB	10:88635687-88821658	1.23	2.75	1.16	0.0040
PRR32	X:10981500-10983418	4.18	9.23	1.14	0.0040
CREB5	4:67765963-68122263	2.33	5.08	1.12	0.0040
ENAH	16:29233872-29442785	9.77	21.31	1.12	0.0040
SOCS3	19:54458148-54461241	1.58	3.45	1.12	0.0040
PRUNE2	8:52958250-53248274	4.27	9.24	1.11	0.0040
FAM101A	17:53715007-53718387	2.89	6.15	1.09	0.0040
ARHGAP36	X:14981844-15013247	0.86	1.84	1.09	0.0434
ACOT11	3:92272943-92309036	1.52	3.19	1.07	0.0040
LOC104976310	29:50168856-50171743	1.11	2.31	1.06	0.0120
YBX2	19:27622772-27628780	1.56	3.25	1.06	0.0040
COL18A1	1:146989202-147041015	2.62	5.42	1.05	0.0040
LOXL1	21:35058002-35081938	2.48	5.13	1.05	0.0040
C15H11orf52	15:22575870-22582949	3.07	6.32	1.04	0.0040
POSTN	12:24241681-24276734	2.06	4.21	1.03	0.0456
RGS16	16:65128460-65134057	1.88	3.81	1.02	0.0040
LOC101907824	2:16965046-17449870	5.02	10.09	1.01	0.0040
MRPS6	1:669918-733729	28.68	57.06	0.99	0.0040
LOC782233	3:28623286-28623761	30.62	60.75	0.99	0.0095
SMPDL3A	9:28811575-28829543	5.28	10.47	0.99	0.0040
LOC615989	3:92245938-92260697	1.40	2.77	0.99	0.0402
CPE	17:541678-697925	1.07	2.10	0.97	0.0040

<i>RFX2</i>	7:19579227-19640152	2.02	3.95	0.97	0.0040
<i>LOC613595</i>	1:84267948-84528019	3.66	7.15	0.96	0.0449
<i>MYOM3</i>	2:129386329-129438822	62.17	120.58	0.96	0.0040
<i>NES</i>	3:14208427-14215602	17.15	33.12	0.95	0.0040
<i>GALNT14</i>	11:68745031-68963586	1.00	1.94	0.95	0.0249
<i>RSPO3</i>	9:24364562-24365792	9.04	17.22	0.93	0.0040
<i>MYL6B</i>	5:57489467-57493545	1224.42	2319.87	0.92	0.0040
<i>XIRP1</i>	22:12549675-12558908	45.71	86.30	0.92	0.0168
<i>RRAD</i>	18:34746054-34749250	122.39	229.21	0.91	0.0040
<i>RGS2</i>	4:58723599-58724929	5.75	10.77	0.91	0.0040
<i>METTL21C</i>	12:82906942-82918374	6.65	12.44	0.90	0.0268
<i>SLC1A4</i>	11:63337654-63367593	2.54	4.73	0.90	0.0040
<i>TNFRSF12A</i>	25:2435179-2441777	34.99	65.16	0.90	0.0040
<i>COL8A1</i>	1:43541935-43717620	1.86	3.40	0.87	0.0040
<i>LAYN</i>	15:22174757-22198295	0.99	1.80	0.87	0.0307
<i>LMCD1</i>	22:17961209-18018234	92.96	169.26	0.86	0.0040
<i>ANGPTL2</i>	11:97780156-97975506	10.99	19.74	0.84	0.0040
<i>ACTN1</i>	10:81023509-81121726	3.18	5.71	0.84	0.0040
<i>CSRP3</i>	29:25994181-26014860	1696.94	3038.41	0.84	0.0402
<i>HCLS1</i>	1:66720532-66754184	1.49	2.68	0.84	0.0040
<i>TRIM55</i>	14:32151429-32211766	12.91	23.03	0.84	0.0040
<i>EHD4</i>	10:37412299-37496211	1.81	3.23	0.83	0.0095
<i>ST8SIA5</i>	24:46884520-46951725	2.96	5.26	0.83	0.0040
<i>MYH7B</i>	13:64888552-64981581	11.18	19.85	0.83	0.0449
<i>SCD5</i>	6:99233278-99410753	1.76	3.13	0.83	0.0120
<i>CDC42EP1</i>	5:109925708-109933543	13.62	24.17	0.83	0.0040
<i>LOC101907485</i>	18:23736539-23760919	6.15	10.89	0.82	0.0040
<i>MYC</i>	14:13769241-13775688	3.95	6.94	0.81	0.0040
<i>HSPB7</i>	2:136635160-136638958	314.90	553.24	0.81	0.0040
<i>TUBA1A</i>	5:30798091-30825706	6.11	10.61	0.80	0.0040
<i>STAC</i>	22:9920968-10273713	0.97	1.69	0.79	0.0192
<i>GCNT4</i>	10:6429194-6461833	1.88	3.25	0.79	0.0456
<i>SERPINH1</i>	15:55514869-55525178	18.69	32.16	0.78	0.0040
<i>MT2A</i>	18:24125365-24126333	134.08	229.78	0.78	0.0040
<i>TSHZ1</i>	24:3673496-3752935	2.37	4.06	0.78	0.0402
<i>UCK2</i>	3:3045780-3119590	4.80	8.21	0.77	0.0382
<i>RCAN2</i>	23:19435240-19714193	4.48	7.65	0.77	0.0040
<i>LOC786565</i>	7:54283885-54352001	33.28	56.63	0.77	0.0040
<i>ADAMTSL4</i>	3:20187037-20195550	8.59	14.61	0.77	0.0040
<i>LHX6</i>	11:93066822-93090379	1.17	1.98	0.75	0.0249
<i>VASH1</i>	10:88994234-89048248	2.69	4.51	0.75	0.0040
<i>S100B</i>	1:148009628-148016981	33.30	55.92	0.75	0.0040
<i>RAI14</i>	20:39351347-39515309	7.02	11.70	0.74	0.0040
<i>COL5A1</i>	11:106987521-107031394	2.66	4.44	0.74	0.0268
<i>TPPP</i>	20:71561963-71577469	2.14	3.57	0.74	0.0070
<i>TRIM47</i>	19:56373235-56377735	1.84	3.04	0.72	0.0147

<i>MXRA5</i>	X:138875004-138904034	1.30	2.15	0.72	0.0040
<i>CNN1</i>	7:17106393-17114224	4.52	7.45	0.72	0.0070
<i>PRRX1</i>	16:39014879-39101813	7.10	11.68	0.72	0.0040
<i>FLNA</i>	X:40310749-40335761	6.97	11.45	0.72	0.0040
<i>AVPI1</i>	26:18711828-18725849	18.76	30.77	0.71	0.0040
<i>RALGAPA2</i>	13:40247867-40535389	1.29	2.11	0.71	0.0040
<i>PALMD</i>	3:43688096-43748131	50.01	81.88	0.71	0.0040
<i>ABLIM1</i>	26:35193851-35406860	23.71	38.63	0.70	0.0040
<i>LDHB</i>	5:88962684-88981219	220.50	357.95	0.70	0.0040
<i>COL4A1</i>	12:88876124-89009422	45.46	73.75	0.70	0.0070
<i>CASQ2</i>	3:27658861-27729417	38.57	62.38	0.69	0.0040
<i>FADS3</i>	29:41087508-41102524	4.82	7.78	0.69	0.0040
<i>C1QB</i>	2:130763198-130776066	6.02	9.67	0.68	0.0168
<i>KLHL40</i>	22:15444469-15451285	86.82	139.29	0.68	0.0040
<i>PROX1</i>	16:71408517-71468630	2.84	4.54	0.68	0.0040
<i>SIK1</i>	1:146208318-146230363	8.13	12.98	0.68	0.0040
<i>RRBP1</i>	13:38289295-38331995	2.36	3.75	0.67	0.0040
<i>GAS7</i>	19:29666087-29876800	4.32	6.84	0.66	0.0325
<i>TPM4</i>	7:7923140-7948408	13.06	20.64	0.66	0.0040
<i>BGN</i>	X:39639856-39653691	10.00	15.78	0.66	0.0040
<i>DMPK</i>	18:53744226-53769643	39.74	62.60	0.66	0.0040
<i>GPRC5B</i>	25:17587460-17608885	6.47	10.12	0.65	0.0040
<i>C15H11orf96</i>	15:74915917-74917297	7.41	11.55	0.64	0.0095
<i>LACC1</i>	12:14134523-14147904	3.84	5.96	0.63	0.0382
<i>AEBP1</i>	4:77881912-77891706	2.72	4.21	0.63	0.0095
<i>PPM1K</i>	6:37876469-37899686	9.01	13.94	0.63	0.0095
<i>IMPA2</i>	24:43189672-43209175	7.08	10.92	0.63	0.0120
<i>LOXL2</i>	8:71278581-71390893	3.85	5.93	0.62	0.0095
<i>CKAP4</i>	5:69971126-69979910	2.22	3.43	0.62	0.0040
<i>SLC16A1</i>	3:30533844-30563304	24.26	37.35	0.62	0.0040
<i>MTURN</i>	4:66610953-66643765	5.71	8.76	0.62	0.0040
<i>AGPAT9</i>	6:100070391-100141603	4.36	6.69	0.62	0.0449
<i>LAMA4</i>	9:38644190-38744427	2.57	3.93	0.61	0.0208
<i>LOC533821</i>	27:23463143-23510403	1.32	2.01	0.61	0.0395
<i>NDRG4</i>	18:26340278-26383765	3.20	4.87	0.60	0.0268
<i>MYH6</i>	10:21350290-21377103	13.36	20.25	0.60	0.0040
<i>MCM10</i>	13:28132518-28170527	3.57	5.40	0.60	0.0168
<i>ATOX8</i>	11:48954368-48996567	4.24	6.41	0.60	0.0382
<i>SEPT9</i>	19:55116876-55296503	5.99	9.06	0.60	0.0070
<i>DOT1L</i>	7:22629627-22682425	7.22	10.88	0.59	0.0192
<i>DNAJA4</i>	21:31178543-31194565	88.38	133.09	0.59	0.0040
<i>EFHD2</i>	16:53306774-53324889	2.23	3.36	0.59	0.0356
<i>SACS</i>	12:34836469-34918947	2.46	3.69	0.59	0.0192
<i>CSRNP1</i>	22:12519353-12532806	6.01	9.00	0.58	0.0208
<i>COL4A2</i>	12:89050616-89166035	27.19	40.72	0.58	0.0070
<i>TUBB4B</i>	11:106854338-106940904	5.39	8.05	0.58	0.0040

<i>CD248</i>	29:45015998-45018647	2.62	3.90	0.57	0.0382
<i>VIM</i>	13:31945011-31952941	75.77	112.29	0.57	0.0095
<i>C1QA</i>	2:130792853-130795743	11.09	16.41	0.57	0.0467
<i>SH2D3C</i>	11:98435178-98467957	3.48	5.14	0.56	0.0344
<i>ZNF385A</i>	5:25802666-25823618	6.34	9.35	0.56	0.0382
<i>COL6A3</i>	3:117408827-117499850	17.29	25.49	0.56	0.0168
<i>LOC782954</i>	22:47474704-47477289	8.09	11.91	0.56	0.0496
<i>DIXDC1</i>	15:22583276-22662895	9.18	13.51	0.56	0.0040
<i>SRXN1</i>	13:60944182-60951049	7.14	10.48	0.55	0.0467
<i>CD44</i>	15:66454329-66541792	9.80	14.38	0.55	0.0208
<i>ABRA</i>	14:59769918-59989290	11.93	17.46	0.55	0.0120
<i>LOC100335819</i>	18:57081668-57082884	28.90	42.27	0.55	0.0070
<i>FHOD3</i>	24:20649325-21172169	9.53	13.93	0.55	0.0120
<i>HSPA2</i>	10:76992201-76994112	5.95	8.70	0.55	0.0372
<i>PRKCDBP</i>	15:47329715-47331336	7.41	10.83	0.55	0.0344
<i>KLHL34</i>	X:128884891-129177766	23.96	34.95	0.54	0.0168
<i>ACTA2</i>	26:10662359-10748026	42.26	61.62	0.54	0.0070
<i>SPTBN1</i>	11:37028023-37241384	31.19	45.43	0.54	0.0168
<i>FN1</i>	2:103881401-103950586	12.70	18.48	0.54	0.0168
<i>SPECC1</i>	19:34121513-34328508	15.80	22.97	0.54	0.0040
<i>TSPAN9</i>	5:107044037-107237024	5.31	7.70	0.54	0.0095
<i>PLAU</i>	28:29962475-29971037	5.35	7.76	0.54	0.0496
<i>IGF2</i>	29:50037925-50065231	26.39	38.26	0.54	0.0070
<i>ACTN4</i>	18:48668481-48766658	10.24	14.83	0.53	0.0070
<i>DPYSL3</i>	7:60586187-60703487	18.67	27.03	0.53	0.0040
<i>PLXNA2</i>	16:77046044-77250671	2.25	3.25	0.53	0.0325
<i>FAM129B</i>	11:98231337-98284464	2.81	4.05	0.53	0.0231
<i>NT5C2</i>	26:23983197-24080607	23.84	34.33	0.53	0.0120
<i>CD97</i>	7:12451148-12469080	8.67	12.47	0.52	0.0307
<i>KLF6</i>	13:44945037-44952166	15.96	22.89	0.52	0.0147
<i>TPPP3</i>	18:35121238-35124832	63.87	91.49	0.52	0.0395
<i>GALNT16</i>	10:81395753-81494776	2.63	3.76	0.52	0.0402
<i>ELN</i>	25:33788120-33820750	14.97	21.40	0.52	0.0120
<i>ANXA2</i>	10:49840792-49904544	63.87	91.11	0.51	0.0307
<i>SESN3</i>	15:15503054-15530868	19.54	27.61	0.50	0.0095
<i>ADRB2</i>	7:62229085-62230992	11.77	16.55	0.49	0.0483
<i>FNIP2</i>	17:40936425-41205627	2.50	3.51	0.49	0.0372
<i>CLIC4</i>	2:128710444-128783921	15.88	22.26	0.49	0.0325
<i>CFL1</i>	29:44632743-44642280	35.82	49.58	0.47	0.0356
<i>MYOG</i>	16:596915-599757	25.84	35.76	0.47	0.0307
<i>VAT1</i>	19:43687402-43694870	11.19	15.44	0.46	0.0356
<i>NID2</i>	10:44894655-44986659	6.75	9.29	0.46	0.0356
<i>ITGA5</i>	5:25778011-25799053	5.20	7.14	0.46	0.0456
<i>COL6A2</i>	1:147542733-147572939	27.15	37.05	0.45	0.0402
<i>HSPA5</i>	11:96115224-96119266	28.65	39.06	0.45	0.0456
<i>FAM134B</i>	20:56709602-56758648	41.90	57.05	0.45	0.0402

<i>ACTR1B</i>	11:2928633-2943363	16.32	22.20	0.44	0.0325
<i>ETS1</i>	29:32355401-32494471	10.31	14.01	0.44	0.0434
<i>MURC</i>	8:91916164-91928341	75.61	56.92	-0.41	0.0434
<i>SEL1L3</i>	6:46790207-46881531	39.28	29.43	-0.42	0.0467
<i>TMBIM4</i>	5:47877714-47894126	227.58	169.57	-0.42	0.0420
<i>SHISA4</i>	16:49618804-49621012	240.76	179.37	-0.42	0.0402
<i>UBE2O</i>	19:55928512-55977810	6.62	4.88	-0.44	0.0456
<i>NEURL1</i>	26:24330860-24404657	56.10	41.31	-0.44	0.0356
<i>GADL1</i>	22:5258456-5452375	47.26	34.48	-0.45	0.0483
<i>LRRC30</i>	24:40563770-40572754	34.61	25.22	-0.46	0.0372
<i>NEO1</i>	10:19722879-19959963	12.76	9.28	-0.46	0.0496
<i>CARNS1</i>	29:45955672-45964991	7.46	5.43	-0.46	0.0356
<i>ATL2</i>	11:20689675-20759067	148.21	107.66	-0.46	0.0382
<i>DUSP26</i>	27:28963363-28970161	167.88	121.95	-0.46	0.0456
<i>ODC1</i>	11:87175388-87182661	90.05	65.40	-0.46	0.0208
<i>EYA1</i>	14:36896159-37268911	10.03	7.28	-0.46	0.0290
<i>UBE2G1</i>	19:25407194-25496726	420.47	305.11	-0.46	0.0434
<i>DCUN1D2</i>	12:90679652-90698678	92.70	67.18	-0.46	0.0268
<i>UBE2D4</i>	22:322994-348275	176.21	127.19	-0.47	0.0395
<i>PPAPDC3</i>	11:101502264-101517589	104.63	75.43	-0.47	0.0249
<i>TNK2</i>	1:71207393-71247202	12.11	8.65	-0.49	0.0249
<i>TMEM233</i>	17:58031788-58079512	69.57	49.55	-0.49	0.0208
<i>SLC12A2</i>	7:26970750-27064924	6.37	4.52	-0.49	0.0231
<i>YPEL3</i>	25:26451071-26454436	111.55	79.15	-0.49	0.0095
<i>CFAP97</i>	27:14553980-14579529	11.16	7.91	-0.50	0.0402
<i>UXS1</i>	11:45583096-45640370	49.36	34.94	-0.50	0.0095
<i>SLC3A2</i>	29:41855870-41896761	37.98	26.87	-0.50	0.0208
<i>SOX6</i>	15:36371591-37088565	12.56	8.87	-0.50	0.0070
<i>DAPK2</i>	10:46004744-46147029	15.02	10.59	-0.50	0.0249
<i>EYA4</i>	9:72371316-72738259	104.74	73.81	-0.50	0.0095
<i>LMOD3</i>	22:32534125-32550297	190.88	134.43	-0.51	0.0249
<i>PPP3R1</i>	11:66582635-66642276	190.61	134.10	-0.51	0.0040
<i>CPED1</i>	4:86249248-86574486	16.66	11.68	-0.51	0.0208
<i>H1FO</i>	5:110122672-110124893	48.00	33.40	-0.52	0.0095
<i>PDE4C</i>	7:4927692-4939931	13.83	9.61	-0.53	0.0070
<i>FAM69A</i>	3:50658411-50787192	15.39	10.68	-0.53	0.0307
<i>PDLIM7</i>	7:40333460-40353588	485.38	336.62	-0.53	0.0395
<i>KLHL23</i>	2:26617808-26637610	32.78	22.67	-0.53	0.0070
<i>MLF1</i>	1:109762074-109793811	129.29	89.35	-0.53	0.0040
<i>CCDC88C</i>	21:56629745-56645863	42.37	29.25	-0.53	0.0040
<i>SH3BP5</i>	1:154023223-154103515	70.53	48.40	-0.54	0.0040
<i>SCUBE2</i>	15:44090649-44162339	8.34	5.70	-0.55	0.0070
<i>MAFA</i>	14:2424939-2428006	9.83	6.72	-0.55	0.0070
<i>RABGGTA</i>	10:20726119-20733208	25.08	17.13	-0.55	0.0040
<i>GANC</i>	10:37754342-37828667	8.53	5.82	-0.55	0.0192
<i>ASB16</i>	19:44614051-44621590	59.69	40.67	-0.55	0.0070

<i>CRTC1</i>	7:4362476-4447554	4.59	3.13	-0.56	0.0095
<i>KCNJ12</i>	19:35955279-35993796	41.08	27.95	-0.56	0.0070
<i>IPP</i>	3:100935508-100982144	12.89	8.76	-0.56	0.0208
<i>GADD45A</i>	3:77972151-77975240	51.46	34.94	-0.56	0.0070
<i>LRRC14B</i>	20:71983691-71991212	7.21	4.89	-0.56	0.0147
<i>ASB12</i>	X:101172768-101492861	329.16	223.12	-0.56	0.0040
<i>LOC104971550</i>	3:28736437-28745657	111.23	75.16	-0.57	0.0040
<i>PRKAG3</i>	2:107507840-107517485	34.73	23.37	-0.57	0.0040
<i>LOC618664</i>	15:52255741-52259548	22.71	15.28	-0.57	0.0192
<i>OAT</i>	26:44378155-44397476	64.53	43.24	-0.58	0.0040
<i>AMPD1</i>	3:28756907-28768515	187.51	125.59	-0.58	0.0040
<i>KLHL33</i>	10:26728542-26738309	38.33	25.56	-0.58	0.0040
<i>MLXIP</i>	17:55472096-55528500	126.23	83.32	-0.60	0.0040
<i>DHDH</i>	18:55964767-55975733	16.80	11.09	-0.60	0.0070
<i>DECR1</i>	14:76022897-76066405	12.90	8.49	-0.60	0.0356
<i>PLCL1</i>	2:86717861-87088818	8.46	5.56	-0.60	0.0168
<i>GPD2</i>	2:39762794-39913649	16.63	10.92	-0.61	0.0040
<i>DENND2A</i>	4:104715492-104805618	4.37	2.85	-0.61	0.0040
<i>ACSM5</i>	25:18206792-18235137	6.74	4.37	-0.62	0.0095
<i>RGS14</i>	7:40214151-40229119	4.98	3.23	-0.63	0.0040
<i>LOC104976285</i>	29:44770857-44771525	71.36	45.87	-0.64	0.0382
<i>WFIKKN2</i>	19:36536564-36543238	6.34	4.06	-0.64	0.0040
<i>BCL2L11</i>	11:1152978-1203900	3.35	2.13	-0.65	0.0372
<i>KBTD13</i>	10:11967407-11970038	2.94	1.86	-0.66	0.0208
<i>SHISA2</i>	12:33578133-33583716	45.95	28.81	-0.67	0.0040
<i>WWP1</i>	14:78599362-78699730	189.42	117.89	-0.68	0.0040
<i>SNTB1</i>	14:84251467-84504187	55.76	34.70	-0.68	0.0040
<i>IYD</i>	9:88612922-88624698	13.55	8.36	-0.70	0.0402
<i>GSTM1</i>	3:33805750-33816442	94.96	58.57	-0.70	0.0040
<i>NPR3</i>	20:40965737-41042094	3.51	2.15	-0.70	0.0344
<i>LOC615514</i>	3:33780025-33800900	61.04	37.42	-0.71	0.0040
<i>AGTPBP1</i>	8:80235081-80433168	18.91	11.47	-0.72	0.0040
<i>NPNT</i>	6:20433658-20530094	14.58	8.83	-0.72	0.0040
<i>KLF10</i>	14:63954120-63957276	133.52	80.20	-0.74	0.0040
<i>PADI2</i>	2:136049637-136106203	20.77	12.44	-0.74	0.0040
<i>PFKFB3</i>	13:17380739-17458018	173.42	103.33	-0.75	0.0040
<i>RAB3A</i>	7:4944324-4950014	4.86	2.87	-0.76	0.0231
<i>RPS6KA1</i>	2:127120755-127202561	7.26	4.29	-0.76	0.0040
<i>LOC100300896</i>	11:44706146-44717569	10.09	5.94	-0.76	0.0147
<i>LOC104972782</i>	18:57057264-57058171	125.03	73.62	-0.76	0.0040
<i>FRZB</i>	2:13723775-13761804	7.90	4.63	-0.77	0.0040
<i>MSTN</i>	2:6213565-6220196	11.50	6.71	-0.78	0.0040
<i>IDO1</i>	27:34686576-34699492	8.69	5.05	-0.78	0.0456
<i>METTL7A</i>	5:29161689-29172757	39.83	22.72	-0.81	0.0040
<i>LIPT2</i>	15:54616101-54618044	4.18	2.32	-0.85	0.0290
<i>MYLK3</i>	18:15086316-15143371	3.56	1.96	-0.86	0.0095

<i>LOC781439</i>	5:29081457-29082134	9.08	4.99	-0.86	0.0467
<i>HES1</i>	1:73974239-73976720	31.42	16.98	-0.89	0.0040
<i>TSPAN33</i>	4:93800785-93901224	1.93	1.03	-0.90	0.0382
<i>LOC101904082</i>	8:99160768-99165239	9.12	4.75	-0.94	0.0040
<i>ATP2B2</i>	22:54995630-55122063	5.42	2.79	-0.96	0.0040
<i>KY</i>	1:135841442-135852560	7.39	3.74	-0.98	0.0040
<i>LOC104975597</i>	22:55141082-55153028	1.50	0.75	-1.00	0.0467
<i>KIAA1211</i>	6:73120829-73398915	3.70	1.75	-1.08	0.0040
<i>LOC100196901</i>	19:55822701-55827196	32.90	15.56	-1.08	0.0040
<i>PLEK</i>	11:66762700-66789951	5.75	2.63	-1.13	0.0040
<i>SH3RF2</i>	7:59052290-59286495	2.46	1.12	-1.13	0.0040
<i>LOC782631</i>	13:59957495-59958191	20.58	9.32	-1.14	0.0040
<i>SPOCK2</i>	28:28304693-28329944	11.26	4.45	-1.34	0.0040
<i>CATHL1</i>	22:52148147-52149590	6.57	2.46	-1.42	0.0120
<i>MAOB</i>	X:105235854-105359155	3.32	0.99	-1.75	0.0040
<i>OXT</i>	13:52575289-52576188	69.20	17.11	-2.02	0.0040

Symbol of the differentially expressed gene (Gene), location of the gene in the *Bos taurus* genome, FPKM values obtained for HEREA and LREA, relative expression (log2 fold change) and q-value.

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

Table S3: Biological Process (GO terms) and pathways (bta) containing genes differentially expressed in ribeye muscle of Nelore bulls selected to be divergent for ribeye muscle area

Terms	Genes	p-value
GO:0045214~sarcomere organization	<i>MYLK3, FHOD3, ACTN1, MYH6, CASQ2</i>	2,5E-4
GO:0048741~skeletal muscle fiber development	<i>ACTA2, KLHL40, ACTN1, MYOG</i>	0.003
GO:0030198~extracellular matrix organization	<i>COL18A1, BGN, SPOCK2, NPNT, ADAMTSL4, ELN</i>	0.005
GO:0055003~cardiac myofibril assembly	<i>FHOD3, CSRP3, MYH10</i>	0.006
GO:0035987~endodermal cell differentiation	<i>COL4A2, ITGA5, COL8A1, FN1</i>	0.007
GO:0051764~actin crosslink formation	<i>ACTN1, DPYSL3, FLNA</i>	0.010
GO:0030220~platelet formation	<i>ACTN1, MYH9, ZNF385A</i>	0.020
GO:0045765~regulation of angiogenesis	<i>TNFRSF12A, ETS1, VASH1</i>	0.032
GO:0009409~response to cold	<i>ADRB2, HSPA2, ACOT11</i>	0.036
GO:0097150~neuronal stem cell population maintenance	<i>HES1, PRRX1, PROX1</i>	0.039
GO:0033631~cell-cell adhesion mediated by integrin	<i>ITGA5, NPNT</i>	0.046
GO:0016567~protein ubiquitination	<i>IPP, KBTBD13, SOCS3, ASB12, KLHL23, KLHL34, KLHL33</i>	0.050
GO:0007155~cell adhesion	<i>HES1, COL18A1, LAMA4, CD44, TNC, FN1, MYH10</i>	0.058
GO:0070884~regulation of calcineurin-NFAT signaling cascade	<i>RCAN1, RCAN2</i>	0.062
GO:0021707~cerebellar granule cell differentiation	<i>ATP2B2, PROX1</i>	0.062
GO:0051271~negative regulation of cellular component movement	<i>ACTN4, ACTN1</i>	0.062
GO:0016576~histone dephosphorylation	<i>EYA4, EYA1</i>	0.062
GO:0001525~angiogenesis	<i>COL18A1, CLIC4, MYH9, VASH1, FN1, ANXA2</i>	0.062
GO:0051017~actin filament bundle assembly	<i>ACTN4, ACTN1, DPYSL3</i>	0.063
GO:0007517~muscle organ development	<i>MURC, MYOG, CSRP3</i>	0.063
GO:0050873~brown fat cell differentiation	<i>ADRB2, LAMA4, RGS2</i>	0.063
GO:0055114~oxidation-reduction process	<i>GPD2, LDHB, MAOB, CYP26B1, FADS3, SCD5, LOXL2, BDH1, LOXL1, DHDH</i>	0.066
GO:0060048~cardiac muscle contraction	<i>MYH6, CSRP3, CASQ2</i>	0.068
GO:0030154~cell differentiation	<i>MURC, EYA4, EYA1, PDLIM7, HSPA2, ETS1, TNK2, ATOH8</i>	0.072

GO:0070534~protein K63-linked ubiquitination	<i>UBE2D4, UBE2O, UBE2G1</i>	0.072
GO:0010811~positive regulation of cell-substrate adhesion	<i>SPOCK2, NPNT, COL8A1</i>	0.072
GO:0001768~establishment of T cell polarity	<i>CYP26B1, MYH9</i>	0.077
GO:0007044~cell-substrate junction assembly	<i>ITGA5, FN1</i>	0.077
GO:0019800~peptide cross-linking via chondroitin 4-sulfate glycosaminoglycan	<i>BGN, SPOCK2</i>	0.077
GO:0030240~skeletal muscle thin filament assembly	<i>ACTA2, PROX1</i>	0.077
GO:0055009~atrial cardiac muscle tissue morphogenesis	<i>MYH6, PROX1</i>	0.077
GO:0030199~collagen fibril organization	<i>LOXL2, SERPINH1, ANXA2</i>	0.077
GO:0045747~positive regulation of Notch signaling pathway	<i>HES1, EYA1, TSPAN33</i>	0.082
GO:0010595~positive regulation of endothelial cell migration	<i>ETS1, ATOH8, PROX1</i>	0.082
GO:0008360~regulation of cell shape	<i>CDC42EP1, PALMD, MYH9, FN1, MYH10</i>	0.086
GO:0070527~platelet aggregation	<i>PLEK, MYH9, FLNA</i>	0.086
GO:0051216~cartilage development	<i>CD44, PRRX1, SOX6</i>	0.091
GO:0033173~calcineurin-NFAT signaling cascade	<i>PPP3R1, RCAN1</i>	0.091
GO:0090084~negative regulation of inclusion body assembly	<i>SACS, DNAJA4</i>	0.091
GO:0043462~regulation of ATPase activity	<i>ACTN1, MYH6</i>	0.091
GO:0060070~canonical Wnt signaling pathway	<i>DIXDC1, MYH6, FRZB, MYC</i>	0.094
bta04512~ECM-receptor interaction	<i>COL4A2, LAMA4, COL4A1, CD44, ITGA5, TNC, COL6A3, COL6A2, COL5A1, FN1</i>	5,49E-06
bta04510~Focal adhesion	<i>COL4A2, COL4A1, ACTN4, MYLK3, TNC, ACTN1, FLNA, COL5A1, LAMA4, ITGA5, COL6A3, COL6A2, FN1</i>	6,24E-05
bta04974~Protein digestion and absorption	<i>COL18A1, COL4A2, COL4A1, COL6A3, ELN, COL6A2, SLC3A2, COL5A1</i>	2,52E-04
bta04921~Oxytocin signaling pathway	<i>PRKAG3, RGS2, MYL6B, OXT, MYLK3, PPP3R1, RCAN1, KCNJ12</i>	2,07E-03
bta05146~Amoebiasis	<i>COL4A2, LAMA4, COL4A1, ACTN4, ACTN1, COL5A1, FN1</i>	6,66E-03
bta04151~PI3K-Akt signaling pathway	<i>COL4A2, LAMA4, COL4A1, ITGA5, TNC, COL6A3, COL6A2, CREB5, MYC, BCL2L11, COL5A1, FN1</i>	1,49E-02
bta04530~Tight junction	<i>ACTN4, HCLS1, ACTN1, MYH6, MYH9, MYH7B, MYH10</i>	1,69E-02
bta05205~Proteoglycans in cancer	<i>CD44, ITGA5, HCLS1, IGF2, MYC, FLNA, PLAU, FN1</i>	3,29E-02
bta00330~Arginine and proline metabolism	<i>ODC1, MAOB, OAT, CARN1</i>	3,89E-02
bta04022~cGMP-PKG signaling pathway	<i>ATP2B2, ADRB2, RGS2, MYLK3, PPP3R1, CREB5</i>	3,96E-02

bta05222~Small cell lung cancer	<i>COL4A2, LAMA4, COL4A1, MYC, FN1</i>	4,31E-02
bta00480~Glutathione metabolism	<i>GSTM1, LOC615514, ODC1</i>	4,72E-02
bta04922~Glucagon signaling pathway	<i>PRKAG3, LDHB, PPP3R1, CREB5, SIK1</i>	5,44E-02
bta00982~Drug metabolism - cytochrome P450	<i>GSTM1, LOC615514, MAOB</i>	5,63E-02
bta00980~Metabolism of xenobiotics by cytochrome P450	<i>GSTM1, LOC615514, DHDH</i>	5,87E-02
bta00512~Mucin type O-Glycan biosynthesis	<i>GCNT4, GALNT16, GALNT14</i>	8,30E-02
bta05020~Prion diseases	<i>C1QA, C1QB, HSPA5</i>	8,30E-02
bta04931~Insulin resistance	<i>PRKAG3, RPS6KA1, SOCS3, MLXIP, CREB5</i>	8,56E-02

CHAPTER 3 - Prediction of hub genes associated with intramuscular fat content in Nelore cattle^b

ABSTRACT - Background: The aim of this study was to use transcriptome RNA-Seq data from *longissimus thoracis* muscle of uncastrated Nelore males to identify hub genes based on co-expression networks obtained from differentially expressed genes (DEGs) associated with intramuscular fat content. **Results:** A total of 30 transcriptomics datasets (RNA-Seq) obtained from *longissimus thoracis* muscle were selected based on the phenotypic value of divergent intramuscular fat content: 15 with the highest intramuscular fat content (HIF) and 15 with the lowest intramuscular fat content (LIF). The transcriptomics datasets were aligned with a reference genome and DEGs were identified. Upon analysis, 65 DEGs were identified, including 21 upregulated and 44 downregulated genes in HIF animals. Gene set enrichment analysis was performed and showed that the gene ontology (GO) terms that were significantly enriched are closely associated with lipid and muscle metabolism, as well as genes related to the immune system. The normalized count data from DEGs was then used for co-expression network construction. The topological properties of the network were analyzed; those genes engaging in the most interactions with other genes were predicted to be hub genes. Gene co-expression network analysis showed three top hub genes (*PDE4D*, *KLHL30*, and *IL1RAP*) found to be the main regulators of other DEGs identified in the analysis which consequently may play a role in cellular and/or systemic lipid biology in Nelore cattle. The hub genes were downregulated in HIF animals. *PDE4D* and *IL1RAP* have known effects on lipid metabolism and the immune system through the regulation of cAMP signaling. *PDE4D* and *IL1RAP* downregulation may contribute to increased levels of intracellular cAMP and thus may have effects on IF content differences in Nelore cattle, given that cAMP is known to play a role in lipid systems. *KLHL30* may have effects on muscle metabolism and Khl protein families play a role in protein degradation. **Conclusions:** The results reported in this study indicate candidate genes and a better understanding of the molecular mechanisms of intramuscular fat deposition in Nelore cattle. **Keywords:** Bovine, co-expression, genes, lipids, transcriptome

Background

Intramuscular fat (IF) content is an important meat quality trait that represents the amount of fat accumulated between muscle fibers or inside muscle cells; it is the sum of phospholipids and triglycerides. IF plays a key role in meat quality, as it is mainly responsible for the palatability, tenderness, and nutritional value of meat [1, 2, 3]. IF content may be measured via chemical methods, which evaluate total lipids or by marbling scores, which may be assessed after slaughter or by ultrasonography. IF content is a polygenic trait regulated by many genes involved directly or indirectly in adipogenesis and fat

^b This chapter corresponds to manuscript submitted to "BMC Genomics" at November 1, 2018.

metabolism [4]. Therefore, understanding the biological and functional mechanisms that regulate IF content is a compelling question in meat science.

Next generation sequencing applied to tissue transcriptomes (RNA-Seq) is an approach for screening the expression of functional candidate genes and identifying important molecular mechanisms that generate variation in pathways resulting in different tissue phenotypes [5, 6]. This technology has been used to identify differentially expressed genes (DEGs) in muscle tissue relate to: tenderness [7] and fatty acid composition [8] in a commercial Nelore population (uncastrated males); intramuscular fat content (measured by marbling scores) [9]; residual feed intake [10]; iron content [11]; and ribeye area and backfat thickness [12] in Nelore steers from an Embrapa (Brazilian Agricultural Research Corporation) breeding herd (castrated males).

Using RNA-seq analysis, Cesar et al. [9] identified differentially expressed genes and elucidated some of the molecular mechanisms involved in the lipid metabolism of muscle and adipogenesis in Nelore cattle (castrated males) genetically divergent for intramuscular fat (measured by marbling scores). When castrated and uncastrated Nelore or Nelore-cross males are compared, there are differences in intramuscular fat deposition [13, 14]. Therefore, gene expression and the molecular mechanisms regulating differences in intramuscular fat deposition may be different in castrated and uncastrated animals.

The development of IF is influenced by different genes and a complex network of genetic interactions. The biological interpretation of the results obtained from the analysis of RNA-seq, especially for DEGs, remains a challenge. Additional studies are necessary, since many genes and mechanisms that induce differences in uncastrated Nelore male IF content ratios are still unknown, especially hubs genes (biomarkers). Hub genes are highly correlated with a large number of genes, have been shown to play key regulatory roles in gene expression networks [15, 16, 17], and are proposed to play an important role in the overall biology of organisms [18]. Network approaches have been used to identify complex transcriptional regulation, i.e., in the identification of hub genes. Thus, co-expression analysis may facilitate the detection of important biological pathways involved in targeted phenotypes. Consequently, the aim of this study was to use transcriptome RNA-Seq data from *longissimus thoracis*

muscle of uncastrated Nelore males to identify hub genes based on a gene co-expression network constructed from DEGs associated with IF content.

Methods

Sample collection and phenotype

All animals (N = 80), uncastrated Nelore males belonging to the same contemporary group (i.e., remaining together from birth until slaughter), were from the Capivara Farm, which participates in the Nelore Qualitas Breeding Program. They were reared on grazing systems (*Brachiaria* sp. and *Panicum* sp. forage and free access to mineral salt) and finished in confinement for approximately 90 days. The animals were slaughtered at a mean age of 24 months - all on the same day and under the same conditions.

Longissimus thoracis muscle samples were collected from an area between the 12th and 13th ribs of the left half of each carcass two times: 1) at slaughter, stored in 15-mL Falcon tubes containing 5 mL RNA holder (BioAgency, São Paulo, SP, Brazil) at -80 °C until total RNA extraction was performed for RNA sequencing analysis; and 2) 24 hours after slaughter for an analysis of the IF content.

IF content was quantified for all animals, through chemical method for total lipid content, according to the methodology described by Bligh & Dyer [19]. The animals were ranked in accordance with their IF phenotype and the samples derived from the animals with the 15 highest and 15 lowest values for IF content were selected for RNA-seq analysis (Table 1). A Student's t-test was performed to evaluate whether there were significant differences in IF between the selected groups.

Table 1. Descriptive statistics for intramuscular fat content of Nelore cattle

Phenotype*	N	Mean ± standard deviation	Minimum	Maximum	p-value
HIF	15	0.101 ± 0.0095	0.094	0.126	0.05
LIF	15	0.063 ± 0.0049	0.051	0.067	

*The data were transformed by the square root of the percentage of the arcsine function. HIF = Highest intramuscular fat content; LIF = Lowest intramuscular fat content; N = Number of animals.

RNA-seq library construction

Total RNA was isolated from the *longissimus thoracis* samples (an average of 50 mg each) using the RNeasy Lipid Tissue Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol. The following three methods were used for RNA quantification and qualification: RNA purity was determined by evaluating absorbance using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Santa Clara, CA, USA; 2007); RNA concentration was measured using a Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA; 2010), and RNA integrity was assessed using an RNA Nano 6000 Assay Kit with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA; 2009).

Sequencing libraries were prepared using the TruSeq RNA Sample Preparation Kit® (Illumina, San Diego, CA) following the manufacturer's protocol. Libraries were pooled to enable multiplexed sequencing and generated an average of approximately 25 M reads per sample. RNA sequencing (RNA-Seq) was carried out on a HiSeq 2500 System (Illumina®) that generated 100bp paired-end reads.

Data filtering and alignment of reads

Trimmed data (trimmed reads) were obtained by removing low quality reads (adapter sequence and reads containing poly-N) from raw data using Trimmomatic v.0.36 [20]. All downstream analysis was based on the trimmed data with high quality reads. HISAT2 v.2.0.5 [21] was used to align the paired-end trimmed reads to the bovine reference genome (UMD3.1.1 *Bos taurus*) and chromosome Y (Btau 4.6.1), both deposited in National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/>).

Differentially expressed genes (DEGs) analysis

The Cufflinks2 v.2.1.1 suite of tools [22] was used for transcriptome assembly and differential expression analysis. Cufflinks assembles transcriptomes from RNA-Seq data and quantifies their expression in fragments per kilobase of transcript per million reads mapped (FPKM). After assembly, the Cufflinks output per sample was merged together by Cuffmerge. A uniform set of transcripts was obtained for all samples and then Cuffdiff was used to test for

genes that were differentially expressed between the IF groups. False discovery rates (FDR) were controlled using the Benjamini-Hochberg procedure in which we considered a gene to be differentially expressed if it had a *q-value* ≤ 0.05 [23].

Co-expression network analysis and prediction of hub genes

All analyses were performed with plugins or applications from Cytoscape v.3.4, a free software package for visualizing, modeling, and analyzing the integration of biomolecular interaction networks with high-throughput expression data and other molecular states [24]. The co-expression network was constructed from the ExpressionCorrelation plugin [25]. The similarity matrix of count data from DEGs (normalized by FPKM) was computed using a Pearson correlation. A histogram tool was used for the screening criteria at node score cut-offs > 0.75 and < -0.75 . Using the Network Analyzer plugin [26], regulatory relationships among various genes were analyzed by calculating the topological properties of the network such as degree distribution, betweenness centrality, network density, and clustering coefficient [27, 28]. The genes that engaged in the most interactions with other genes, maximal clique centrality (MCC), were identified as hub genes using the cytoHubba plugin [29].

Gene set enrichment analysis

The enrichment and pathway analysis of DEG sets was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID 6.8) [30]. *p-values* were determined using the Fisher Exact test and adjusted using the Benjamini-Hochberg method [23]. The DAVID Pathway was used to map the enriched pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [31].

Results

RNA-seq data alignment

After quality control of the raw reads (approximately 25 million reads), the mean number of reads per sample (paired-end) was approximately 21.9 million. In the library, 88% of the clean reads were uniquely mapped to bovine reference genome UMD3.1.1 *Bos taurus* and chrY of Btau 4.6.1. The mean number of reads mapped in pairs using Hisat2 software was approximately 20.7 million

(96.44%) with 45x sequencing coverage (coverage for all transcripts of all samples), for more details, see Additional File 1. Box plot containing the transformed FPKM values (\log_{10}) for each group and the plot of principal component analysis (PCA), constructed using the cummerbund package, are in Additional file 2. 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 and close to -1 , indicating that the level of sequencing coverage permitted the identification of low-expressed genes. PCA showed the formation of different groups (highest and lowest IF content), indicating differences in the expression of genes between the HIF and LIF groups.

Differentially expressed genes

A total of 65 DEGs ($q \leq 0.05$) were found, including 21 upregulated genes and 44 downregulated genes (Table 2) in the HIF group. The genes were distributed across almost all bovine chromosomes in different proportions, except chromosomes 2, 4, 11, 14, 24, mitochondrial (MT), and Y (Figure 1).

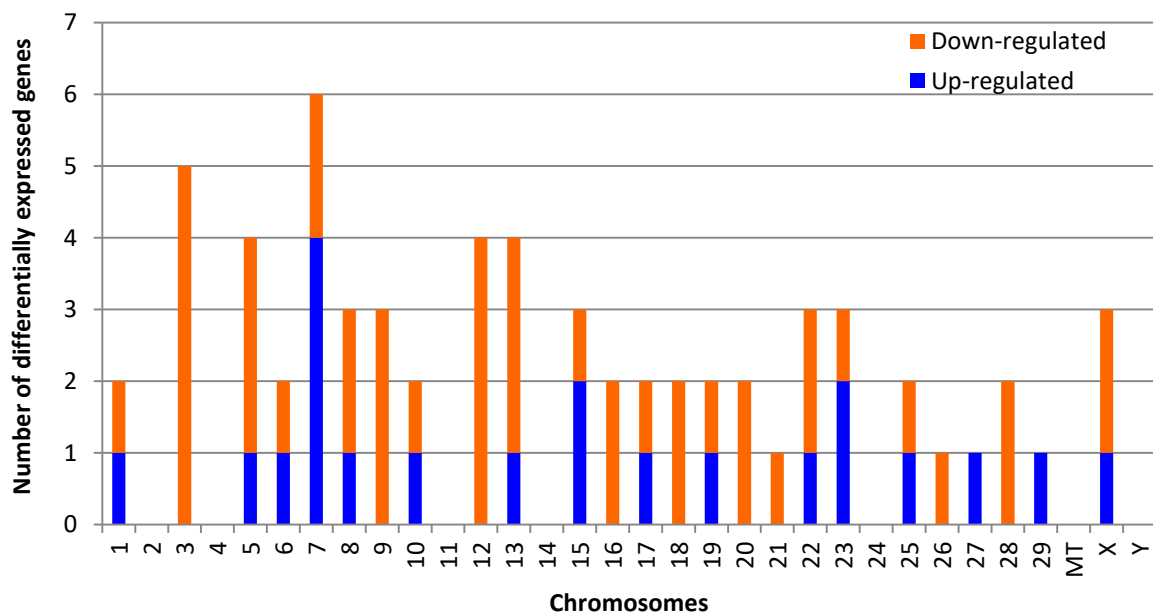


Figure 1. Distribution of the differentially expression genes across the chromosomes. Blue bars represent up-regulated genes and red bars down-regulated genes in the highest ribeye muscle area group.

Important genes that play a role in lipid metabolism were differentially expressed, including CDP-diacylglycerol synthase 1 (*CDS1*) and solute carrier superfamily genes (*SLC22A4* and *SLC2A3*), all of which were upregulated in HIF animals. Genes known to be associated with growth and muscle development were also found, such as collagen type XVIII alpha 1 chain (*COL18A1*) and myosin binding protein H (*MYBPH*). These genes were up- and down-regulated in HIF animals, respectively. In addition to the DEGs related to lipid and muscle metabolism, those related to the innate immune system were also identified: boLA family member (*BOLA*) genes and NFkB inhibitor alpha (*NFKBIA*). These genes were up- and down-regulated in HIF animals, respectively.

Among the DEGs, eight transcription factors were identified: mediator complex subunit 25 (*MED25*); BARX homeobox 2 (*BARX2*); early growth response 1 (*EGR1*); forkhead box O1 (*FOXO1*); CLOCK interacting pacemaker (*CIPC*); nuclear factor and interleukin 3 regulated (NFIL3/ E4BP4); immediate early response 2 (*IER2*), and GlcNAc-1-phosphotransferase subunits alpha/beta (*GNPTAB*). Transcription factors are proteins that play a role in gene expression control systems which are responsible for either positively or negatively influencing the transcription of specific genes, determining whether a particular gene will be turned "on" or "off" in an organism [32].

Table 2. Differentially expressed genes in the *longissimus thoracis* muscle of Nelore cattle divergent for intramuscular fat content phenotypes

Gene	Locus	HIF	LIF	Fold change (log ₂)*	q value
<i>CDS1</i>	6:101195739-101273209	0.56	1.92	1.79	0.02
<i>TNC</i>	8:105965489-106067896	0.71	1.83	1.37	0.01
<i>BARX2</i>	29:33240583-33256128	1.49	3.60	1.28	0.04
<i>EGR1</i>	7:51438971-51441815	12.59	30.35	1.27	0.01
<i>CCL2</i>	19:16232955-16234822	9.64	21.62	1.16	0.01
<i>MXRA5</i>	X:138875004-138904034	1.06	2.20	1.06	0.01
<i>PCDH12</i>	7:54685624-54703271	1.10	2.18	0.99	0.01
<i>SLC22A4</i>	7:23382631-23425462	3.93	7.81	0.99	0.04
<i>SLC2A3</i>	5:101896620-101909565	4.74	9.39	0.99	0.01
<i>APLN</i>	15:81734844-81738412	1.37	2.68	0.96	0.01
<i>FSCN1</i>	25:39292718-39302192	4.08	7.02	0.78	0.01
<i>COL18A1</i>	1:146989202-147041015	2.88	4.90	0.77	0.01
<i>CLDN5</i>	17:74749249-74750524	14.81	24.51	0.73	0.01
<i>CD93</i>	13:42244261-42247908	6.31	10.10	0.68	0.01

<i>MCAM</i>	15:30414471-30422213	5.15	8.12	0.66	0.04
<i>DLL4</i>	10:36577284-36586360	2.68	4.21	0.65	0.04
<i>BOLA</i>	23:28502516-28506343	41.94	65.42	0.64	0.01
<i>IER2</i>	7:13543658-13548059	4.94	7.62	0.63	0.05
<i>IER3</i>	23:28088454-28089726	10.11	15.49	0.62	0.04
<i>PLXND1</i>	22:56776197-56825067	2.38	3.58	0.59	0.01
<i>PPP1R3B</i>	27:24314781-24323720	13.34	19.54	0.55	0.04
<i>NFKBIA</i>	21:46065545-46068942	32.35	21.85	-0.57	0.04
<i>CRISPLD2</i>	18:10985131-11050908	13.29	8.90	-0.58	0.04
<i>ANKRD52</i>	5:57399491-57415909	9.99	6.62	-0.59	0.01
<i>MED25</i>	18:56632972-56645624	13.37	8.79	-0.61	0.04
<i>NT5C2</i>	26:23983197-24080607	33.01	21.57	-0.61	0.04
<i>FOXO1</i>	12:21915289-22009079	11.38	7.37	-0.63	0.01
<i>PMP22</i>	19:33357138-33386333	36.62	23.67	-0.63	0.01
<i>RPS6KA3</i>	X:130145854-130305125	90.58	58.17	-0.64	0.01
<i>GADD45A</i>	3:77972151-77975240	44.72	28.68	-0.64	0.01
<i>SFMBT1</i>	22:48444152-48549526	6.77	4.33	-0.65	0.04
<i>LOC786565</i>	7:54392624-54437291	39.05	24.83	-0.65	0.01
<i>CIPC</i>	10:89365865-89381544	55.74	35.31	-0.66	0.02
<i>LOC782855</i>	6:81188550-81188972	114.47	72.10	-0.67	0.02
<i>LOC104973907</i>	13:78185912-78193935	6.04	3.78	-0.68	0.01
<i>DNAJA1</i>	8:76107881-76117637	49.92	31.06	-0.68	0.01
<i>NFIL3</i>	8:87878330-87893188	21.56	13.36	-0.69	0.01
<i>TMCO3</i>	12:90698789-90720756	44.09	27.02	-0.71	0.01
<i>EEF1A1</i>	9:13230571-13237071	210.35	128.79	-0.71	0.04
<i>RGCC</i>	12:11625777-11641054	27.50	16.64	-0.72	0.01
<i>MAOA</i>	X:105380190-105462564	9.07	5.47	-0.73	0.02
<i>ARID5B</i>	28:18003670-18195880	12.92	7.75	-0.74	0.01
<i>SDC4</i>	13:74391486-74412899	35.88	21.36	-0.75	0.01
<i>CISH</i>	22:50320204-50325618	11.74	6.97	-0.75	0.01
<i>PDE4D</i>	20:18748597-20326423	52.28	30.49	-0.78	0.01
<i>ADAMTS20</i>	5:37102284-37333841	4.75	2.75	-0.79	0.01
<i>HSPH1</i>	12:29820753-29844227	12.76	7.39	-0.79	0.01
<i>KIAA1671</i>	17:67368397-67541098	8.79	5.08	-0.79	0.01
<i>CD160</i>	3:21711758-21725178	3.40	1.88	-0.85	0.01
<i>SRXN1</i>	13:60944182-60951049	9.83	5.29	-0.89	0.01
<i>LOC513548</i>	9:88273884-88299107	9.26	4.94	-0.91	0.01
<i>KLHL30</i>	3:118105157-118112080	17.70	9.10	-0.96	0.01
<i>TNFRSF12A</i>	25:2435179-2441777	64.37	32.72	-0.98	0.01
<i>CDKN1A</i>	23:10551754-10568782	27.80	13.85	-1.01	0.01
<i>MYBPH</i>	16:760047-769469	79.15	39.33	-1.01	0.02
<i>SMPDL3A</i>	9:28811575-28829543	11.21	5.53	-1.02	0.01
<i>LOC104971692</i>	3:66997750-67054450	169.28	80.23	-1.08	0.01
<i>GFPT2</i>	7:772036-817756	5.13	2.42	-1.08	0.01
<i>GNPTAB</i>	5:65912911-65996254	13.49	6.24	-1.11	0.01
<i>SPOCK2</i>	28:28304693-28329944	6.10	2.82	-1.11	0.01

<i>MICAL2</i>	15:40914779-41195753	21.93	9.98	-1.14	0.01
<i>RNF115</i>	3:21626658-21706296	86.33	39.05	-1.14	0.01
<i>LOC781186</i>	20:3132654-3195037	34.42	15.42	-1.16	0.02
<i>IL1RAP</i>	1:77187089-77343695	2.36	1.04	-1.18	0.01
<i>RGS2</i>	16:13041891-13045246	14.64	6.41	-1.19	0.01

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

HIF = Highest intramuscular fat content; LIF = Lowest intramuscular fat content; Gene: Symbol of the differentially expressed gene; Locus: location of the gene in the *Bos taurus* genome; log2 fold change: FPKM values obtained for HIF and LIF, relative expression; *q-value*: *p-value* adjusted.

Gene co-expression network analysis and prediction of hub genes

The expression levels of transcripts changed dynamically between HIF and LIF animals as shown in the heat map (Figure 2A). We computed the correlation matrix using DEG expression profiles (Figure 2B). Most DEGs have a moderately high correlation, positive or negative. We used the correlation matrix to construct the co-expression network (Figure 3). The network has 274 significantly correlated gene pairs (edges) that were discovered for 54 genes (nodes). The density and the clustering coefficient were 0.19 and 0.59, respectively. Networks with clustering and density values close to 1 contain many edges and no isolated nodes [27, 28].

When network maximal clique centrality (MCC) was examined, the genes with the highest MCC score were considered to be the hub genes. In this study, three top hub genes were selected: phosphodiesterase 4D (*PDE4D*), kelch-Like family member 30 (*KLHL30*), and interleukin 1 receptor accessory protein (*IL1RAP*). *PDE4D* belongs to a family of four *PDE4* genes, all encoding phosphodiesterases that hydrolyze the second messenger cyclic adenosine monophosphate (cAMP) [33, 34]. KLHL proteins are known to be involved in the ubiquitination process [35]. *IL1RAP* encodes the interleukin 1 receptor accessory protein [36]. Hub genes, highly interconnected with nodes in a network, have been shown to be functionally significant [37]. The hub genes found may have a putative effect on lipid metabolism in Nelore cattle.

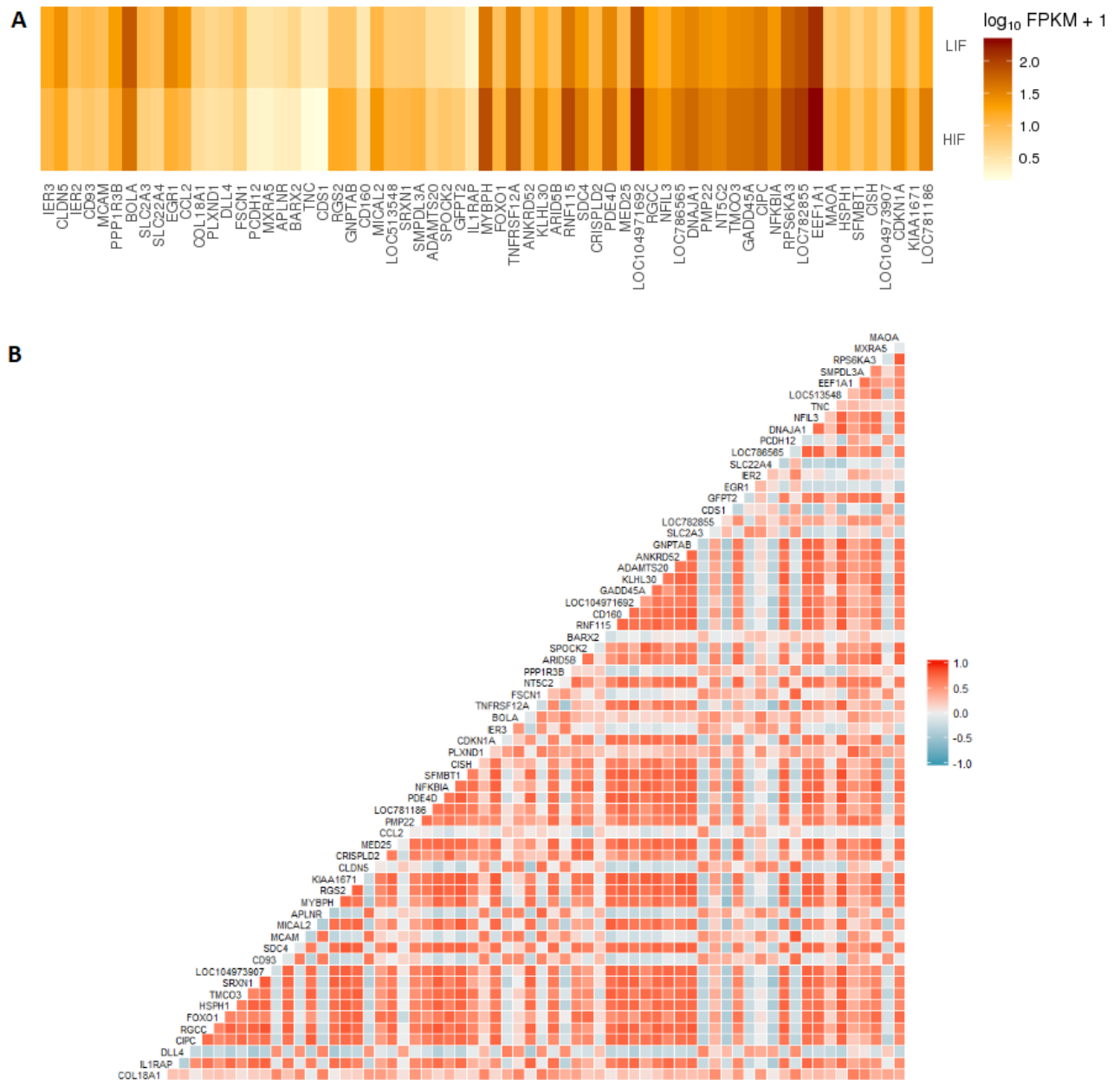


Figure 2. A) Heat map displays differentially expressed genes found in Nelore cattle divergent for intramuscular fat content. The differing colors represent differing levels of expression of those genes. B) Correlation matrix plot. The differing colors represent differing levels of correlation between genes.

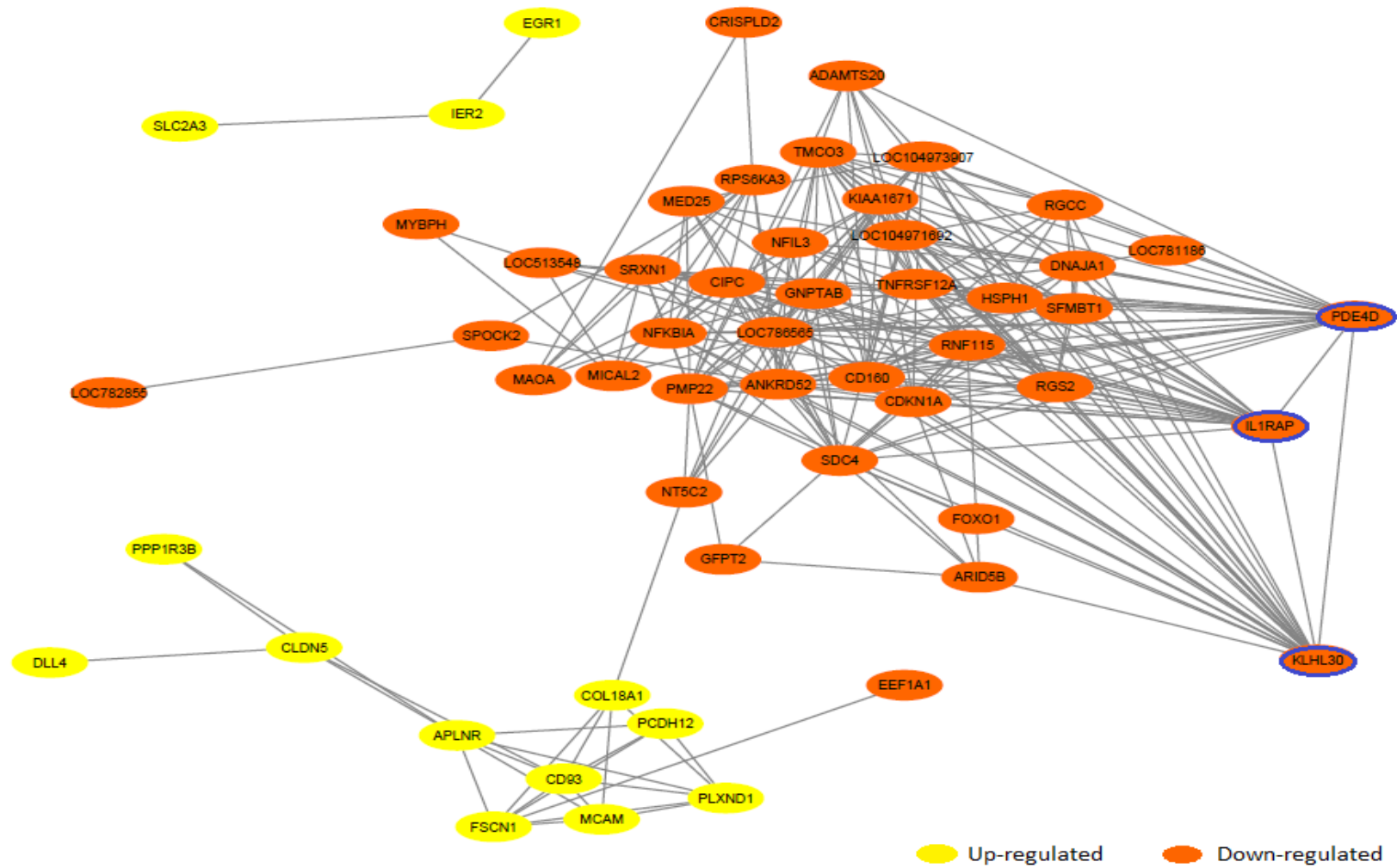


Figure 3. Gene co-expression network analysis for the differentially expressed genes in muscle of Nelore cattle divergent for intramuscular fat content. Genes circled in blue were predicted as hubs.

Gene set enrichment analysis

GO terms significantly enriched in the biological processes and pathways categories are shown in Additional file 3. The GO terms were closely associated with cell regulation and may have a putative association with lipid and muscle metabolism, for example: negative (GO:0000122) and positive (GO:0045944) regulation of transcription from RNA polymerase II promoter; the glycogen metabolic process (GO:0005977); negative regulation of the Notch signaling pathway (GO:0045746); cellular response to fibroblast growth factor stimulus (GO:0044344); and positive regulation of the reactive oxygen species metabolic process (GO:2000379). The transcription factors *BARX2*, *EGR1*, *MDE5*, and *FOXO1* were found in most of the GO terms mentioned above.

GO terms and pathways related to the immune system were also observed: antigen processing and presentation of peptide antigen via MHC class I (GO:0002474), immune response (GO:0006955), positive regulation of interleukin-5 production (GO:0032754), and cell adhesion molecules (CAMs) (bta04514), among others. *BOLA* and *NFKBIA* genes were found in most GO terms significantly enriched in the biological processes category and pathways associated with the immune system. *BOLA* genes, upregulated in HIF animals, were previously associated with intramuscular fat deposition in highly divergent marbling phenotypes of Hanwoo cattle [38] and with growth traits in *Bos Taurus* [39]. *NFKBIA* was previously reported to be associated with gain in crossbred beef steers [40].

Discussion

Few studies have investigated how the muscle transcriptome affects intramuscular fat deposition in Nelore cattle. Our current gene expression analyses show that DEGs, biological processes, and pathways may play very important roles in IF deposition; furthermore, DEGs may play roles in muscle metabolism and the immune system. Three hub genes were selected as a result of co-expression analysis: *PDE4D*, *KLHL30*, and *IL1RAP* (downregulated in HIF animals). *PDE4D* play a role in lipid metabolism [41], *KLHL30* has functions on muscle metabolism [35], and *IL1RAP* play a role in immune system [42]. The following discussion explains the biological processes and potential roles of the

DEGs identified in this study, especially the hub genes, in lipid and muscle metabolism and the immune system.

Lipid Metabolism

The protein coding gene *PDE4D* was identified as a hub gene. This gene was downregulated in HIF animals and was previously associated in genome-wide association study (GWAS), with marbling score in a commercial Hanwoo cattle population [43]. GO annotations related to *PDE4D* include 3'–5'-cyclic adenosine monophosphate (cAMP) binding [44]. Xu et al. [41] showed that where the expression of *PDE4* was inhibited in mice, intracellular cAMP levels increased. cAMP is known to have a significant role in adipogenesis [45].

The cAMP signaling pathways are among the well-characterized mechanisms controlling adipocyte differentiation. PDE inhibitors and synthetic cAMP analogs are commonly employed to switch on the adipogenic program in vitro [46]. According to our results, the *PDE4D* gene exhibited the highest centrality, indicating that it has influence on the other DEGs identified in this study. The fact that *PDE4D* was downregulated in HIF animals may increase the intracellular cAMP and thus may contribute to an increase in IF content of Nelore cattle. In this study, *PDE4D* participated in the positive regulation of interleukin-5 production (GO:0032754), revealing that this gene also has effects on the immune system (details will be discussed later).

Other DEGs were also associated with lipid metabolism, for example: *CDS1*, *SLC22A4*, *SLC2A3*, *FOXO1*, and *MED25*. According to our results, these genes may play a putative role in the regulation of other expressed genes and may be candidates for different traits in cattle, including IF content in Nelore cattle. *CDS1*, upregulated in HIF animals, plays an important role in the synthesis of triglycerides. Qi et al. [47] elucidated the role of *CDS*, a key enzyme in phospholipid metabolism, in cellular lipid storage as well as in the differentiation of adipocytes. They demonstrated that *CDS1* may regulate the expansion of lipid droplets essential for adipogenesis.

The *SLC* family genes are responsible for the transport of glucose and other sugars, bile salts, and organic acids [48, 49]. *SLC22A4* (*OCTN1*), differentially expressed in this study and upregulated in HIF animals, is a plasma membrane carnitine transporter (an organic cation transporter) which is

an important component of the carnitine system. Adequate carnitine levels are required for normal lipid metabolism and are important for energy metabolism [50, 51]. Members of the *SLC2A* family (facilitating hexose transporters) are essential in the regulation of cellular metabolism [52]. *SLC2A3* (*GLUT3*), upregulated in HIF animals, was classically defined as a neuronal glucose transporter due to its high level of expression and initial characterization in nervous tissue [53]. However, *SLC2A3* is also highly expressed in other tissues with high energy needs, including muscle [54, 55]. The animals in the HIF group had higher expression levels of genes associated with lipid metabolism, indicating that these genes may contribute to adipogenesis and the maintenance of lipid metabolism in HIF animals.

A few notable examples of transcription factors that regulate lipid metabolism include forkhead O box proteins (FoxOs) and hepatic nuclear factors (HNFs) [56]. The *FOXO1* gene (belonging to the FoxOs proteins family), downregulated in HIF animals, is one of the most important transcriptional effectors, promoting the expression of genes involved in gluconeogenesis [57, 58]. This gene is widely expressed in all tissue types. *FOXO1* plays a role in hepatic lipid and lipoprotein pathways, potentially considered the central link for the regulation and coordination of insulin action on carbohydrate and lipid metabolism [59, 60]. Previous studies showed that lipogenesis is modulated by *FOXO1* through sterol regulatory element binding protein 1c (SREBP-1c) [59, 60].

The most common function of mediator complex (MED) is to regulate RNA polymerase II-dependent gene transcription [61]. In this study, *MED25* was in three GO terms associated with cell regulation: negative (GO:0000122) and positive (GO:0045944) regulation of transcription from RNA polymerase II promoter and positive regulation of cell cycle arrest (GO:0071158). *MED25*, downregulated in HIF animals, is a transcription factor that belongs to a variable functional complex that controls the constitutive expression of genes. Rana et al. [62] showed that *MED25* is a cofactor of HNF4 α (nuclear factor), i.e., *MED25* plays a vital role in the modulation of the transcriptional activity of HNF4 α . HNF4 α regulates genes that are responsible for lipid and drug metabolism, such as cytochrome P-450. These researchers also showed that down regulation of *MED25* impairs a specific set of HNF4 α target genes, suggesting a role for

MED25 in metabolism and lipid homeostasis. Our results identified DEGs related to lipid metabolism that may be potential biomarkers of IF content in Nelore cattle.

Muscle Metabolism

KLHL30, downregulated in HIF animals, was identified as a hub gene belonging to the kelch (Klhl) superfamily that consists of a large number of structurally and functionally diverse proteins characterized by the presence of a kelch-repeat domain [35]. Klhl proteins are involved in a number of cellular and molecular processes such as cell migration, cytoskeletal arrangement, regulation of cell morphology, protein degradation (ubiquitination process), and gene expression [63]. The ubiquitin-proteasome system degrades the major proteins of contractile skeletal muscle and plays an important role in muscle loss [64]. Piórkowska et al [65] showed that *KLHL30* was differentially expressed in muscle tissue of chickens selected for differential shear force. This result, together with ours, shows the pleiotropic activities of *KLHL30* in producing qualitative traits in animals.

Among the DEGs related to muscle metabolism, we found known genes associated with growth and muscle development such as: *COL18A1* (upregulated in HIF animals) and *MYBPH* (downregulated in HIF animals). *COL18A1* encodes the alpha chain of type XVIII collagen. This collagen is one of the multiplexins, extracellular matrix proteins that contain multiple triple-helix domains (collagenous domains) interrupted by non-collagenous domains [66]. *COL18A1* was differentially expressed in *longissimus thoracis* muscle of Qinchuan male cattle [67]. *MYBPH* is a protein coding gene that binds to myosin. This gene is involved in early skeletal muscle development in humans [68]. *MYBPH* might also signal reduced muscle regeneration [68]. *MYBPH* protein was negatively correlated with tenderness in Charolais cattle [69].

BARX2 and *EGR1* (transcriptional factors) were upregulated in HIF animals. *BARX2* encodes a member of the homeobox transcription factor family. Makarenkova et al. [70] reported that *BARX2* cooperates with other transcription factors expressed in muscle tissue to regulate cytoskeletal remodeling events of precocious myoblastic differentiation. Previous studies have reported that the homeobox *Barx2* protein is expressed during skeletal muscle development and

plays an important role in the integrity of myofibrils [71, 72]. *BARX2* was associated with puberty in *Bos indicus* (Brahman) in age and weight-matched heifers [73] and expressed in the epithelial and intralobular stromal compartments in Holstein cows [74].

EGR1 is a nuclear protein and functions as a transcriptional regulator. This gene is involved in the differentiation of satellite cells derived from bovine skeletal muscle. The over-expression of *EGR1* increased the expression of *MYOG* mRNA and protein [75]. *ERG1* was expressed in muscle and has been associated with residual feed intake of the Nelore cattle [10].

Cellular response to fibroblast growth factor stimulus (GO:0044344) was overrepresented in this study. Two genes were related to this GO term: C-C motif chemokine ligand 2 (*CCL2*) and delta-like canonical notch ligand 4 (*DLL4*), both of which were upregulated in HIF animals. *CCL2* was associated with intramuscular adipocyte differentiation in Japanese Black cattle [76] and *DLL4* was associated with beef tenderness induced by acute stress in Angus cattle [77].

Our results showed that IF content differences in Nelore cattle may be controlled by genes that regulate growth and muscle development, if the IF content results from the balance between the absorption, synthesis, and degradation of triacylglycerols, which involve many metabolic pathways in both adipocytes and myofibers [78].

Immune system

IL1RAP, downregulated in HIF, was identified as a hub gene encoding the interleukin 1 receptor accessory protein. It is a necessary part of the interleukin 1 receptor complex which initiates signaling events that result in the activation of interleukin 1-responsive genes. Interleukins are a group of cytokines which play a role, primarily, in the immune system and in glucose/lipid metabolism [42]. Interleukins are expressed in a wide range of cells of the immune, neural, and endocrine systems, reflecting the pleiotropic activities of this molecule [79]. *IL1RAP* and *PDE4D* (hub genes) were present in the positive regulation of interleukin-5 production (GO:0032754). *PDE4D* plays a role in lipid metabolism and also has regulatory effects on innate immunity [80]. Where *PDE4* is inhibited, cAMP levels are increased, resulting in inhibition of NF-kappa-B inhibitor-driven

gene transcription [80]. *NFKBIA* (NF-kappa-B inhibitor) was differentially expressed and downregulated in HIF animals. The *IL1RAP*, *PDE4D*, and *NFKBIA* downregulation may contribute to intracellular the increase in cAMP levels and thus may have effects on IF content differences in Nelore cattle, if cAMP plays a role in lipid systems.

Immune cells such as macrophages depend of the recognition of lipid ligands by membrane proteins, surface/extracellular, and intracellular immune receptors. Involvement of the lipid receptors triggers an immune response. Oxidized lipids activate nuclear receptors, which play a role in lipid homeostasis and also regulate, for example, the immune response directed by *NFKB*. Activation of macrophages promotes the production of cytokines and the induction of acute phase response, accompanied by systemic lipid changes [81]. The *NFKBIA* gene was found in two GO terms and six pathways. Insulin resistance (bta04931) was among the pathways found. Insulin is synthesized by beta cells of islets of Langerhans in the pancreas and is the most important hormone in the regulation of energy metabolism. This hormone is essential for carbohydrate intake, protein synthesis, and fat storage [82]. Insulin resistance causes some disorders of lipid metabolism: increased triglyceride and decreased HDL levels [83].

In addition to *NFKBIA*, the *BOLA* gene (upregulated in HIF animals) was present in two GO terms (antigen processing and presentation of peptide antigen via MHC class I - GO:0002474 and immune response - GO:0006955) and five pathways related to the immune system (see Additional file 3), for example, cell adhesion molecules (CAMs) (bta04514). *BOLA* has been associated with marbling in Hanwoo (Korean Cattle) [38], tenderness [7], and reproductive performance [84] in Nelore cattle.

The immune system can influence lipids and lipoprotein levels [85]. Lipids, besides being structural components of cellular membranes and serving as fuel stores, also play roles as effectors and second messengers associated with the immune system [86]. Our results identified the candidate genes for IF content in Nelore cattle which play a role in the immune system and have a putative effect in lipid metabolism.

Conclusions

Our results show that muscle cells of Nelore cattle phenotypically divergent for IF expressed genes related to lipid and muscle metabolism, as well as genes related to the immune system. Gene co-expression network analysis showed hub genes, which are the main regulators of other DEGs identified and consequently may play a role in Nelore cattle's cellular or systemic lipid biology. This study identified potential biomarkers and molecular mechanisms involved in IF content difference in Nelore cattle.

List of abbreviations

Differentially expressed genes (DEGs); Intramuscular fat (IF); Highest intramuscular fat content (HIF); Lowest intramuscular fat content (LIF); Maximal clique centrality (MCC); False discovery rates (FDR); Fragments per kilobase of transcript per million reads mapped (FPKM); Gene ontology (GO); Genome-wide association study (GWAS); CDP-diacylglycerol synthase 1 (*CDS1*); solute carrier superfamily genes (*SLC22A4* and *SLC2A3*); collagen type XVIII alpha 1 chain (*COL18A1*); myosin binding protein H (*MYBPH*); bola family member (*BOLA*); NFkB inhibitor alpha (*NFKBIA*); mediator complex subunit 25 (*MED25*); BARX homeobox 2 (*BARX2*); early growth response 1 (*EGR1*); forkhead box O1 (*FOXO1*); CLOCK interacting pacemaker (*CIPC*); nuclear factor and interleukin 3 regulated (NFIL3/ E4BP4); immediate early response 2 (*IER2*); GlcNAc-1-phosphotransferase subunits alpha/beta (*GNPTAB*); forkhead O box proteins (FoxOs); phosphodiesterase 4D (*PDE4D*); kelch-Like family member 30 (*KLHL30*); and interleukin 1 receptor accessory protein (*IL1RAP*).

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). The animals were provided by Qualitas Nelore breeding program company and they were slaughtered in commercial slaughterhouses. These slaughterhouses have animal welfare departments staffed by professionals trained by WAG (World Animal

Protection) to ensure that the animals are killed humanely using a captive bolt pistol for stunning.

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 and grant #2015/16850-9).

Acknowledgements

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

Additional Files

Additional File 1

Additional File 2

Additional File 3

References

1. Wood JD, Enser M, Fisher AV, Nute GR, Sheard PR, Richardson RI, et al. Fat deposition, fatty acid composition and meat quality: A review. *Meat Sci.* 2008;78(4):343-358.
2. Pickworth CL, Loerch SC, Velleman SG, Pate JL, Poole DH, Fluharty FL. Adipogenic differentiation state-specific gene expression as related to bovine carcass adiposity. *J Ani Sci.* 2010;89(2):355-366.
3. Costa ASH, Pires VMR, Fontes CMGA, Prates JAM. Expression of genes controlling fat deposition in two genetically diverse beef cattle breeds fed high or low silage diets. *BMC Veterinary Research.* 2013;9(118):1-16.
4. Michal JJ, Zhang ZW, Gaskins CT, Jiang Z. The bovine fatty acid binding protein 4 gene is significantly associated with marbling and subcutaneous fat depth in Wagyu x Limousin F2 crosses. *Anim Genet.* 2006;37(4):400-2.
5. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotech.* 2010;28(5):511-515.
6. He H, Liu X. Characterization of transcriptional complexity during longissimus muscle development in bovines using high-throughput sequencing. *PLoS One.* 2013;8:e64356.
7. Fonseca LFS, Gimenez DFJ, Silva DBS, Barthelson R, Baldi F, Ferro JA, et al. Differences in global gene expression in muscle tissue of Nelore cattle with divergent meat tenderness. *BMC Genomics.* 2017;18(945):1-12.

8. Berton MP, Fonseca LFS, Gimenez DFJ, Utembergue BL, Cesar ASM, Coutinho LL, et al. Gene expression profile of intramuscular muscle in Nelore cattle with extreme values of fatty acid. *BMC Genomics*. 2016;17:972.
9. Cesar AS, Regitano LC, Koltjes JE, Fritz-Waters ER, Lanna DP, Gasparin G, et al. Putative regulatory factors associated with intramuscular fat content. *PLoS One*. 2015;10 (6):1-21.
10. Tizioto PC, Coutinho LL, Oliveira PSN, Cesar AS, Diniz WJ, Lima AO, et al. Gene expression differences in Longissimus muscle of Nelore steers genetically divergent for residual feed intake. *Sci Rep*. 2016; 6:39493.
11. Diniz WJ, Coutinho LL, Tizioto PC, Cesar AS, Gromboni CF, Nogueira AR, et al. Iron content affects Lipogenic Gene expression in the muscle of Nelore beef cattle. *PLoS One*. 2016;11(8):e0161160.
12. Silva-Vignato B, Coutinho LL, Cesar ASM, Poleti MD, Regitano LCA, Balieiro JCC. Comparative muscle transcriptome associated with carcass traits of Nelore cattle. *BMC Genomics*. 2017;18:506.
13. Vaz FN, Restle J, Feijó GLD, Brondani IL, Rosa JRP, Santos AP. Qualidade e composição química da carne de bovinos de corte inteiros ou castrados de diferentes grupos genéticos Charolês x Nelore. *Rev Bras Zootec*. 2001;30(2):518-525.
14. Ribeiro EL, Hernandez JA, Zanella EL, Shimokomaki M, Prudêncio-Ferreira SH, Youssef E, et al. Growth and carcass characteristics of pasture fed LHRH immunocastrated, castrated and intact *Bos indicus* bulls. *Meat Sci*. 2004;68:285-290.
15. Filteau M, Pavey SA, St-Cyr J, Bernatchez L. Gene Coexpression networks reveal key drivers of phenotypic divergence in Lake whitefish. *Mol Biol Evol*. 2013;30(6):1384–96.
16. Lim D, Lee S-H, Kim N-K, Cho Y-M, Chai H-H, Seong H-H, et al. Gene co-expression analysis to characterize genes related to marbling trait in Hanwoo (Korean) cattle. *Asian Australas J Anim Sci*. 2013;26(1):19–29.
17. Kogelman LJA, Cirera S, Zhernakova DV, Fredholm M, Franke L, Kadarmideen HN. Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue RNA sequencing in a porcine model. *BMC Med Genet*. 2014;7(1):57.
18. Langfelder P, Mischel PS, Horvath S. When is hub gene selection better than standard meta-analysis? *PLoS One*. 2013;8:e61505.
19. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. *Can J Biochem Phys*. 1959;37(8):911–917.
20. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*. 2014;30:2114-20.
21. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12:357–360.
22. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks. *Nat Protocols*. 2012;(3):562-578
23. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc*. 1995;57(1):1-12.
24. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-504.
25. Hui S, Sander C, Potylitsine E, Whitaker W, Bader G, Morrison L, et al. ExpressionCorrelation. 2015. Available online at: <http://apps.cytoscape.org/apps/expressioncorrelation>
26. Assenov Y, Ramírez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. *Bioinformatics*. 2008;24(2):282-4.

27. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet.* 2004;5:101-113.
28. Dong J, Horvath S. Understanding network concepts in modules. *BMC Syst Biol.* 2007;1:24.
29. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol.* 2014;8(4):11.
30. Huang W, Sherman BT, Lempick RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocols.* 2009;4:44 – 57.
31. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000;28(1):27-30.
32. Phillips T, Hoopes L. Transcription factors and transcriptional control in eukaryotic cells. *Nat Education.* 2008;1(1):119.
33. Kim JH, Ovilo C, Park EW, Fernandez A, Lee JH, Jeon JT, et al. Minimizing a QTL region for intramuscular fat content by characterizing the porcine Phosphodiesterase 4B (PDE4B) gene. *BMB Rep.* 2008;41(6):466-471.
34. Lee KT, Byun MJ, Kang KS, Park EW, Lee SH, Cho S., et al. Neuronal genes for subcutaneous fat thickness in human and pig are identified by local genomic sequencing and combined SNP association study. *PLoS One.* 2011;6(2):e16356.
35. Dhanoa BS, Cogliati T, Satish AG, Bruford EA, Friedman JS. Update on the Kelch-like (KLHL) gene family. *Hum Genomics.* 2013;7:13.
36. Matsuki T, Horai R, Sudo K, Iwakura Y. IL-1 Plays an Important Role in Lipid Metabolism by Regulating Insulin Levels under Physiological Conditions. *J Exp Med.* 2003;198(6):877-888.
37. Casci T. Network fundamentals, via hub genes. *Nat Rev Genet.* 2006;7:664–665.
38. Lee SH, Gondro C, Van der Werf J, Kim NK, Lim DJ, Park EW, et al. Use of a bovine genome array to identify new biological pathways for beef marbling in Hanwoo (Korean Cattle). *BMC Genomics.* 2010;11:623.
39. Batra TR, Lee AJ, Gavora JS, Stear MT. Class 1 alleles of the bovine major histocompatibility system and their association with economic traits. *J. Dairy Sci.* 1989;72:2115.
40. Lindholm-Perry AK, Keel BN, Zarek CM, Keele JW, Kuehn LA, Snelling WM, et al. 325 Genes in skeletal muscle associated with gain and intake identified in a multiseason study of crossbred beef steers. *J Ani Sci.* 2017;95(4):161.
41. Xu Z, Huo J, Ding X, Yang M, Li L, Dai J, Hosoe K, et al. Coenzyme Q10 Improves Lipid Metabolism and Ameliorates Obesity by Regulating CaMKII-Mediated PDE4 Inhibition. *Sci Rep.* 2017;7(1):8253.
42. Tsao CH, Shiau MY, Chuang PH, Chang YH, Hwang J. Interleukin-4 regulates lipid metabolism by inhibiting adipogenesis and promoting lipolysis. *J Lipid Res.* 2014;55(3):385-397.
43. Sudrajat P, Sharma A, Dang CG, Kim JJ, Kim KS, Lee JH, et al. Validation of Single Nucleotide Polymorphisms Associated with Carcass Traits in a Commercial Hanwoo Population. *Asian-Australas J Anim Sci.* 2016;29(11):1541-1546.
44. Fertig BA, Baillie GS. PDE4-Mediated cAMP Signalling. *JCDD.* 2018;5(1):8.
45. Ravnskjaer K, Madiraju A, Montminy M. Role of the cAMP Pathway in Glucose and Lipid Metabolism. *Handb Exp Pharmacol.* 2016;233:29-49.
46. Russell TR, Ho R. Conversion of 3T3 fibroblasts into adipose cells: triggering of differentiation by prostaglandin F₂α and 1-methyl-3-isobutyl xanthine. *Proc Natl Acad Sci USA.* 1976; 73(12):4516-20.
47. Qi Y, Kapterian TS, Du X, Ma Q, Fei W, Zhang Y, et al. CDP-diacylglycerol synthases regulate the growth of lipid droplets and adipocyte development. *J Lipid Res.* 2016;57(5):767-80.

48. Halestrap AP, Meredith D. The SLC16 gene family—from monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond. *Pflugers Arch.* 2004;447(5): 619-628.
49. Pérez de Heredia F, Wood IS, Trayhurn P. Hypoxia stimulates lactate release and modulates monocarboxylate transporter (MCT1, MCT2, and MCT4) expression in human adipocytes. *Pflugers Arch.* 2010;459,509–518.
50. Akisu M, Kultursay N, Coker I, Huseyinov A. Myocardial and hepatic free carnitine concentrations in pups of diabetic female rats. *Ann Nutr Metab.* 2002;46:45–48.
51. Santiago JL, Martínez A, de la Calle H, Fernández-Arquero M, Figueredo MÁ, de la Concha EG, et al. Evidence for the association of the SLC22A4 and SLC22A5 genes with Type 1 Diabetes: a case control study. *BMC Medical Genetics.* 2006;7:54.
52. Hresko RC, Kraft TE, Quigley A, Carpenter EP, Hruz PW. Mammalian Glucose Transporter Activity Is Dependent upon Anionic and Conical Phospholipids. *The J Biol Chem.* 2016;291(33):17271-17282.
53. Vannucci SJ, Maher F, Simpson IA. Glucose transporter proteins in brain: delivery of glucose to neurons and glia. *Glia.* 1997;21:2–21.
54. Zhao FQ, Glimm DR, Kennelly JJ. Distribution of mammalian facilitative glucose transporter messenger RNA in bovine tissue. *Int J Biochem.* 1993;25(12):1897-903.
55. Shimamoto S, Ijiri D, Kawaguchi M, Nakashima K, Ohtsuka A. Gene expression pattern of glucose transporters in the skeletal muscles of newly hatched chicks. *Biosci Biotechnol Biochem.* 2016;80(7):1382–1385.
56. Brown MS, Goldstein JL. Selective versus total insulin resistance: a pathogenic paradox. *Cell Metab.* 2008;7(2):95-6.
57. Matsumoto M, Poci A, Rossetti L, Depinho RA, Accili D. Impaired regulation of hepatic glucose production in mice lacking the forkhead transcription factor Foxo1 in liver. *Cell Metab.* 2007;6:208–216.
58. Munekata K, Sakamoto K. Forkhead transcription factor Foxo1 is essential for adipocyte differentiation. *In Vitro Cell Dev Biol Anim.* 2009;45:642-51.
59. Sparks JD, Dong HH. FoxO1 and hepatic lipid metabolism. *Curr Opin Lipidol.* 2009;20(3):217-226.
60. Li Y, Ma Z, Jiang S, Hu W, Li T, Di S, et al. A global perspective on FOXO1 in lipid metabolism and lipid-related diseases. *Prog Lipid Res.* 2017;66:42-49.
61. Youn DY, Xiaoli AM, Pessin JE, Yang F. Regulation of metabolism by the Mediator complex. *Biophysics Reports.* 2016;2(2):69-77.
62. Rana R, Surapureddi S, Kam W, Ferguson S, Goldstein JA. Med25 is required for RNA polymerase ii recruitment to specific promoters, thus regulating xenobiotic and lipid metabolism in human liver. *Mol Cell Biol.* 2011;31(3):466-481.
63. Gupta VA, Beggs AH. Kelch proteins: emerging roles in skeletal muscle development and diseases. *Skelet Muscle.* 2014;4:11.
64. Attaix D, Ventadour S, Codran A, Béchet D, Taillandier D, Combaret L. The ubiquitin-proteasome system and skeletal muscle wasting. *Essays Biochem.* 2005;41:173-86.
65. Piórkowska K, Żukowski K, Nowak J, Połtowicz K, Ropka-Molik K, Gurgul A. Genome-wide RNA-Seq analysis of breast muscles of two broiler chicken groups differing in shear force. *Anim Genet.* 2016;47(1):68-80.
66. Passos-Bueno MR, Suzuki OT, Armelin-Correa LM, Sertié AL, Errera FI, Bagatini K, et al. Mutations in collagen 18A1 and their relevance to the human phenotype. *An Acad Bras Cienc.* 2006;78(1):123-31.
67. Zhang Y, Zan LA, Wang HB, Qing L, Wu KX, Quan SA. Differentially expressed genes in skeletal muscle tissues from castrated Qinchuan cattle males compared with those from intact males Y. *Livest Sci.* 2011;135(1):76-83.
68. Loell I, Raouf J, Chen YW, Chen YW, Shi R, Nennesmo I, Alexanderson H, et al. Effects on muscle tissue remodeling and lipid metabolism in muscle tissue from

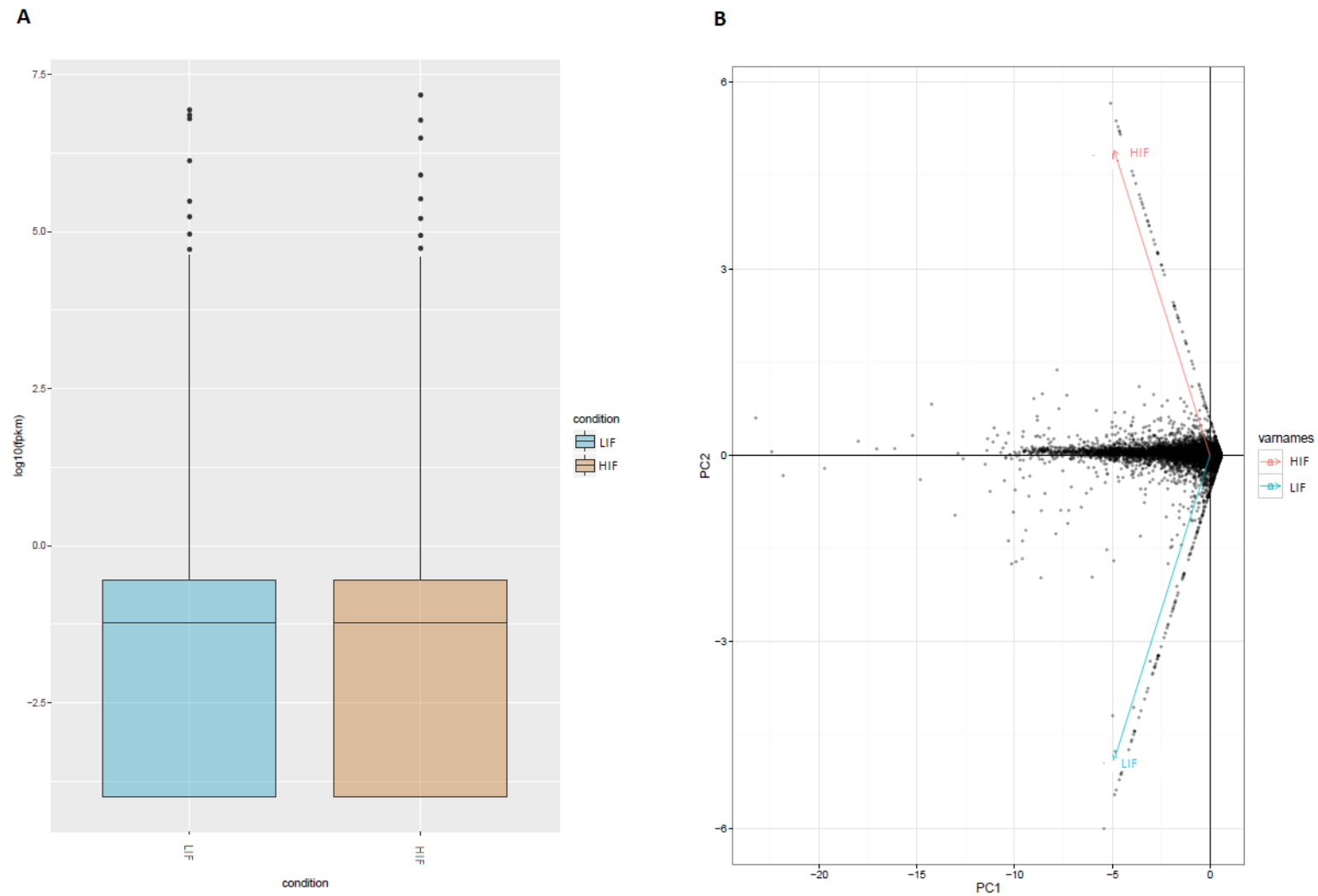
- adult patients with polymyositis or dermatomyositis treated with immunosuppressive agents. *Arthritis Res Ther.* 2016;18(1):136.
69. Guillemin N, Bonnet M, Jurie C, Picard B. Functional analysis of beef tenderness. *J Proteomics.* 2011;75(2):352-65.
 70. Makarenkova HP, Gonzalez KN, Kiosses WB, Meech R. Barx2 controls myoblast fusion and promotes MyoD-mediated activation of the smooth muscle alpha actin gene. *J Biol Chem.* 2009;284(22):14866-74.
 71. Meech R, Makarenkova H, Edelman DB, Jones FS. The homeodomain protein Barx2 promotes myogenic differentiation and is regulated by myogenic regulatory factors. *J Biol Chem.* 2003;278:8269–8278.
 72. Meech R, Gomez M, Woolley C, Barro M, Hulin JA, Walcott EC, et al. The Homeobox Transcription Factor Barx2 Regulates Plasticity of Young Primary Myofibers. *PLoS One.* 2010;5(7):e11612.
 73. Fortes MR, Nguyen LT, Weller MM, Cánovas A, Islas-Trejo A, Porto-Neto LR, et al. Transcriptome analyses identify five transcription factors differentially expressed in the hypothalamus of post-versus prepubertal Brahman heifers. *J Ani Sci.* 2016;94(9): 3693-3702.
 74. Casey T, Dover H, Liesman J, DeVries L, Kiupel M, Vandehaar M, et al. Transcriptome analysis of epithelial and stromal contributions to mammaryogenesis in three week prepartum cows. *PLoS One.* 2011;6(7):e22541.
 75. Zhang W, Tong H, Zhang Z, Shao S, Liu D, Li S, Yan Y. Transcription factor EGR1 promotes differentiation of bovine skeletal muscle satellite cells by regulating MyoG gene expression. *J Cell Physiol.* 2018;233(1):350-362.
 76. Mizoguchi Y, Hirano T, Itoh T, Aso H, Takasuga A, Sugimoto Y, et al. Differentially expressed genes during bovine intramuscular adipocyte differentiation profiled by serial analysis of gene expression. *Anim Genet.* 2010;41(4):436-41.
 77. Zhao C, Tian F, Yu Y, Luo J, Mitra A, Zhan F, et al. Functional genomic analysis of variation on beef tenderness induced by acute stress in angus cattle. *Comparative and Functional Genomics.* 2012;2012:756284.
 78. Hocquette J, Gondret F, Baéza E, Médale F, Jurie C, Pethick DW. Intramuscular fat content in meat-producing animals: Development, genetic and nutritional control, and identification of putative markers. *Animal.* 2010;4(2) 303-319.
 79. Tocci MJ, Schmidt J.A. Interleukin-1: Structure and function. *Cytokines in health and disease, second edition.* In Remick DG, Friedland, JS, editors. Marcel Dekker, Inc., New York; 1997. p.1–27.
 80. Schafer PH, Parton A, Capone L, Cedzik D, Brady H, Evans JF, et al. Apremilast is a selective PDE4 inhibitor with regulatory effects on innate immunity. *Cell Signal.* 2016;26(9):2016-29.
 81. Neyen CD, Gordon S. *Macrófagos em lipid and immune homeostasis.* eLS. John Wiley & Sons Ltd 1. 2014;1.
 82. Dimitriadis G, Mitrou P, Lambadiari V, Maratou E, Raptis SA. Insulin effects in muscle and adipose tissue. *Diabetes Res Clin Pract.* 2011;93 Suppl 1:S52-9.
 83. Ikeoka D, Krusinova E. Insulin resistance and lipid metabolism. *Rev. Assoc. Med. Bras.* 2009;55(3):1.
 84. Melo TP, Camargo GMF, Albuquerque LG, Carneiro R. Genome-wide association study provides strong evidence of genes affecting the reproductive performance of Nelore beef cows. *PLoS One.* 2017;12(5):e0178551.
 85. Getz GS, Reardon CA. The mutual interplay of lipid metabolism and the cells of the immune system in relation to atherosclerosis. *Clin Lipidol.* 2014;9(6):657-671.
 86. Dowds CM, Kornell SC, Blumberg RS, Zeissig S. Lipid antigens in immunity. *Biol Chem.* 2014;395(1):61-81.

Additional Files

Additional File 1. Descriptive statistics for alignment parameters

Highest intramuscular fat content (HIF)				Lowest intramuscular fat content (LIF)			
Sequence ID	Reads raw	Reads trimmed	Overall alignment rate (%)	Sequence ID	Reads raw	Reads trimmed	Overall alignment rate (%)
1N	28,723,235	26,921,059	96.23	16N	32,171,041	30,151,500	96.7
2N	33,597,350	30,905,727	96.62	17N	12,069,555	9,706,541	96.18
3N	32,645,772	30,068,141	96.22	18N	30,857,496	28,477,140	96.34
4N	35,748,151	32,838,508	96.57	19N	14,954,299	12,425,519	96.62
5N	13,898,951	11,215,972	96.34	20N	24,889,177	20,258,917	96.58
6N	37,716,840	30,583,246	96.4	21N	10,183,749	8,341,290	96.38
7N	33,083,107	30,3670,38	96.73	22N	15,361,700	12,673,331	96.14
8N	11,862,753	9,841036	96.31	23N	31,126,455	28,771,526	96.39
9N	27,719,519	25,817,447	96.23	24N	14,738,117	11,991,827	96.6
10N	14,386,753	25,817,447	96.23	25N	27,567,691	22,970,733	96.55
11N	14,790,642	11,559,729	96.57	26N	12,301,661	99,333,68	96.63
12N	42,738,914	34,531,411	96.81	27N	13,674,494	11,290,360	95.9
13N	51,165,141	41,041,717	96.68	28N	26,558,216	21,997,335	96.52
14N	31,387,621	26,298,812	96.31	29N	13,133,847	10,580,378	96.45
15N	29,626,011	27,625,207	96.56	30N	28,665,948	23,466,657	96.49
Mean	29,272,717	26,362,166	96.454	Mean	20,550,230	17,535,761	96.43

*IF was measured in percentage and did not present normal distribution of the data and then we performed the transformation of these data by the square root of the percentage of the arcsine function. SD = Standard deviation



Additional File 2. Global statistics and quality control. (A) Box plot of FPKM distributions for individual conditions. (B) PCA plot for gene-level features.

Additional File 3. Biological Process GO terms and pathways obtained with the DAVID software for differentially expressed genes in Nelore cattle muscle divergent for intramuscular fat content phenotypes

Category	Terms	Genes	p-value	Benjamini-Hochberg	FDR (False Discovery rate)
	GO:0002474~antigen processing and presentation of peptide antigen via MHC class I	<i>BOLA</i>	1,56E-07	5,90E-05	2,16E-04
	GO:0006955~immune response	<i>BOLA, CCL2, NFIL3</i>	2,03E-04	3,76E-02	2,80E-01
	GO:0071850~mitotic cell cycle arrest	<i>CDKN1A, RGCC, GADD45A</i>	5,50E-04	6,69E-02	7,58E-01
	GO:0030198~extracellular matrix organization	<i>COL18A1, ADAMTS20, CRISPLD2, SPOCK2</i>	2,81E-03	2,34E-01	3,82E+00
	GO:1903748~negative regulation of establishment of protein localization to mitochondrion	<i>HSPH1, DNAJA1</i>	7,94E-03	4,53E-01	1,04E+01
	GO:0032754~positive regulation of interleukin-5 production	<i>IL1RAP, PDE4D</i>	2,36E-02	7,78E-01	2,82E+01
	GO:0000122~negative regulation of transcription from RNA polymerase II promoter	<i>EGR1, BARX2, DLL4, MED25, FOXO1, NFIL3</i>	2,54E-02	7,50E-01	2,99E+01
	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	<i>EGR1, HSPH1, RPS6KA3, RGCC, MED25, FOXO1, NFKBIA</i>	2,70E-02	7,26E-01	3,15E+01
Biological Process	GO:2000353~positive regulation of endothelial cell apoptotic process	<i>COL18A1, RGCC</i>	3,53E-02	7,79E-01	3,91E+01
	GO:0044344~cellular response to fibroblast growth factor stimulus	<i>CCL2, DLL4</i>	4,29E-02	8,10E-01	4,55E+01
	GO:0005977~glycogen metabolic process	<i>PPP1R3B, PCDH12</i>	6,56E-02	9,03E-01	6,09E+01
	GO:2000379~positive regulation of reactive oxygen species metabolic process	<i>CDKN1A, GADD45A</i>	6,56E-02	9,03E-01	6,09E+01
	GO:0045765~regulation of angiogenesis	<i>TNFRSF12A, PLXND1</i>	6,93E-02	8,96E-01	6,30E+01
	GO:0043066~negative regulation of apoptotic process	<i>IER3, ADAMTS20, DNAJA1, FOXO1</i>	7,44E-02	8,94E-01	6,57E+01
	GO:0045746~negative regulation of Notch signaling pathway	<i>DLL4, NFKBIA</i>	8,40E-02	9,07E-01	7,03E+01
	GO:0030217~T cell differentiation	<i>EGR1, DLL4</i>	8,40E-02	9,07E-01	7,03E+01
	GO:0071158~positive regulation of cell cycle arrest	<i>RGCC, MED25</i>	8,40E-02	9,07E-01	7,03E+01
	GO:0071479~cellular response to ionizing radiation	<i>CDKN1A, GADD45A</i>	9,49E-02	9,19E-01	7,48E+01
	GO:0043065~positive regulation of apoptotic process	<i>DNAJA1, FOXO1, GADD45A</i>	9,82E-02	9,13E-01	7,60E+01

	GO:0001569~patterning of blood vessels	<i>DLL4, PLXND1</i>	9,85E-02	9,00E-01	7,62E+01
	bta04931:Insulin resistance	<i>RPS6KA3, PPP1R3B, GFPT2, FOXO1, NFKBIA</i>	1,01E-03	9,97E-02	1,12E+00
	bta05166:HTLV-I infection	<i>EGR1, BOLA, CDKN1A, NFKBIA</i>	2,29E-02	7,01E-01	2,29E+01
	bta04514:Cell adhesion molecules (CAMs)	<i>BOLA, CLDN5, SDC4</i>	2,44E-02	5,75E-01	2,41E+01
Pathways	bta05168:Herpes simplex infection	<i>BOLA, CCL2, NFKBIA</i>	4,16E-02	6,69E-01	3,78E+01
	bta05169:Epstein-Barr virus infection	<i>BOLA, CDKN1A, NFKBIA</i>	4,27E-02	5,97E-01	3,86E+01
	bta05215:Prostate cancer	<i>CDKN1A, FOXO1, NFKBIA</i>	4,78E-02	5,72E-01	4,22E+01
	bta05203:Viral carcinogenesis	<i>BOLA, CDKN1A, NFKBIA</i>	6,92E-02	6,55E-01	5,52E+01

CHAPTER 4 - Differentially expressed alternatively spliced genes associated with carcass and meat traits in Nelore cattle^c

ABSTRACT – Transcript data obtained by RNA-Seq were used to identify genes for which alternatively spliced transcripts were differentially expressed in ribeye muscle tissue between Nelore cattle that differed in their ribeye muscle area (REA) or intramuscular fat content (IF). A total of 166 alternatively spliced transcripts from 125 genes were differentially expressed ($p \leq 0.05$) in ribeye muscle between the highest and lowest REA groups. For animals selected on their IF content, 269 alternatively spliced transcripts from 219 genes were differentially expressed in ribeye muscle between the highest and lowest IF groups. Cassette exons and alternative 3' splice sites were the most frequently found alternatively spliced transcripts for both traits. For both traits, some differentially expressed alternatively spliced transcripts belonged to the myosin and myotilin gene families. The hub genes identified for REA (*LRRFIP1*, *RCAN1* and *RHOBTB1*) and IF (*TRIP12*, *ANXA11* and *MAP2K6*) play a role in motility and muscle cell development. In general, genes were found for both traits with biological process GO terms that were involved in pathways related to protein ubiquitination, muscle differentiation, lipids and hormonal systems. Our results reinforce the biological importance of these known processes but also reveal new insights into the complexity of the muscle tissue transcriptome of Nelore cattle.

Introduction

The alternative processing of messenger RNAs (mRNAs) produces multiple mRNA and protein isoforms that may have similar or quite different functions in mammalian tissues^{1,2}. Next generation sequencing, applied to tissue transcriptomes (RNA-Seq), is an approach for gene expression and it's been shown to be very accurate and can measure all genes in the transcriptome at the same time. RNA-Seq could to screening the expression of functional candidate genes and identifying important molecular mechanisms creating variation in pathways those results in different tissue phenotypes. This technology has also been used to identify alternative splicing events^{3,4}.

The spliceosomal machinery may generate multiple mature mRNAs isoforms by interpreting exon/intron boundaries in several distinct ways. These different interpretations, for the most part, fall into subgroups of alternative splicing events⁵ which may be classified into: cassette exon, mutually exclusive exon, coordinate cassette exon, alternative 5' and 3' splice sites, intron retention, or alternative first and last exon events^{6,7}. The most common

^cThis chapter corresponds to the manuscript submitted to "Scientific Reports" at November 12, 2018.

alternative splicing events in mammals are cassette exon and the alternative usage of distinct 3' or 5' splice sites. Intron retention is a rare alternative splicing event in mammals^{5,8}.

The global analysis of alternative splicing by splicing sensitive microarrays or RNA-Seq have provided new insights into the contribution of different alternative splicing mechanisms to the establishment of tissue or developmental-specific gene expression programs⁵. Several studies have reported the frequencies of different types of alternative splicing events in different species^{2,4,7,9,10}. In bovine, alternative splicing events have been associated with variation in mastitis susceptibility², embryonic development¹¹ and muscle development¹². Currently, knowledge concerning the molecular mechanisms regulated by alternative splicing events, especially within the context of individual differences in carcass and meat traits in cattle, is limited, particularly for Nelore cattle.

Ribeye muscle area (REA) in conjunction with carcass weight provides an indication of carcass muscularity and meat yield¹³. The extent of intramuscular fat (IF) content of the ribeye muscle is a primary determinant of meat palatability and consumer satisfaction with the eating experience. Muscle lipid traits, which determine meat flavor and contribute to beef color, can influence the juiciness and tenderness of the meat¹⁴.

The identification of candidate genes that are differentially expressed as alternatively spliced isoforms in beef that differs for carcass muscularity and fat content may elucidate the mechanisms that lead to the regulation of these traits in Nelore cattle. The aim of this study was to utilize RNA-Seq to identify candidate genes that produced alternatively spliced transcripts that were differentially expressed in the ribeye muscle of Nelore animals that differed for REA and IF content.

Results

Differentially expressed alternatively spliced transcripts

There were 166 alternatively spliced transcripts produced from 125 genes that were differentially expressed between the high ribeye area (HREA) and low ribeye muscle area (LREA) groups (the comprehensive list of the differentially expressed alternatively spliced transcripts associated with REA is provided in

Supplementary Tables: Table S1). For REA, cassette exon was the most frequently found alternative splicing event (Figure 1). Of the differentially expressed alternatively spliced transcripts, 54 were known and 112 were novel transcripts. A total of 269 (78 known and 191 novel) alternatively spliced transcripts from 219 genes were differentially expressed between the high intramuscular fat (HIF) and low intramuscular fat (LIF) (see Supplementary Tables: Table S2). Alternative 3' splice site was the most frequently observed alternate splicing event found in the IF content analysis (Figure 1).

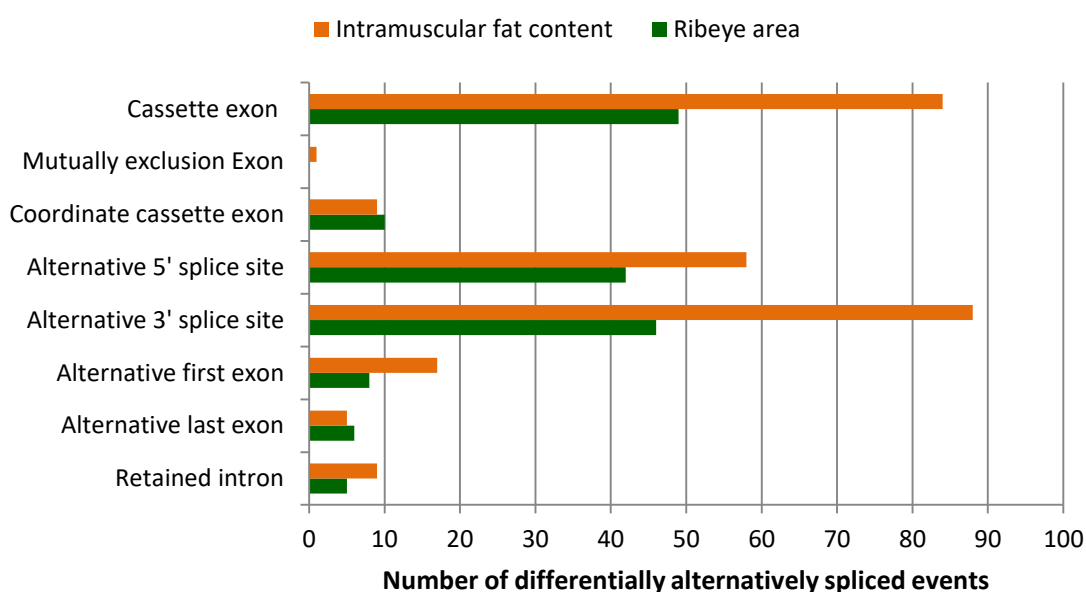


Figure 1. Alternatively spliced and differentially expressed ($p < 0.05$, t-test) transcripts identified in the ribeye muscle of Nelore bulls with the highest and lowest ribeye muscle area animals (green); or highest and lowest intramuscular fat content (orange).

For REA, some differentially expressed alternatively spliced transcripts belong to important gene families, such as the myosin (*MYBPC1* and *MYH1*), myotilin (*MYOT*) and myomesin (*MYOM2*) and play a significant role in the stability of thin filaments during muscle contraction. The transcript isoforms produced by these gene families differ by cassette exon, intron retention, alternative 3' splice site and alternative 5' splice site events (see Supplementary Tables: Table S1). The solute carrier (SLC) genes family (*SLC25A2*, *SLC25A25*, *SLC29A1* and *SLC29A2*) produced differentially expressed alternatively spliced transcripts with alternative 3' splice sites. Genes belonging to the solute carrier family play roles as transporters¹⁵. Among differentially expressed alternatively

spliced transcripts, we found *BOLA-DOA* (alternative 3' splice site). In cattle, *BOLA* is a major histocompatibility complex class II gene that exhibits very little sequence variation, especially at the protein level¹⁶.

For IF, members of the myosin (*MYH7B*, *MYLK2* and *MYO18A*) and myotilin (*MYOT*) families were among the differentially expressed alternatively spliced transcripts. These genes produced transcripts that differed by alternative first exons alternative 3' splice sites, alternative 5' splice sites and retained introns. In addition to these, members of the myozenin (*MYOZ1* and *MYOZ3*) family were found to produce differentially expressed alternatively spliced transcripts: alternative 3' splice sites, alternative 5' splice sites and alternative first exons. These genes play an important role in the modulation of calcineurin signaling. Other differentially expressed alternatively spliced transcripts belong to ubiquitin family (*UBAC2* and *UBE3C*). These genes produced transcripts with cassette exons, alternative 3' splice sites and alternative 5' splice sites. The ubiquitin-proteasome system plays an important role in muscle loss¹⁷.

Co-expression network analyses

Co-expression network analyses found 65 significantly correlated gene pairs (edges) involving 66 genes (nodes), density = 0.030 and clustering coefficient = 0.130 for REA (Figure 2). For IF, there were 92 significantly correlated gene pairs (edges) involving 101 genes (nodes), density = 0.018 and a clustering coefficient = 0.152 (Figure 3). Networks with clustering and density values close to 1 contain many edges and no isolated nodes¹⁸.

The topological properties of each network were analyzed and those genes engaging in the most interactions with other genes were predicted to be hub genes. For REA, the three hub genes with the largest numbers of interactions were: LRR Binding FLII Interacting Protein 1 (*LRRFIP1*), Regulator of Calcineurin 1 (*RCAN1*) and Rho Related BTB Domain Containing 1 (*RHOBTB1*). These genes play a role in the proliferation of smooth muscle cells¹⁹, and interact with the calcineurin A²⁰ and actin filament systems²¹, respectively. The analyses predicted that several different mRNA transcripts were processed from *LRRFIP1*: alternative last exon, cassette exon and coordinate cassette exon transcripts (see Supplementary Figures: Figure S1A). For *LRRFIP1*, cassette exon transcripts were more abundant in the HREA animals (see Supplementary

Tables: Table S1), whereas alternative last exon and coordinate cassette exon transcripts were more abundant in the LREA animals. *RCAN1* produces transcripts that differ by known alternative first exons (see Supplementary Figures: Figure S1B), which were more abundant in the LREA animals (see Supplementary Tables: Table S1). The transcript isoforms produced by *RHOBTB1* differ by alternative 3' splice sites (see Supplementary Figures: Figure S1C), and were more abundant in the LREA animals (see Supplementary Tables: Table S1).

For IF content, the three hub genes with the largest numbers of interactions were: Thyroid Hormone Receptor Interactor 12 (*TRIP12*), Annexin A11 (*ANXA11*) and Mitogen-Activated Protein Kinase Kinase 6 (*MAP2K6*). *TRIP12*, *ANXA11* and *MAP2K6* genes are involved in ubiquitin mediated proteolysis²², calcium-dependent phospholipid-binding²³ and the regulation of the mitogen-activated protein kinase pathway²⁴, respectively. *TRIP12* and *ANXA11* produce transcripts that differ by their use of novel alternative 3' splice sites (see Supplementary Figures: Figure S2A and S2B), which were more abundant in the HIF animals (see Supplementary Tables: Table S2). *MAP2K6* produces transcripts that differ by known alternative first exons, which were more abundant in the LIF animals (see Supplementary Tables and Figures: Table S2 and Figure S2C).

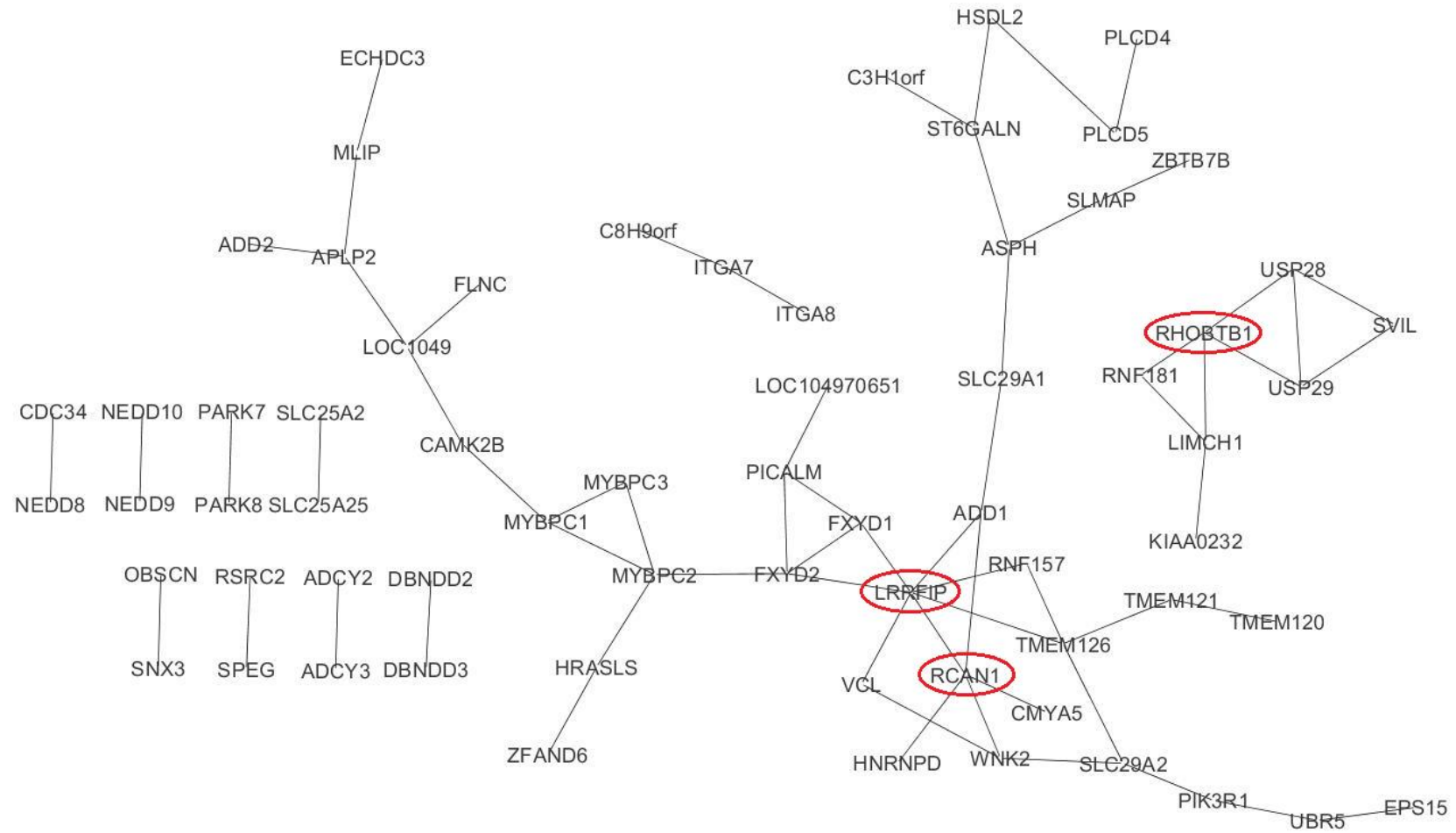


Figure 2. Transcript correlation network for the genes with differentially expressed isoforms in the ribeye muscle of Nelore cattle differing for ribeye muscle area. Hub genes are circled in red.

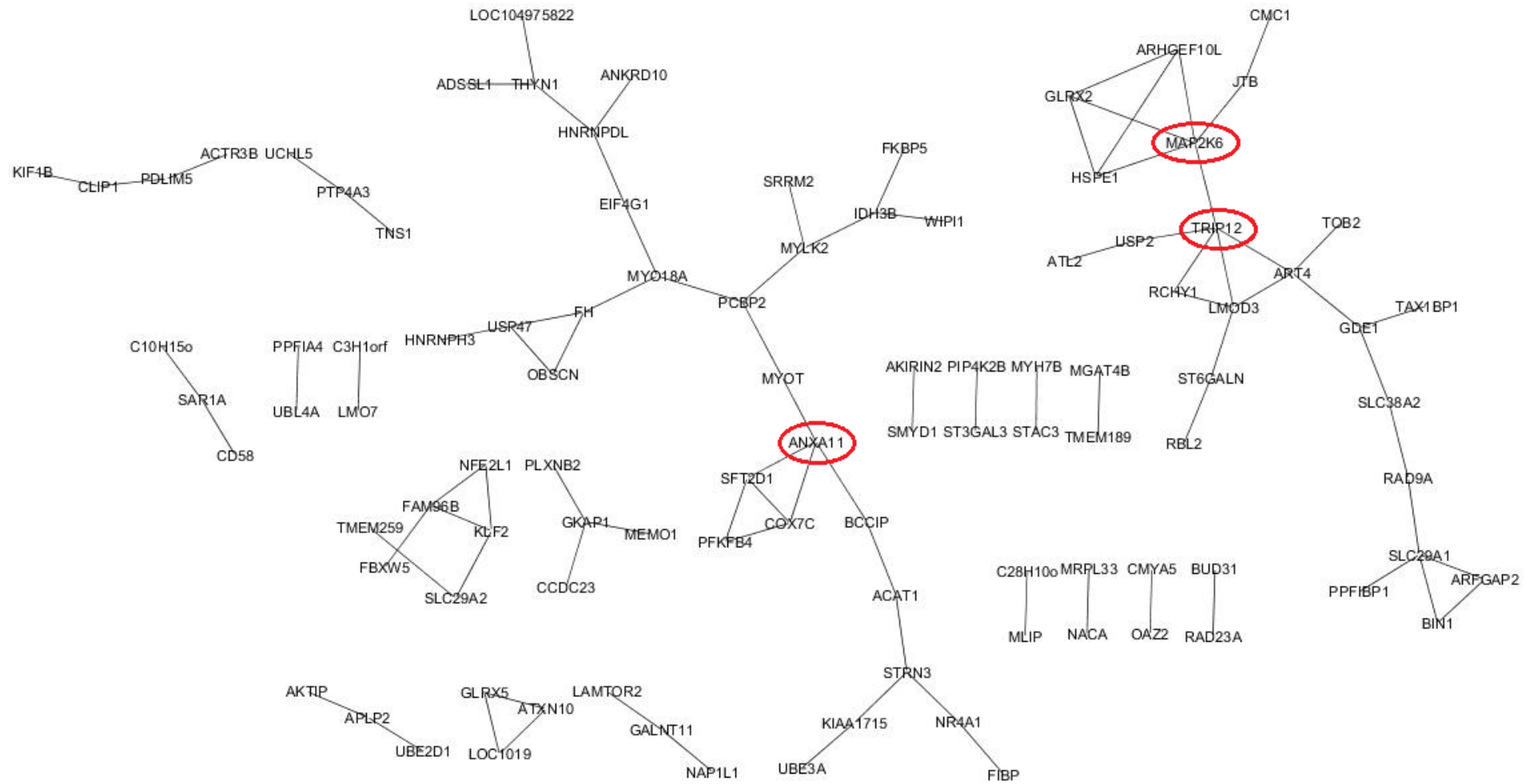


Figure 3. Transcript correlation network for the genes with differentially expressed isoforms in the ribeye muscle of Nelore bulls divergent for intramuscular fat content. Hub genes are circled in red.

Gene set enrichment analyses

For bulls selected on REA, the biological process Gene Ontology (GO) terms for the genes identified in the pairwise comparison are summarized in Supplementary Tables: Table S3. We observed GO terms and pathways related to muscle growth and muscle protein ubiquitination, for example: regulation of protein stability (GO:0031647), smooth muscle tissue development (GO:0048745), protein ubiquitination involved in ubiquitin-dependent protein catabolic process (GO:0042787), cAMP (3',5'-cyclic adenosine monophosphate) signaling (bta04024), focal adhesion (bta04510) and ECM-receptor (extracellular matrix-receptor) interaction (bta04512), among others. The analysis also found biological GO terms and pathways related to neurological, hormonal, and immune systems, such as: adrenergic signaling in cardiomyocytes (bta04261); oxytocin signaling (bta04921) and inflammatory mediator regulation of TRP channels (bta04750). Calcium/Calmodulin Dependent Protein Kinase II Beta (*CAMK2B*) – represented by a novel cassette exon transcript isoform – and Adenylate Cyclase 2 (*ADCY2*) – represented by novel alternative 3' splice site isoform – were present in the pathways mentioned above and also in six other pathways (Supplementary Tables: Table S3). The hub gene *RCAN1* was found in the oxytocin signaling pathway. In addition to oxytocin, the differentially expressed transcript isoforms were also overrepresented in the glucagon signaling pathway (bta04922). Both of these hormones are important for the maintenance of the functional integrity and quality of skeletal muscle²⁵.

Most biological terms found for animals selected to differ for IF content were associated with growth, regulation and proteolysis of muscle cells (Supplementary Tables: Table S4), for example: myofibril assembly (GO:0030239); skeletal muscle cell differentiation (GO:0035914) and ubiquitin-dependent protein catabolic process (GO:0006511). The analysis found GO terms and pathways related to lipid metabolism, such as PI3K-Akt (bta04151), mechanistic target of rapamycin (mTOR) signaling (bta04150) and negative regulation of TORC1 signaling (GO:1904262). Serine/Threonine Kinase 11 (*STK11*) and Eukaryotic Translation Initiation Factor 4B (*EIF4B*) produce transcript isoforms with alternative 3' splice sites and were present in these pathways. All of the GO terms and pathways mentioned above play important roles in muscle and lipid metabolism.

Discussion

Differentially expressed alternatively spliced transcripts and their associated genes were identified in this study, using RNA sequencing data from the ribeye muscle of Nelore bulls phenotypically divergent for REA or IF content. In cattle, most genes have multiple isoforms characterized by different event types. The cassette exon and alternative 3'-splice site events were frequently responsible for the transcript isoforms found to be differentially expressed in the ribeye muscle of bulls with the highest and lowest REA or IF content. Cassette exon and alternative 3'-splice sites are the most of the common alternative splicing events in mammalian pre-mRNAs⁸. In a study of mice, Bland et al.²⁶ identified 95 alternative splicing events that occur during the course of myogenic differentiation. Most of the splicing events involved the alternative use of cassette exons (86%).

The co-expression analyses were performed with gene sets that contained the isoforms that were differentially expressed between the highest and lowest REA and IF content groups. Through co-expression analyses hub genes for each trait were predicted that may be the primary regulators of the other genes with differentially expressed transcript isoforms found in this study. Hub genes are expected to play an important role in the biology of organisms²⁷. In addition, a functional enrichment analysis was performed for the gene sets found for each trait. In general, for both traits, the results indicated that hub genes had functions in protein ubiquitination, muscle differentiation and regulation of lipolysis in adipocytes.

This study also found biological terms and pathways related to immune function systems, indicating the pleiotropic nature of the genes and their isoforms in several 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 carcass or meat traits of beef cattle. For example, *CAMK2B* (differentially expressed in animals phenotypically divergent for REA and represented by novel cassette exon events) was present in nine pathways (Supplementary Tables: Table S3) related to the neurological, hormonal, and immune systems, including: adrenergic signaling in cardiomyocytes (bta04261); oxytocin signaling (bta04921) and inflammatory mediator regulation of TRP channels (bta04750).

CAMK2B has been associated with neurodevelopmental disorders in humans²⁸. This gene is a serine/threonine kinase that is highly abundant in brain, especially in the post synaptic density and it is possible that distinct isoforms of this gene have different cellular localizations and interact differently with calmodulin²⁸. Calmodulin mediates many crucial processes such as inflammation, metabolism, apoptosis, smooth muscle contraction, intracellular movement, short-term and long-term memory, and the immune response²⁹.

Another gene that has not previously been reported to be associated with carcass or meat traits of beef cattle is Serine/Threonine Kinase 11 (*STK11*). This gene was differentially expressed between animals phenotypically divergent for IF content and by a known isoform with an alternative 3' splice site, regulates cell polarity and functions as a tumor suppressor³⁰. *STK11* is involved in TORC1 signaling (GO:1904262); mTOR signaling (bta04150) and PI3K-Akt signaling (bta04151) (see Supplementary Tables: Table S4). All of these processes have important roles in the growth, proliferation, motility and survival of cells.

Several other genes that were predicted to produce alternatively spliced isoforms that were differentially expressed have also been implicated in muscle differentiation and growth; protein ubiquitination in muscle; and lipid metabolism. All of these processes could promote changes in ribeye muscle area and intramuscular fat content phenotypes.

Ribeye muscle area

Several differentially expressed alternatively spliced transcripts were produced by the myosin (*MYBPC1* and *MYH1*), myotilin (*MYOT*), myomesin (*MYOM2*), solute carrier (*SLC25A2*, *SLC25A25*, *SLC29A1* and *SLC29A2*) and *BOLA* (*BOLA-DOA*) genes families. Myosin, myotilin and myomesin proteins play a significant role in the stability of thin filaments during muscle contraction and are the main components of myofilaments^{31,32,33}. Genes belonging to the solute carrier family play roles as transporters¹⁵. Solute carrier genes have been associated with REA³⁴ in Nelore cattle. *BOLA-DOA* belongs to the major histocompatibility complex and has effects in the immune system. Members of the *BOLA* gene family have previously been associated with tenderness³⁵ in

Nelore cattle and growth traits in Holstein and Angus, Charolais, Hereford, Limousin and Simmental beef cattle^{36,37}.

Three hub genes (*LRRFIP1*, *RCAN1* and *RHOBTB1*) with alternatively spliced transcripts were differentially expressed between Nelore bulls selected to differ for REA. With the exception of the cassette exon splicing event, found for *LRRFIP1*, transcripts with the other four splicing events were more abundant in animals with LREA (see Supplementary Tables: Table S1). The differences in transcript splicing event usage for *LRRFIP1*, *RCAN1* and *RHOBTB1* found here may be due to the tuning of the spliceosome machinery.

RCAN1 produces a protein that inhibits calcineurin-dependent transcriptional responses by binding to the catalytic domain of calcineurin A. Calcineurin A signaling is known to be important for the proper functioning of both cardiac and skeletal muscle. In skeletal muscle, calcineurin A participates in a variety of processes including myotube differentiation, fiber type specification and dystrophic muscle damage³⁸. *RCAN1* has previously been associated with REA in a Wagyu x Angus F1 population²⁰.

Hub genes *LRRFIP1* and *RHOBTB1* play roles in muscle cell development and regulation in humans and mice^{19,21}. *LRRFIP1* is a transcriptional repressor which preferentially binds to a GC-rich consensus sequence and may control the proliferation of smooth muscle cells¹⁹. The protein encoded by *RHOBTB1* plays a role in the organization of the actin filament system and small GTPase-mediated signal transduction²¹.

The gene set enrichment analysis based on transcript isoforms that were differentially expressed in animals differing for REA revealed GO terms and pathways related to muscle growth and muscle protein ubiquitination (Supplementary Tables: Table S3). Skeletal muscle hypertrophy is an adaptive process that results largely from protein synthesis by pathways subjected to external stimuli, growth factors and other anabolic hormones, and neuronal inputs³⁹. For example, we found the known and important pathways: cAMP signaling (bta04024), focal adhesion (bta04510) and ECM-receptor (extracellular matrix-receptor) interaction (bta04512). The cAMP pathway contributes to an increase in myofiber size and metabolic phenotype over the long term⁴⁰. It also participates in the differentiation of cell muscle precursors⁴¹. All of these cellular events are necessary for the efficient regeneration of skeletal muscle⁴⁰.

Focal adhesions, serve as mechanical links to ECM-receptor and other molecules, playing key roles in important biological processes⁴². Focal adhesion kinase is involved in mammalian myoblast fusion both *in vitro* and *in vivo*⁴³. Our results indicate the biological importance of cAMP signaling, focal adhesion and ECM-receptor signaling in determining the composition and organization of bovine muscle tissue.

Pathways related to hormonal systems were also found, among them: oxytocin (bta04921) and glucagon signaling (bta04922). *RCAN1* is a member of the oxytocin signaling pathway (see Supplementary Tables: Table S3) which is essential for the regeneration of skeletal muscle tissue and maintenance of homeostasis in mice⁴⁴. Glucagon signals liver and muscle cells to change stored glycogen into glucose which is released into the bloodstream and used by other cells for energy⁴⁵. A continuous turnover of protein (synthesis and breakdown) maintains the functional integrity and quality of skeletal muscle. Hormones are important regulators of this remodeling process²⁵.

In contrast to muscle growth, muscle loss is a debilitating consequence of a range of pathologic and metabolic conditions, for example, an increase in protein degradation. We found processes (protein polyubiquitination - GO:0000209, protein ubiquitination involved in ubiquitin-dependent protein catabolic process - GO:0042787 and ubiquitin mediated proteolysis - bta04120) related to proteolytic systems some of which are involved in muscle degradation⁴⁶. Among these, the ubiquitin-proteasome system is highly conserved across vertebrates, degrades the major proteins of contractile skeletal muscle and plays an important role in muscle loss¹⁷.

Genes encoding differentially expressed alternatively spliced transcripts associated with REA were also related to immune function, such as the Adenylate Cyclase 2 (*ADCY2*) gene which also is annotated with several other GO terms and is a member of other pathways including those related to the neural and hormonal systems. There is a complex interaction between the skeletal muscle and the immune systems that regulates muscle regeneration⁴⁷. *ADCY2* is a membrane-associated enzyme and catalyzes the formation of the secondary messenger cyclic adenosine monophosphate⁴⁸. An alternative 3' splice site event was found for this gene which is a member of the cAMP signaling pathway (bta04024). *ADCY2* may be involved in multiple physiological

processes, and has previously been associated with meat tenderness in Yanbian cattle⁴⁸.

Intramuscular fat content

We found differentially expressed alternatively spliced transcripts belonging to members of the myosin (*MYH7B*, *MYLK2* and *MYO18A*), myotilin (*MYOT*), myozenin (*MYOZ1* and *MYOZ3*) and ubiquitin (*UBAC2* and *UBE3C*) families. A factor that may affect intramuscular fat deposition is the skeletal muscle fiber metabolism. The extent of intramuscular fat is positively correlated with the percentage of oxidative muscle fiber⁴⁹. When the rate of protein deposition decreases (possibly due to ubiquitination processes) lipid deposition becomes the major component of weight gain and the energy request to fat tissues increases⁵⁰.

TRIP12, *ANXA11* and *MAP2K6* which have functions related to cell regulation and maintenance^{22,23,24} were predicted to be hub genes associated with IF content. *TRIP12* and *ANXA11* encode transcripts with novel alternative 3' splice sites, which were more abundant in the HIF animals. *MAP2K6* encodes a transcript with a known alternative first exon which was more abundant in the LIF animals (see Supplementary Tables: Table S2). Alternative 3'-splice site events found in the hub genes were predominantly expressed in the HIF animals. The function that these hub genes and alternatively spliced transcripts may have on lipid metabolism is unknown.

TRIP12 is a protein coding gene that is a key regulator of the response to DNA damage. Among its related pathways are Class I MHC mediated antigen processing and presentation, and ubiquitin mediated proteolysis²². *TRIP12* was found to be expressed in the skeletal muscle of Holstein-Friesian, Limousin, Hereford and Polish Red bulls⁵¹. In humans, *TRIP12* plays a role in the ubiquitin fusion degradation pathway, which is a proteolytic system that is conserved in mammals⁵².

ANXA11 encodes a member of the annexin family, a group of calcium-dependent phospholipid-binding proteins²³. *ANXA11* isoforms play receptor roles that regulate membrane-related events, such as exocytosis and endocytosis⁵³. *MAP2K6* is a member of the dual specificity protein kinase family and plays a role in the regulation of the mitogen-activated protein kinase pathway²⁴. Single

nucleotide polymorphisms in *MAP2K6* have previously been associated with carcass weight, back fat thickness and marbling score in Hanwoo cattle⁵⁴.

The enrichment analysis suggested that the most relevant function of the genes with significant isoform expression shifts for IF was related to the growth, regulation and proteolysis of muscle cells (Supplementary Tables: Table S4), for example: myofibril assembly (GO:0030239); skeletal muscle cell differentiation (GO:0035914) and ubiquitin-dependent protein catabolic process (GO:0006511). This suggests that variation in IF content may be due to a balance between the absorption, synthesis and degradation of intramyocellular and extramyocellular lipids, which involves many metabolic pathways in myofibers⁵⁵. Intramuscular fat accumulates both within (intramyocellular) and externally (extramyocellular) of muscle fibers⁵⁶. The intramyocellular lipids are stored as spherical droplets in muscle cells and are reportedly associated with aerobic metabolism, whereas the extramyocellular lipids are located in long fatty septa of laminar shape, along muscle fiber bundles. Intramyocellular lipids play an important role in muscle metabolism⁵⁷.

The PI3K-Akt (bta04151), mTOR signaling (bta04150) and negative regulation of TORC1 signaling (GO:1904262) pathways were overrepresented in this study. *EIF4B* which is a member of these pathways, produces a transcript with a novel alternative 3'-splice site, that was found to be more abundant in the LIF animals. *EIF4B* is a protein coding gene that is required for the binding of mRNA to ribosomes. The activity of EIF4 family genes is controlled by mTOR serine threonine kinase⁵⁸ which forms the mTORC1 and mTORC2 complexes. The role of mTORC1 in regulating lipid synthesis, which is required for cell growth and proliferation, is beginning to be appreciated⁵⁹. It has been demonstrated that mTORC1 positively regulates the transcription factors (sterol regulatory element binding protein 1 - *SREBP1* and peroxisome proliferator-activated receptor- γ - *PPARG* γ) that control the expression of genes encoding proteins involved in lipid and cholesterol homeostasis⁶⁰. The PI3K-AKT pathway is an intracellular signaling pathway important in regulating the cell cycle. PI3K activation phosphorylates and activates AKT. AKT can have a number of downstream effects such as activating mTOR⁶¹.

We found candidate genes that encode differentially expressed alternatively spliced transcripts associated with IF content in Nelore cattle.

Previous studies have identified differentially expressed genes in Simmental-Luxi cross⁶², Wagyu and Angus⁶³, Wagyu and Holstein⁶⁴ and Nelore⁶⁵, suggesting that some of these genes regulate lipid composition and deposition in intramuscular fat^{62,63,64}. These findings and our results enable a better understanding of the mechanisms underlying the gene transcription associated with IF content in Nelore cattle.

Conclusion

In summary, RNA-Seq was used to identify candidate genes that encode alternatively spliced transcripts that were associated with REA and IF content in Nelore bulls. The results suggest that REA and IF content may be influenced by sets of genes that encode alternatively spliced transcripts (both known and novel transcripts). For both traits, we found genes with biological process GO terms that are involved in pathways related to protein ubiquitination, muscle differentiation, regulation of lipolysis in adipocytes, and hormonal systems. These results provide a foundation for further research into the specific functions of candidate genes encoding alternatively spliced transcripts that are differentially expressed in animals differing for REA and IF content.

Methods

Animals and phenotypes

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). The intact Nelore bulls used in this study (N=80) were from the Capivara Farm, which participates in the Nelore Qualitas Breeding Program. The animals from a single contemporary group (that remained together from birth until slaughter) were reared on *Brachiaria* sp. and *Panicum* sp. forages, and had free access to mineral salt and were finished in a feedlot for approximately 90 days. The animals were slaughtered at an average age of 24 months in a commercial slaughterhouse, in accordance with Brazilian Federal Inspection Service procedures.

Samples from the *longissimus thoracis* muscle were collected, between the 12th and 13th ribs of the left half of each carcass, at two time points:

immediately following slaughter for the RNA sequencing analysis (details previously described by Fonseca et al.³⁵) and again at 24 hours post-slaughter for the evaluation of REA and determination of IF content. REA was measured by the grid method in units of centimeters squared⁶⁶. The Bligh and Dyer⁶⁷ method was used to determine IF content. From the 80 animals, groups of animals were selected for RNA Seq analysis of ribeye muscle tissue: 1) REA, the 15 animals with the highest and 15 with the lowest REA (HREA and LREA); and 2) IF, 15 animals with the highest and 15 with the lowest IF content (HIF and LIF). A Student's t-test was applied to determine if there were statistically significant differences between the selected groups (Table 1).

Phenotype	N	Mean \pm standard deviation	Minimum	Maximum	p-value
HREA (cm ²)	15	83.66 \pm 2.55	81	88	0.01
LREA (cm ²)	15	65.73 \pm 2.05	61	68	
HIF*	15	0.101 \pm 0.0095	0.094	0.126	0.05
LIF*	15	0.063 \pm 0.0049	0.051	0.067	

Table 1. Descriptive statistics for ribeye muscle area and intramuscular fat content of Nelore cattle. *The % data were transformed using the arcsine square root function. HREA = High ribeye muscle area; LREA = Low ribeye muscle area; HIF = High intramuscular fat content; LIF = Low intramuscular fat content; N = Number of animals.

RNA Seq library preparation and data processing

Total RNA was extracted from the samples (an average of 50 mg of the muscle tissue previously stored in RNA holder, BioAgency, SP, Brazil) using the RNeasy Lipid Tissue Mini Kit (Qiagen, Valencia, 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). The quality of the RNA samples was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and the RNA concentration and genomic DNA contamination were measured using a Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA).

High-quality total RNA (approximately 500 ng) was processed using a TruSeq RNA Sample Preparation Kit® (Illumina, San Diego, CA) according to the manufacturer's protocol. Differently bar-coded libraries were pooled to enable

multiplexed sequencing and generated, on average, approximately 25M read-pairs per sample. RNA sequencing (RNA-Seq) was performed using a HiSeq 2500 System (Illumina®) that generated 100 bp paired-end reads.

Quality of RNA-Seq reads (quality scores, GC content, N content, length distributions, duplication levels, overrepresented sequences and K-mer content) was checked using FastQC (v.0.11.4) software⁶⁸. Trimmed read data were obtained by removing reads containing low quality and adapter sequence from the raw data using Trimmomatic v.0.36⁶⁹ following parameters: cut of adapter, other illumina-specific sequences from the read, bases off the start of a read, bases off the end of a read; read to a specified length by removing bases from the end, sliding window trimming, drop the read if the average quality is below the specified level; and the read if it is below a specified length.

All downstream analyses were performed on the trimmed data with high quality reads. HISAT2 v.2.0.5⁷⁰ was used to align the paired-end trimmed data to the bovine reference genome (UMD3.1.1 *Bos taurus*) and chrY (*Btau* 4.6.1), both deposited in National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/>). Table 2 shows the descriptive statistics for the mean alignment rates for the samples by trait (REA and IF).

Animals group	Mean of raw reads (Mb)	Mean trimmed reads (Mb)	Overall alignment rate (%)	Uniquely mapped (%)	Coverage
REA (N=30)	25,798,305	22,468,708	96.5	88	45X
IF (N=30)	24,911,474	21,976,535	96.4		

Table 2. Descriptive statistics for the overall alignment percentage for RNA-Seq data from Nelore cattle samples. REA = Ribeye muscle area; IF = Intramuscular fat content; Mb= Megabase pair

Identification of differentially expressed alternatively spliced transcripts

JuncBASEv.0.9 software (Junction-Based Analysis of Splicing Events)³ was used to identify and classify exon-centric alternative splicing events (cassette exons, alternative 5' splice site, alternative 3' splice site, mutually exclusive exons, coordinate cassette exons, alternative first exons, alternative

last exons and intron retention) based on splice junction reads predicted using the Cufflinks and Cuffmerge tools⁷¹.

The Percentage Spliced Index (PSI) was estimated as the ratio of the number of reads including exons to the sum of the numbers of reads including or excluding exons³. The PSI value was calculated and Fisher's exact test was used to test for differences in alternative splicing events between the high and low REA and IF groups ($p < 0.05$). Only alternatively spliced transcripts that were supported by at least 10 reads and that had PSI differences (Δ PSI) between the respective high and low groups of greater than 5% were considered significant.

Co-expression networks analyses

For co-expression network analyses, the ExpressionCorrelation plugin⁷² within Cytoscape v.3.4⁷³ was used to identify consensus networks of gene sets in which differentially expressed alternatively spliced transcripts were associated with REA or IF content. Similarity matrices were computed using the Pearson correlations among PSI values. A histogram tool was used for choosing a similarity strength threshold (cut-off of > 0.75 or < -0.75), to enable the creation of reasonably sized networks. The Network Analyzer plugin⁷⁴ in Cytoscape software, was used to obtain the network metrics (clustering and density coefficient). The network clustering coefficient is the average of the clustering coefficients for all nodes in the network. The network density is an average numbers of neighbors a normalized value that indicates the average connectivity of a node in the network. The clustering and density coefficients are values between 0 and 1¹⁸. The genes that were engaged in the most interactions with other genes, defined as maximal clique centrality (MCC), were identified as hub genes using the cytoHubba plugin⁷⁵ in Cytoscape software.

Gene set enrichment analyses

The gene sets in which differentially expressed alternatively spliced transcripts associated with REA or IF content were identified, and then the enrichment and pathway analyses of these gene sets were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID 6.8)⁷⁶. Fisher's exact test was used and p values were adjusted for multiple testing using the Benjamini-Hochberg method⁷⁷. DAVID Pathway was used to map the

gene enriched pathways using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database⁷⁸.

References

1. Lerch, J. K. et al. Isoform Diversity and Regulation in Peripheral and Central Neurons Revealed through RNA-Seq. *PLoS One*. 7, e30417 (2012).
2. 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*. 11(7), e0159719. (2016).
3. Brooks, A. N. et al. Estimates of genetic parameters for carcass growth and reproductive traits in Nellore cattle. *Livest Sci*. 155, 1–7 (2013).
4. Guan, Y., Liang, G., Martin, G. B. & Guan, L. L. Functional changes in mRNA expression and alternative pre-mRNA splicing associated with the effects of nutrition on apoptosis and spermatogenesis in the adult testis. *BMC Genomics*. 18, 64 (2017).
5. Nakka, K., Ghigna, C., Gabellini, D. & Dilworth, F. J. Diversification of the muscle proteome through alternative splicing. *Skelet Muscle*. 8, 8 (2018).
6. Blencowe, B. J. Alternative splicing: new insights from global analyses. *Cell*. 126, 37-47 (2006).
7. Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature*. 463, 457-63 (2010).
8. Sammeth, M., Foissac, S. & Guigó, R. A general definition and nomenclature for alternative splicing events. *Plos Comp Biol*. 4:e1000147 (2008).
9. Potenza, E. et al. Exploration of alternative splicing events in ten different grapevine cultivars. *BMC Genomics*. 16, 706 (2015).
10. Liang, G. et al. Altered microRNA expression and pre-mRNA splicing events reveal new mechanisms associated with early stage *Mycobacterium avium* subspecies paratuberculosis infection. *Sci Rep*. 6, 24964 (2016).
11. Huang, W. & Khatib, H. Comparison of transcriptomic landscapes of bovine embryos using RNA-Seq. *BMC Genomics*. 11, 1-10 (2010).
12. He, H. & Liu, X. Characterization of transcriptional complexity during longissimus muscle development in bovines using high-throughput sequencing. *PLoS One*. 8, e64356 (2013).
13. Scholz, A. M., Bünger, L., Kongsro, J., Baulain, U. & Mitchell, A. D. Non-invasive methods for the determination of body and carcass composition in livestock: dual-energy X-ray absorptiometry, computed tomography, magnetic resonance imaging and ultrasound: invited review. *Animal*. 9, 1250–1264 (2015).
14. Shingfield, K. J., Bonnet, M. & Scollan, N. D. Recent developments in altering the fatty acid composition of ruminant-derived foods. *Animal*. 7, 132–162 (2013).
15. Pérez de Heredia, F., Wood, I. S. & Trayhurn, P. Hypoxia stimulates lactate release and modulates monocarboxylate transporter (MCT1, MCT2, and MCT4) expression in human adipocytes. *Pflugers Arch*. 459, 509–518 (2010).
16. Klein, J. et al. Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics*. 4, 217–9 (1990).
17. Attaix, D. et al. The ubiquitin-proteasome system and skeletal muscle wasting. *Essays Biochem*. 41, 173-186 (2005).
18. Dong, J. & Horvath, S. Understanding network concepts in modules. *BMC Syst Biol* 1, 24 (2007).
19. Mankodi, A. et al. Ribonuclear inclusions in skeletal muscle in myotonic dystrophy types 1 and 2. *Ann Neurol*. 54, 760-8 (2003).
20. Zhang, L. & Michal, J. J. Quantitative genomics of 30 complex phenotypes in Wagyu x Angus F1 progeny. *Int J Biol Sci*. 8, 838–858 (2012).

21. Ramos, S., Khademi, F., Somesh, B. P. & Rivero, F. Genomic organization and expression profile of the small GTPases of the RhoBTB family in human and mouse. *Gene*. 298, 147-57 (2002).
22. Gudjonsson, T. et al. TRIP12 and UBR5 suppress spreading of chromatin ubiquitylation at damaged chromosomes. *Cell*. 150, 697-709 (2012).
23. Farnaes, L. & Ditzel, H. J. Dissecting the cellular functions of annexin XI using recombinant human annexin XI-specific autoantibodies cloned by phage display. *J Biol Chem*. 278, 33120-33126 (2003).
24. Remy, G. et al. Differential activation of p38MAPK isoforms by MKK6 and MKK3. *Cell Signal*. 22, 660–667 (2010).
25. Rooyackers, O. E. & Nair, K. S. Hormonal regulation of human muscle protein metabolism. *Annu Rev Nutr*. 17, 457-85 (1997).
26. Bland, C. S. et al. Global regulation of alternative splicing during myogenic differentiation. *Nucleic Acids Res*. 38, 7651-7664 (2010).
27. Langfelder, P., Mischel, P. S. & Horvath, S. When is hub gene selection better than standard meta-analysis? *PLoS One*. 8, e61505 (2013).
28. Akita, T. et al. De novo variants in CAMK2A and CAMK2B cause neuro developmental disorders. *Ann Clin Trans Neurol*. 5, 280-296 (2018).
29. Martinsen, A., Dessy, C. & Morel, N. Regulation of calcium channels in smooth muscle: New insights into the role of myosin light chain kinase. *Channels*. 8:402-413 (2014).
30. Collins, S. P., Reoma, J. L., Gamm, D. M. & Uhler, M. D. LKB1, a novel serine/threonine protein kinase and potential tumour suppressor, is phosphorylated by cAMP-dependent protein kinase (PKA) and prenylated in vivo. *Biochem J*. 345, 673–80 (2000).
31. Salmikangas, P. et al. O.Myotilin a novel sarcomeric protein with two Ig-like domains, is encoded by a candidate gene for limb-girdle muscular dystrophy. *Hum Mol Genet*. 8(7), 1329-36 (1999).
32. Van der Ven, P. F. et al. Assignment of the human gene for endosarcomeric cytoskeletal M-protein (MYOM2) to 8p23.3. *Genomics* 55(2), 253-255 (1999).
33. Sgarbieri, V. C. Proteínas em alimentos protéicos (ed. Sgarbieri, V. C) 518 (Varela, 1996).
34. Fernandes Júnior, G. A. et al. Genome scan for postmortem carcass traits in Nellore cattle. *J Anim Sci*. 94(10), 4087-4095 (2016).
35. Fonseca, L. F. S. et al. Differences in global gene expression in muscle tissue of Nellore cattle with divergent meat tenderness. *BMC Genomics*. 18, 1-12 (2017).
36. Batra, T. R., Lee, A. J., Gavora, J. S. & Stear, M. J. CLASS I alleles of the bovine major histocompatibility system and their association with economic traits. *J Dairy Sci*. 72, 2115–2124 (1989).
37. Stear, M. J., Pokorny, T. S., Muggli, N. E. & Stone, R. T. The relationships of birth weight, preweaning gain and postweaning gain with the bovine major histocompatibility system. *J Anim Sci*. 67, 641–9 (1989).
38. Stupka, N. et al. Activated calcineurin ameliorates contraction-induced injury to skeletal muscles of mdx dystrophic mice. *J Physiol*. 575, 645-56 (2006).
39. Hudson, M. B & Price, S. R. Calcineurin: A poorly understood regulator of muscle mass. *Int J Biochem Cell Biol*. 45, 2173-2178 (2013).
40. Berdeaux, R. & Stewart, R. cAMP signaling in skeletal muscle adaptation: hypertrophy, metabolism, and regeneration. *Am J Physiol Endocrinol Metab*. 303, E1-E17 (2012).
41. Chen, A. E., Ginty, D. D. & Fan, C. M. Protein kinase A signalling via CREB controls myogenesis induced by Wnt proteins. *Nature*. 433, 317-22 (2005).
42. Jiang, C. et al. Gene expression profiling of skeletal muscle of nursing piglets. *Int J Biol Sci*. 6, 627-38 (2010).

43. Quach, N. L., Biressi, S., Reichardt, L. F., Keller, C. & Rando, T. A. Focal adhesion kinase signaling regulates the expression of Caveolin 3 and β 1 Integrin, genes essential for normal myoblast fusion. *Mol Biol Cell*. 20, 3422-3435 (2009).
44. Elabd, C. et al. Oxytocin is an age-specific circulating hormone that is necessary for muscle maintenance and regeneration. *Nat Comm*. 5, 4082 (2014).
45. Holst, J. J. et al. Insulin and Glucagon: Partners for Life. *Endocrinology*. 158, 696-701 (2017).
46. Purintrapiban, J., Wang, M. C. & Forsberg, N. E. Degradation of sarcomeric and cytoskeletal proteins in cultured skeletal muscle cells. *Comp Biochem Physio B Biochem Mol Biol*. 136, 393-401 (2003).
47. Tidball, J. G. & Villalta, S. A. Regulatory interactions between muscle and the immune system during muscle regeneration. *Am J Physiol Regul Integr Comp Physiol*. 298, R1173-87. (2010).
48. Li, Y. X. et al. Molecular cloning, sequence identification, and gene expression analysis of bovine ADCY2 gene. *Mol Biol Rep*. 41, 3561-8 (2014).
49. Veloso, R. C. et al. Expression of lipid metabolism and myosin heavy chain genes in pigs is affected by genotype and dietary lysine. *Gen Mol Res*. 17(2), gmr16039904 (2018).
50. Hocquette, J. F. et al. Energy metabolism in skeletal muscle of meat-producing animals. *INRA Prod. Anim*. 13, 185-200 (2000).
51. Sadkowski, T., Jank, M., Zwierzchowski, L., Oprzadek, J. & Motyl, T. Transcriptomic index of skeletal muscle of beef breeds bulls. *J Physiol Pharm* 60:15–28 (2009).
52. Park, Y., Yoon, S. K. & Yoon, J. B. The HECT domain of TRIP12 ubiquitinates substrates of the ubiquitin fusion degradation pathway. *J Biol Chem*. 284, 1540-9 (2009).
53. Rescher, U. & Gerke, V. Annexins—unique membrane binding proteins with diverse functions. *J Cell Sci*. 117, 2631–2639 2003.
54. Ryu, J., Kim, Y., Kim, C., Kim, J. & Lee, C. Association of bovine carcass phenotypes with genes in an adaptive thermogenesis pathway. *Mol Biol Rep* 39, 1441-5 (2012).
55. Hocquette, J. F. et al. Intramuscular fat content in meat-producing animals: Development, genetic and nutritional control, and identification of putative markers. *Animal*. 4, 303-19 (2010).
56. Rahemi, H., Nigam, N. & Wakeling, J. M. The effect of intramuscular fat on skeletal muscle mechanics: implications for the elderly and obese. *J R Soc Interface*. 12, 20150365 (2015).
57. Gambarota, G., Janiczek, R. L., Mulkern, R. V. & Newbould, R. D. An NMR phantom mimicking intramyocellular (IMCL) and extramyocellular lipids (EMCL). *Appl Magn Reson* 43, 451–457 (2012).
58. Gingras, A. C. et al. Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev*. 13, 1422-1437 (1999).
59. Laplante, M. & Sabatini, D. M. mTOR signaling at a glance. *J Cell Sci* 122:3589-3594 (2009).
60. Porstmann, T. et al. SREBP activity is regulated by mTORC1 and contributes to Akt-dependent cell growth. *Cell Metab*. 8, 224-236 (2008).
61. Rafalski, V. A. & Brunet, A. Energy metabolism in adult neural stem cell fate. *Prog Neurobiol*. 93, 182–203. (2011).
62. Sheng, X. et al. RNA-seq analysis of bovine intramuscular, subcutaneous and perirenal adipose tissues. *Mol Biol Rep*. 41, 1631-7 (2014).
63. Wei, S. et al. Enhanced mitogenesis in stromal vascular cells derived from subcutaneous adipose tissue of Wagyu compared with those of Angus cattle. *J Anim Sci*. 93, 1015-24 (2015).

64. Huang, W. et al. Global transcriptome analysis identifies differentially expressed genes related to lipid metabolism in Wagyu and Holstein cattle. *Sci Rep.* 7, 5278 (2017).
65. Cesar, A. S. et al. Putative regulatory factors associated with intramuscular fat content. *PLoS One* 10, 1-21 (2015).
66. United States Department of Agriculture – USDA. *Official United States standards for grades of carcass beef*. Washington, D.C.: Agricultural Marketing Service. (1997).
67. Bligh, E. G & Dyer, W. J. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol.* 37, 911–917 (1959).
68. Andrews, S. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (2010).
69. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics.* 30, 2114-20 (2014).
70. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: A fast spliced aligner with low memory requirements. *Nat Meth.* 12, 357–360 (2015).
71. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks. *Nat Protoc.* 7, 562-578 (2012).
72. Hui, S. et al. ExpressionCorrelation. <http://apps.cytoscape.org/apps/expressioncorrelation> (2015).
73. Shannon, P. et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498-504. (2003).
74. Assenov, Y., Ramírez, F., Schelhorn, S. E., Lengauer, T. & Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* 24, 282-4 (2008).
75. Chin, CH. et al. CytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol.* 8(4), S11 (2014).
76. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 4, 44-57 (2009).
77. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc.* 57, 1-12 (1995).
78. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27-30. (2000).

Acknowledgements

The authors thank the Qualitas Nelore breeding program company for providing the tissue samples and database used in this study. 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 (grants #2009/16118-5 #2015/16850-9 and #2016/22894-1).

Additional Files

Supplementary Tables – S1, S2, S3 and S4

Table S1. Alternatively spliced and differentially expressed ($p < 0.05$, t-test) transcripts identified in the ribeye muscle of Nelore bulls with the highest and lowest ribeye muscle area animals (HREA and LREA). Green highlighted: hub genes

Annotated Junctions (Known and Novel)	Transcripts type	Gene	Chr	Strand	HREA	LREA	Delta (LREA-HREA)	P value
K	cassette exon	<i>ACAA2</i>	chr24	-	100	94.74	-5.26	0.014
N	alternative 3' splice site	<i>ADCY2</i>	chr20	-	78.26	84.21	5.95	0.033
N	alternative 3' splice site	<i>ADCY2</i>	chr20	-	17.46	15.79	-1.67	0.043
K	alternative 5' splice site	<i>ADD1</i>	chr6	-	10	5.26	-4.74	0.048
K	intron retention	<i>ADD1</i>	chr6	-	15.36	8	-7.36	0.038
N	alternative 3' splice site	<i>AKAP1</i>	chr19	+	90.45	93.54	3.08	0.036
N	alternative 3' splice site	<i>ANAPC11</i>	chr19	-	3.45	2.47	-0.98	0.025
K	cassette exon	<i>APLP2</i>	chr29	+	79.37	84.62	5.25	0.022
N	alternative 3' splice site	<i>ASPH</i>	chr14	-	79.26	70.83	-8.43	0.013
N	alternative 3' splice site	<i>BOLA-DOA</i>	chr23	+	97.56	91.45	-6.11	0.023
N	alternative 5' splice site	<i>C23H6or</i>	chr23	-	71.93	64.24	-7.69	0.018
N	alternative 5' splice site	<i>C23H6or</i>	chr23	-	27.63	35.76	8.13	0.023
N	coordinate cassette exon	<i>C23H6or</i>	chr23	-	86.21	81.13	-5.08	0.041
N	cassette exon	<i>C3H1orf</i>	chr3	+	96.36	87.5	-8.86	0.007
N	cassette exon	<i>C3H1orf</i>	chr3	+	100	90.48	-9.52	0.033
N	coordinate cassette exon	<i>C3H1orf</i>	chr3	+	100	91.67	-8.33	0.010
K	cassette exon	<i>C8H9orf</i>	chr8	+	33.69	28.33	-5.36	0.026
N	cassette exon	<i>CAMK2B</i>	chr4	+	76.92	64.19	-12.73	0.005
N	alternative 3' splice site	<i>CARM1</i>	chr7	+	91.45	100	8.55	0.007
N	alternative 3' splice site	<i>CCDC82</i>	chr15	+	6.67	2.94	-3.73	0.029
N	cassette exon	<i>CD36</i>	chr4	-	1.06	3.03	1.97	0.032

N	alternative 5' splice site	<i>CDC34</i>	chr7	+	14.6	10.64	-3.96	0.018
N	alternative 5' splice site	<i>CDC35</i>	chr7	+	76.44	79.41	2.97	0.042
K	alternative 3' splice site	<i>CDV3</i>	chr1	-	17.5	22.78	5.28	0.026
N	alternative 3' splice site	<i>CMYA5</i>	chr10	+	95.84	98.18	2.34	0.022
K	alternative 5' splice site	<i>COPS7A</i>	chr5	-	11.11	15.05	3.95	0.027
K	cassette exon	<i>CSDE1</i>	chr3	+	98.29	97.18	-1.11	0.030
K	alternative last exon	<i>CXCL12</i>	chr28	+	10.99	13.79	2.8	0.016
N	alternative 5' splice site	<i>DBNDD2</i>	chr13	+	34.84	42.47	7.62	0.043
N	alternative 5' splice site	<i>DBNDD2</i>	chr13	+	62.83	56.23	-6.59	0.044
N	alternative 3' splice site	<i>DDX39A</i>	chr7	+	11.76	19.44	7.68	0.035
N	cassette exon	<i>DR1</i>	chr3	-	96.77	99.08	2.31	0.049
N	intron retention	<i>ECHDC3</i>	chr13	-	91.65	95.06	3.41	0.044
N	cassette exon	<i>EGLN1</i>	chr28	+	98.82	96.38	-2.44	0.002
K	alternative 3' splice site	<i>EIF4G3</i>	chr2	+	27.03	29.41	2.38	0.010
N	alternative 5' splice site	<i>EPS15</i>	chr3	+	61.33	51.52	-9.81	0.004
K	alternative first exon	<i>FAM195A</i>	chr25	+	97.78	100	2.22	0.010
N	cassette exon	<i>FAM96B</i>	chr18	-	77.78	81.97	4.19	0.014
N	alternative 3' splice site	<i>FLNC</i>	chr4	+	7.02	4.18	-2.84	0.016
N	alternative 5' splice site	<i>FLNC</i>	chr4	+	5.57	4.53	-1.04	0.031
N	alternative 5' splice site	<i>FLNC</i>	chr4	+	93.74	94.87	1.13	0.037
K	cassette exon	<i>FXR1</i>	chr1	-	100	97.45	-2.55	0.033
K	coordinate cassette exon	<i>FXR1</i>	chr1	-	100	97.89	-2.11	0.012
N	cassette exon	<i>FXYD1</i>	chr18	+	89.8	84.08	-5.72	0.001
N	cassette exon	<i>FXYD1</i>	chr18	+	91.94	87.62	-4.32	0.002
N	alternative 5' splice site	<i>GNL3L</i>	chrX	+	59.18	45.65	-13.53	0.031
K	alternative 5' splice site	<i>HECTD1</i>	chr21	-	64.11	70.45	6.34	0.005
K	cassette exon	<i>HNRNPD</i>	chr6	-	3.03	1.47	-1.56	0.024
K	alternative 5' splice site	<i>HNRNPH3</i>	chr28	+	64.29	59.68	-4.61	0.024

K	cassette exon	<i>HRASLS</i>	chr1	-	2.44	6.76	4.32	0.046
N	alternative 5' splice site	<i>HSDL2</i>	chr8	+	4.76	1.92	-2.84	0.009
N	alternative 3' splice site	<i>HSPE1</i>	chr2	+	1.08	2.22	1.14	0.011
N	alternative 3' splice site	<i>ITGA7</i>	chr5	+	37.64	28.33	-9.31	0.023
N	alternative 5' splice site	<i>ITGA8</i>	chr5	+	38.42	30	-8.42	0.021
N	cassette exon	<i>KIAA0232</i>	chr6	+	22.07	30.7	8.63	0.035
N	alternative 5' splice site	<i>KIF5B</i>	chr13	+	83.02	87.16	4.14	0.027
K	alternative 3' splice site	<i>LIMCH1</i>	chr6	+	14.67	11.76	-2.91	0.031
N	alternative 3' splice site	<i>LOC1008</i>	chr18	+	48.5	52.32	3.82	0.049
N	alternative 5' splice site	<i>LOC1019</i>	chr4	+	0.35	2.4	2.05	0.044
N	alternative 5' splice site	<i>LOC1049</i>	chr14	+	79.31	82.61	3.3	0.049
N	alternative 5' splice site	<i>LOC104970651</i>	chr21	+	100	95.12	-4.88	0.025
K	alternative last exon	<i>LOC104972733</i>	chr6	+	1.61	0.7	-0.91	0.049
N	alternative 5' splice site	<i>LOC7865</i>	chr7	-	42.11	36.21	-5.9	0.020
K	alternative first exon	<i>LRR2</i>	chr22	+	0.75	2.67	1.92	0.047
K	alternative last exon	<i>LRRFIP1</i>	chr3	+	8.33	10.98	2.65	0.042
K	cassette exon	<i>LRRFIP1</i>	chr3	+	93.62	89.92	-3.7	0.013
K	coordinate cassette exon	<i>LRRFIP1</i>	chr3	+	64.91	81.82	16.91	0.023
N	alternative 3' splice site	<i>LRRFIP2</i>	chr22	+	95.19	100	4.81	0.034
K	cassette exon	<i>LRRFIP2</i>	chr22	+	81.98	100	18.02	0.018
N	cassette exon	<i>LRRFIP2</i>	chr22	+	92.86	100	7.14	0.043
N	cassette exon	<i>MFF</i>	chr2	+	84.06	79.13	-4.93	0.019
N	alternative 5' splice site	<i>MGLL</i>	chr22	+	23.2	21.1	-1.1	0.017
N	alternative 5' splice site	<i>MGLL</i>	chr22	+	45.2	50.5	5.3	0.025
N	alternative 3' splice site	<i>MGST2</i>	chr17	-	4.05	6.25	2.2	0.050
N	cassette exon	<i>MIF4GD</i>	chr19	+	73.53	80	6.47	0.032
N	cassette exon	<i>MIF4GD</i>	chr19	+	79.7	69.7	-10	0.038
N	cassette exon	<i>MLIP</i>	chr23	-	75	70.75	-4.25	0.021

N	cassette exon	<i>MLIP</i>	chr23	-	75	70.55	-4.45	0.024
N	coordinate cassette exon	<i>MLIP</i>	chr23	-	85.96	80.39	-5.57	0.029
N	alternative 5' splice site	<i>MPRIP</i>	chr19	-	11.63	10.57	-1.06	0.016
K	alternative 3' splice site	<i>MRPL55</i>	chr7	+	69.46	76.19	6.73	0.039
K	cassette exon	<i>MYBPC1</i>	chr5	+	96.34	84.55	-11.79	0.007
K	cassette exon	<i>MYBPC1</i>	chr5	+	20.88	28.65	7.77	0.015
K	cassette exon	<i>MYBPC1</i>	chr5	+	29.22	37.44	8.22	0.016
N	intron retention	<i>MYH1</i>	chr19	+	91.53	96.09	4.56	0.014
N	intron retention	<i>MYH1</i>	chr19	-	100	98	2	0.030
N	alternative 3' splice site	<i>MYH1</i>	chr19	-	3.34	2.7	-0.64	0.009
N	alternative 3' splice site	<i>MYL3</i>	chr22	+	96.78	100	3.22	0.046
N	alternative 5' splice site	<i>MYOM2</i>	chr27	+	98	95.5	-2.5	0.043
N	alternative 5' splice site	<i>MYOM2</i>	chr27	+	100	93.5	-6.5	0.048
N	cassette exon	<i>MYOM3</i>	chr2	+	100	98.91	-1.09	0.042
N	alternative 5' splice site	<i>MYOT</i>	chr7	+	100	92	-8	0.020
N	alternative 5' splice site	<i>MYOT</i>	chr7	+	2	8	6	0.021
N	alternative 5' splice site	<i>NDRG2</i>	chr10	+	92.86	93.06	0.2	0.044
N	alternative 5' splice site	<i>NDRG2</i>	chr10	+	0.96	0.38	-0.58	0.040
N	alternative 3' splice site	<i>NEDD8</i>	chr10	+	3.17	2.17	-1	0.020
K	alternative first exon	<i>NTMT1</i>	chr11	+	3.43	5.9	-2.47	0.016
K	coordinate cassette exon	<i>OBSCN</i>	chr7	-	10.39	9.36	-1.03	0.043
N	alternative 5' splice site	<i>OBSCN</i>	chr7	-	35.53	39.86	4.33	0.033
N	intron retention	<i>OBSCN</i>	chr7	-	100	81.82	-18.18	0.007
K	cassette exon	<i>OGDH</i>	chr4	-	23.08	29.94	6.86	0.048
N	alternative 5' splice site	<i>PARK7</i>	chr16	-	90.32	93.24	2.92	0.010
N	alternative 5' splice site	<i>PARK7</i>	chr16	-	8.62	6.76	-1.86	0.014
K	cassette exon	<i>PHLDB1</i>	chr15	+	6.25	12.4	6.15	0.031
N	cassette exon	<i>PICALM</i>	chr29	+	58.23	48.13	-10.1	0.048

K	alternative first exon	<i>PIK3R1</i>	chr20	-	11.11	3.98	-7.13	0.045
K	cassette exon	<i>PITPNA</i>	chr19	-	33.33	20.69	-12.64	0.034
N	alternative 3' splice site	<i>PLCD4</i>	chr2	+	17.14	11.11	-6.03	0.007
N	alternative 3' splice site	<i>PLCD4</i>	chr2	+	69.81	77.78	7.97	0.034
K	cassette exon	<i>POLR2E</i>	chr7	-	92	95.42	3.42	0.025
K	cassette exon	<i>PPP3CC</i>	chr8	+	83.97	77.78	-6.19	0.007
K	coordinate cassette exon	<i>PTPRM</i>	chr24	+	7.73	12.9	5.17	0.041
K	alternative first exon	<i>RCAN1</i>	chr1	+	87.5	96.68	9.18	0.017
N	alternative 3' splice site	<i>RHOBTB1</i>	chr28	-	88.65	98.44	9.79	0.027
N	cassette exon	<i>RNF115</i>	chr3	+	100	95	-5	0.008
N	alternative 3' splice site	<i>RNF157</i>	chr19	+	98.63	100	1.37	0.050
K	alternative 3' splice site	<i>RNF181</i>	chr11	-	7.55	10.13	2.58	0.005
N	alternative 3' splice site	<i>RPS4X</i>	chrX	+	96.81	95.26	-1.55	0.018
N	cassette exon	<i>RRAGD</i>	chr9	+	91.3	87.09	-4.21	0.020
N	cassette exon	<i>RSRC2</i>	chr17	+	6.9	1.08	-5.82	0.012
K	cassette exon	<i>SDHAF2</i>	chr29	+	3.64	1.69	-1.95	0.008
N	alternative 3' splice site	<i>SLC25A2</i>	chr11	+	7.55	12.9	-5.35	0.046
N	alternative 3' splice site	<i>SLC25A25</i>	chr11	+	100	94.2	-5.8	0.032
N	alternative 3' splice site	<i>SLC29A1</i>	chr23	+	86.36	87.67	1.31	0.039
N	alternative 3' splice site	<i>SLC29A2</i>	chr29	-	7.55	5.92	4.37	0.006
K	cassette exon	<i>SLMAP</i>	chr22	-	81.31	87.88	6.57	0.043
N	alternative 5' splice site	<i>SMTNL1</i>	chr15	+	90	92.45	2.45	0.008
N	alternative 5' splice site	<i>SNX3</i>	chr9	+	100	97.42	-2.58	0.015
K	alternative last exon	<i>SPECC1</i>	chr19	-	61.54	71.43	9.89	0.001
K	alternative last exon	<i>SPEG</i>	chr2	+	3.63	4.52	0.89	0.005
N	alternative 3' splice site	<i>SPTBN1</i>	chr11	+	72.65	65.88	-6.77	0.037
N	alternative 3' splice site	<i>SPTBN1</i>	chr11	+	18.18	24.14	5.96	0.008
N	alternative 5' splice site	<i>SPTBN1</i>	chr11	+	30.43	40.24	9.81	0.004

N	alternative 3' splice site	<i>SRCAP</i>	chr25	+	100	96	-4	0.043
K	cassette exon	<i>SRRM2</i>	chr25	-	84.96	89.86	4.9	0.035
K	cassette exon	<i>ST6GALN</i>	chr11	-	75	84.7	9.7	0.019
N	alternative 5' splice site	<i>STAT3</i>	chr19	-	4.55	7.14	2.59	0.001
K	coordinate cassette exon	<i>STX8</i>	chr19	-	40	47.11	7.11	0.027
N	alternative 3' splice site	<i>SVIL</i>	chr13	-	94.74	100	5.26	0.005
N	alternative 3' splice site	<i>SVIL</i>	chr13	-	93.2	100	6.8	0.010
N	alternative 3' splice site	<i>TACC2</i>	chr26	+	95.31	89.92	-5.39	0.036
N	alternative 3' splice site	<i>TAX1BP</i>	chr4	-	2.38	4.67	2.29	0.016
N	alternative 3' splice site	<i>TAX1BP</i>	chr4	-	97.3	94.91	-2.39	0.018
K	cassette exon	<i>TMED2</i>	chr17	-	82.61	77.42	-5.19	0.020
N	alternative 5' splice site	<i>TMEM120</i>	chr25	-	16.03	11.24	-4.79	0.017
N	alternative 5' splice site	<i>TMEM120</i>	chr25	-	83.71	88.24	4.53	0.034
N	alternative 5' splice site	<i>TMEM126A</i>	chr29	-	8.54	5.26	-3.28	0.002
N	alternative 5' splice site	<i>TMEM126A</i>	chr29	-	89.47	94.39	4.92	0.000
N	cassette exon	<i>TMEM14C</i>	chr23	-	41.69	47.66	5.97	0.007
N	cassette exon	<i>TMEM14C</i>	chr23	-	45.74	52.27	6.53	0.008
N	coordinate cassette exon	<i>TMEM14C</i>	chr23	-	43.52	50	6.48	0.005
N	alternative 5' splice site	<i>UBR5</i>	chr14	+	50.5	62.9	12.14	0.038
K	alternative first exon	<i>UGP2</i>	chr11	+	12.66	10.64	-2.02	0.023
K	alternative first exon	<i>UGP2</i>	chr11	+	30.67	38.36	7.69	0.027
K	cassette exon	<i>USP28</i>	chr15	-	97.9	100	2.1	0.039
K	cassette exon	<i>USP28</i>	chr15	-	96.3	100	3.7	0.039
N	coordinate cassette exon	<i>USP34</i>	chr11	-	78.17	85.16	6.99	0.008
K	alternative 3' splice site	<i>VCL</i>	chr28	+	71.43	67.43	-4	0.031
N	alternative 3' splice site	<i>VLDLR</i>	chr8	-	100	96.2	-3.8	0.028
N	cassette exon	<i>WNK1</i>	chr5	+	42.97	35.63	-7.34	0.047
N	alternative 3' splice site	<i>WNK2</i>	chr8	+	75	84.7	9.7	0.031

K	alternative first exon	<i>ZBTB7B</i>	chr3	-	82.84	89.66	6.82	0.019
N	alternative last exon	<i>ZFAND2B</i>	chr2	+	55	59.49	4.49	0.037
N	alternative 3' splice site	<i>ZFAND5</i>	chr8	-	85.05	80.39	-4.66	0.019
N	alternative 3' splice site	<i>ZFAND5</i>	chr8	-	2.22	3.66	1.44	0.041
K	cassette exon	<i>ZNF106</i>	chr10	-	8.46	6.25	-2.21	0.012

Table S2. Alternatively spliced and differentially expressed ($p < 0.05$) transcripts identified in the ribeye muscle of Nelore bulls with the highest and lowest intramuscular fat content animals (HIF and LIF). Green highlighted: hub genes

Annotated Junctions (Known and Novel)	Transcripts type	Gene	Chr	Strand	HIF	LIF	Delta (LIF-HIF)	p value
N	alternative 3' splice site	<i>ACAT1</i>	chr15	+	98.4	100	1.6	0.024
K	cassette exon	<i>ACTR3B</i>	chr4	+	3.7	6.15	-2.45	0.021
N	cassette exon	<i>ACTR3B</i>	chr4	+	100	97.2	-2.8	0.025
N	alternative 3' splice site	<i>ADSSL1</i>	chr21	+	92.56	100	7.44	0.020
N	intron retention	<i>AHNAK</i>	chr29	+	98	100	2	0.013
N	intron retention	<i>AHNAK</i>	chr29	+	95.4	100	4.6	0.031
K	cassette exon	<i>AKIRIN2</i>	chr9	+	97.83	100	2.17	0.039
N	alternative 5' splice site	<i>AKTIP</i>	chr18	-	3.77	0	-3.77	0.032
N	alternative 5' splice site	<i>ALPK2</i>	chr24	-	9.09	12.2	3.11	0.045
N	alternative 3' splice site	<i>ANAPC11</i>	chr19	-	3.7	2.38	-1.32	0.007
K	alternative 5' splice site	<i>ANK3</i>	chr28	-	95.65	97.8	2.15	0.026
N	cassette exon	<i>ANKRD10</i>	chr12	-	80.03	85.71	5.68	0.011
N	alternative 3' splice site	<i>ANXA11</i>	chr28	+	7.41	5.41	-2	0.031
K	cassette exon	<i>APLP2</i>	chr29	+	65.69	53.85	-11.84	0.018
N	alternative 5' splice site	<i>APOBEC2</i>	chr23	+	3.04	7.3	4.26	0.004
N	alternative 5' splice site	<i>APOBEC2</i>	chr23	+	96.3	99.09	2.79	0.005
N	alternative 5' splice site	<i>ARFGAP2</i>	chr15	-	3.39	7.2	3.81	0.012
N	alternative 3' splice site	<i>ARHGEF10L</i>	chr2	-	100	94.8	-5.2	0.043

N	alternative 3' splice site	<i>ARIH2</i>	chr22	-	94.67	100	5.33	0.050
K	alternative first exon	<i>ARMCX3</i>	chrX	+	15.91	11.11	-4.8	0.030
N	alternative 3' splice site	<i>ARPP21</i>	chr22	+	9.52	4.4	-5.12	0.032
K	alternative first exon	<i>ART4</i>	chr6	+	15.59	13.96	-1.62	0.041
N	cassette exon	<i>ART4</i>	chr6	+	83.33	89.86	6.53	0.034
K	cassette exon	<i>ATL2</i>	chr11	-	98.4	100	1.6	0.019
K	alternative 3' splice site	<i>ATN1</i>	chr5	-	13.64	15.84	2.2	0.040
N	alternative 5' splice site	<i>ATP6V1C</i>	chr14	-	2.86	1.2	-1.66	0.025
N	cassette exon	<i>ATXN10</i>	chr5	+	96.83	100	3.17	0.017
N	alternative 3' splice site	<i>AURKAIP</i>	chr16	+	15.65	8.12	-7.53	0.038
N	alternative first exon	<i>BCCIP</i>	chr26	+	100	98	-2	0.011
K	coordinate cassette exon	<i>BIN1</i>	chr2	+	97.92	100	2.08	0.048
K	intron retention	<i>BUD31</i>	chr25	-	5.41	2.56	-2.85	0.042
N	alternative 3' splice site	<i>C10H15o</i>	chr10	-	8.77	3.85	-4.92	0.011
N	alternative 5' splice site	<i>C10H15o</i>	chr10	-	15.79	10.1	-5.69	0.006
N	alternative 3' splice site	<i>C13H20o</i>	chr13	+	25	29.29	4.29	0.006
N	alternative 3' splice site	<i>C13H20o</i>	chr13	+	72.41	67.54	-4.87	0.050
K	cassette exon	<i>C28H10o</i>	chr28	+	22.22	13.64	-8.58	0.046
K	cassette exon	<i>C3H1orf</i>	chr3	+	94.03	90.38	-3.65	0.030
N	alternative 3' splice site	<i>CCDC23</i>	chr3	+	96.5	100	3.5	0.039
K	alternative last exon	<i>CD58</i>	chr3	+	88.56	91.29	2.72	0.045
N	intron retention	<i>CKM</i>	chr18	-	76.78	83.87	7.09	0.020
K	cassette exon	<i>CLIP1</i>	chr17	+	96.61	100	3.39	0.009
N	cassette exon	<i>CMC1</i>	chr22	+	100	95.83	-4.17	0.043
N	cassette exon	<i>CMC2</i>	chr18	-	95.65	97.8	2.15	0.044
N	alternative 3' splice site	<i>CMYA5</i>	chr10	+	96.15	100	3.85	0.015
K	alternative last exon	<i>COL6A2</i>	chr1	+	14.72	15.79	1.07	0.040
N	alternative 5' splice site	<i>COPS5</i>	chr14	-	94.74	95.83	1.09	0.042

N	alternative 3' splice site	<i>COX7C</i>	chr7	+	93.87	97.4	3.53	0.017
K	alternative first exon	<i>CS</i>	chr5	+	2.2	1.2	-1	0.007
K	alternative last exon	<i>CTNNB1</i>	chr22	+	79.41	83.33	3.92	0.031
N	cassette exon	<i>CTSD</i>	chr29	-	97.11	100	2.89	0.010
K	alternative 5' splice site	<i>CTTN</i>	chr29	+	13.33	9.68	-3.65	0.019
N	cassette exon	<i>CUTC</i>	chr26	+	96.72	100	3.28	0.010
N	alternative 3' splice site	<i>CUTC</i>	chr26	+	100	91	-9	0.048
N	alternative 3' splice site	<i>DDX39B</i>	chr23	+	7.14	2	-5.14	0.004
N	alternative 3' splice site	<i>DDX39B</i>	chr23	+	90.7	100	9.3	0.002
N	alternative 3' splice site	<i>DES</i>	chr2	+	83.33	89.86	6.53	0.042
N	alternative 3' splice site	<i>DES</i>	chr2	+	96	100	4	0.042
N	alternative 5' splice site	<i>DLGAP4</i>	chr13	+	1.85	0.3	-1.55	0.005
N	alternative 3' splice site	<i>DNM2</i>	chr7	+	11.54	6.51	-5.03	0.007
N	alternative 3' splice site	<i>DNM2</i>	chr7	+	71.43	81.17	9.74	0.017
N	cassette exon	<i>DTNA</i>	chr24	-	15.59	13.96	-1.62	0.049
K	alternative first exon	<i>DTNA</i>	chr24	-	95.65	100	4.35	0.003
K	cassette exon	<i>DUSP26</i>	chr27	-	6.94	3.92	-3.02	0.008
K	cassette exon	<i>DYRK1B</i>	chr18	-	47.37	43.17	-4.2	0.017
N	alternative 3' splice site	<i>EIF4B</i>	chr5	-	94.22	100	5.78	0.023
K	cassette exon	<i>EIF4G1</i>	chr1	-	88.24	85	-3.24	0.027
N	cassette exon	<i>FAM96B</i>	chr18	-	78.62	82.35	3.73	0.038
N	alternative 3' splice site	<i>FBXW5</i>	chr11	+	2.5	0.5	-2	0.019
N	alternative 3' splice site	<i>FBXW5</i>	chr11	+	97	100	3	0.018
N	alternative 3' splice site	<i>FH</i>	chr16	+	100	98.2	-1.8	0.050
N	cassette exon	<i>FIBP</i>	chr29	-	100	93.9	-6.1	0.049
N	alternative 3' splice site	<i>FKBP5</i>	chr23	-	97.92	100	2.08	0.047
N	alternative 3' splice site	<i>FKBP5</i>	chr23	-	100	94	-6	0.047
N	cassette exon	<i>FKBP5</i>	chr23	-	100	92.7	-7.3	0.048

N	alternative 3' splice site	<i>FLNC</i>	chr4	+	98.3	100	1.7	0.017
N	alternative 5' splice site	<i>GALNT11</i>	chr4	+	3.58	2	-0.58	0.002
N	alternative 3' splice site	<i>GBAS</i>	chr25	+	90.7	100	9.3	0.043
K	cassette exon	<i>GDE1</i>	chr25	-	4.17	2	-2.17	0.003
K	cassette exon	<i>GKAP1</i>	chr8	-	96.59	100	3.41	0.022
K	cassette exon	<i>GLRX2</i>	chr16	+	9.09	5.5	-3.58	0.002
N	alternative 5' splice site	<i>GLRX2</i>	chr2	+	97.92	100	2.08	0.005
N	alternative 5' splice site	<i>GLRX5</i>	chr21	+	100	97.83	-2.17	0.026
N	alternative 5' splice site	<i>GLRX5</i>	chr21	+	100	94.8	-5.2	0.045
N	alternative 3' splice site	<i>GNB2</i>	chr25	-	100	97.2	-2.8	0.045
N	alternative 3' splice site	<i>HDLBP</i>	chr3	-	2.2	1.2	-1	0.047
N	alternative first exon	<i>HHATL</i>	chr22	-	2.41	6.2	3.79	0.005
N	alternative 5' splice site	<i>HNRNPA1</i>	chr5	-	11.02	8.65	-2.37	0.034
K	cassette exon	<i>HNRNPD</i>	chr6	-	95.65	98.98	3.33	0.020
N	alternative 5' splice site	<i>HNRNPH1</i>	chr7	+	1.19	3.57	2.38	0.006
N	alternative 5' splice site	<i>HNRNPH1</i>	chr7	+	14.29	8.11	-6.18	0.019
N	alternative 5' splice site	<i>HNRNPH3</i>	chr28	+	3.49	2	-1.49	0.004
K	alternative first exon	<i>HNRNPK</i>	chr8	-	23.33	28.57	5.24	0.029
N	alternative 5' splice site	<i>HSD17B8</i>	chr23	+	5.8	8.7	2.9	0.019
N	alternative 5' splice site	<i>HSDL2</i>	chr8	+	4.55	1.89	-2.66	0.001
K	alternative 3' splice site	<i>HSF1</i>	chr14	-	8.11	7.02	-1.09	0.023
N	alternative 5' splice site	<i>HSP90AB1</i>	chr23	+	96.72	100	3.28	0.008
N	alternative 5' splice site	<i>HSP90AB1</i>	chr23	+	3.28	1.02	-4.3	0.009
N	alternative 5' splice site	<i>HSPE1</i>	chr2	+	2.2	1.7	-0.5	0.006
N	alternative 3' splice site	<i>IDH3B</i>	chr13	+	91.03	93.1	2.07	0.021
N	alternative 3' splice site	<i>IDH3B</i>	chr13	+	2.88	1.59	-1.29	0.015
N	cassette exon	<i>INPPL1</i>	chr15	+	94.28	100	5.72	0.002
N	alternative 5' splice site	<i>ISOC2</i>	chr18	+	8.33	5.33	-3	0.043

N	intron retention	<i>JTB</i>	chr3	+	97.87	100	2.13	0.013
K	cassette exon	<i>KARS</i>	chr18	-	17.65	14.29	-3.36	0.016
N	alternative 5' splice site	<i>KDM1A</i>	chr2	-	1.85	0	-1.85	0.030
N	cassette exon	<i>KIAA0368</i>	chr8	-	97.83	100	2.17	0.014
N	alternative 3' splice site	<i>KIAA1671</i>	chr17	+	8.11	7.02	-1.09	0.037
N	alternative 5' splice site	<i>KIAA1715</i>	chr2	+	23.4	30.43	7.03	0.022
K	cassette exon	<i>KIF1B</i>	chr16	-	96.48	100	3.52	0.031
K	intron retention	<i>KLF2</i>	chr7	-	5.17	0	-5.17	0.024
N	alternative 3' splice site	<i>KLHL30</i>	chr3	-	1.83	5.26	3.42	0.019
N	alternative 3' splice site	<i>KLHL30</i>	chr3	-	1.7	1.05	-0.65	0.021
N	alternative 3' splice site	<i>KLHL30</i>	chr3	-	2.07	5.26	3.19	0.023
K	cassette exon	<i>LAMTOR2</i>	chr3	-	93.44	90	-3.44	0.013
N	alternative 3' splice site	<i>LARP1</i>	chr7	+	100	89.47	-10.53	0.021
K	alternative first exon	<i>LINGO1</i>	chr21	-	9.43	4.82	-4.61	0.007
K	coordinate cassette exon	<i>LMO7</i>	chr12	+	6.06	0	-6.06	0.002
N	alternative 3' splice site	<i>LMOD3</i>	chr22	+	2.94	1.79	-1.15	0.013
N	alternative 3' splice site	<i>LMOD3</i>	chr22	+	97.06	98.22	1.16	0.015
N	cassette exon	<i>LOC1008</i>	chr10	+	92.11	100	7.89	0.033
N	alternative 3' splice site	<i>LOC1019</i>	chr13	+	17.78	10	-7.78	0.009
N	alternative 3' splice site	<i>LOC104975822</i>	chr25	+	1.67	0	-1.67	0.026
N	cassette exon	<i>LRRC2</i>	chr22	+	96.1	100	3.9	0.037
N	alternative 3' splice site	<i>MAF1</i>	chr14	-	81.48	86.84	5.36	0.030
K	alternative first exon	<i>MAP2K6</i>	chr19	-	48.81	62.16	13.35	0.050
N	cassette exon	<i>MAZ</i>	chr25	-	36.08	32.26	-3.82	0.031
N	cassette exon	<i>MBNL1</i>	chr1	-	2.35	0	-2.35	0.014
N	alternative 3' splice site	<i>MEF2C</i>	chr7	-	5.43	8.33	2.9	0.005
K	mutually exclusive exons	<i>MEF2D</i>	chr3	+	15.38	19.35	3.97	0.039
N	coordinate cassette exon	<i>MEMO1</i>	chr11	-	96.3	100	3.7	0.019

N	alternative first exon	<i>METTL21B</i>	chr5	-	95.8	100	4.2	0.003
N	alternative 5' splice site	<i>MGAT4B</i>	chr7	+	1.7	1.05	-0.65	0.033
N	alternative 5' splice site	<i>MID1IP1</i>	chrX	-	100	100	0	0.045
N	cassette exon	<i>MIF4GD</i>	chr19	+	80	67.71	-12.29	0.006
N	intron retention	<i>MIR208A</i>	chr10	+	79.9	100	20.1	0.036
N	alternative 5' splice site	<i>MIR208B</i>	chr10	+	10.3	10.74	0.44	0.034
K	cassette exon	<i>MKNK2</i>	chr7	+	2.78	2.17	-0.61	0.020
N	cassette exon	<i>MLIP</i>	chr23	-	84.81	94.12	9.31	0.008
N	alternative 3' splice site	<i>MPRIP</i>	chr19	-	0.91	2.27	1.36	0.049
N	cassette exon	<i>MPRIP</i>	chr19	-	51.65	58.14	6.49	0.018
N	alternative 5' splice site	<i>MRPL33</i>	chr11	-	87.5	90	2.5	0.023
N	cassette exon	<i>MRPS18A</i>	chr23	-	95.69	100	4.31	0.008
K	cassette exon	<i>MRPS25</i>	chr22	+	92.31	96.08	3.77	0.016
N	alternative 5' splice site	<i>MVP</i>	chr25	-	4	0	-4	0.048
N	alternative 3' splice site	<i>MYH7B</i>	chr13	+	11.67	15.91	4.24	0.014
N	alternative 5' splice site	<i>MYH7B</i>	chr13	+	23.4	30.43	7.03	0.006
N	alternative 3' splice site	<i>MYH7B</i>	chr13	+	80	72.73	-7.27	0.023
N	alternative 3' splice site	<i>MYLK2</i>	chr3	-	10.45	7.14	-3.31	0.028
N	alternative 3' splice site	<i>MYLK2</i>	chr3	-	3.85	1.69	-2.16	0.032
N	alternative 3' splice site	<i>MYLK2</i>	chr3	-	13.55	8	-5.55	0.038
N	intron retention	<i>MYO18A</i>	chr19	-	12.63	6.67	-5.96	0.038
K	intron retention	<i>MYO18A</i>	chr19	-	1.94	0	-1.94	0.040
N	alternative 3' splice site	<i>MYOT</i>	chr7	+	3.97	3.51	-0.46	0.031
N	alternative 3' splice site	<i>MYOZ1</i>	chr28	-	89.37	90.85	1.48	0.008
N	alternative 5' splice site	<i>MYOZ3</i>	chr7	+	2.42	4.35	1.93	0.039
K	alternative first exon	<i>MYOZ3</i>	chr7	+	52.63	43.48	-9.15	0.037
N	alternative 5' splice site	<i>NACA</i>	chr11	-	12.5	9.09	-3.41	0.011
K	alternative last exon	<i>NAP1L1</i>	chr5	-	5	0	-5	0.003

K	alternative 3' splice site	<i>NAP1L4</i>	chr29	+	3.33	2.31	-1.02	0.005
K	cassette exon	<i>NBR1</i>	chr19	+	95.83	100	4.17	0.022
N	cassette exon	<i>NDUFS4</i>	chr20	-	95.24	98.64	3.4	0.039
N	alternative 5' splice site	<i>NFE2L1</i>	chr19	-	10.53	3.36	-7.17	0.010
K	cassette exon	<i>NFE2L1</i>	chr19	-	71.88	68.29	-3.59	0.003
N	alternative 3' splice site	<i>NFE2L1</i>	chr19	-	16.18	6.67	-9.51	0.012
N	alternative 3' splice site	<i>NFE2L1</i>	chr19	-	0	1.92	1.92	0.023
N	alternative 3' splice site	<i>NFE2L1</i>	chr19	-	81.18	90.32	9.14	0.038
N	alternative 5' splice site	<i>NFE2L1</i>	chr19	-	87.47	94.63	7.16	0.011
N	cassette exon	<i>NR4A1</i>	chr5	-	100	92.7	-7.3	0.050
K	cassette exon	<i>NRBP1</i>	chr11	-	2.02	0	-2.02	0.038
K	cassette exon	<i>NTPCR</i>	chr28	+	4.08	0	-4.08	0.046
N	alternative 3' splice site	<i>OAZ2</i>	chr10	+	89.58	100	10.42	0.050
N	alternative 5' splice site	<i>OBSCN</i>	chr7	-	100	92.7	-7.3	0.027
N	alternative 5' splice site	<i>P2RX5</i>	chr19	-	68.58	75	6.42	0.034
K	cassette exon	<i>PCBP2</i>	chr5	-	51.85	56.73	4.88	0.008
K	cassette exon	<i>PCBP2</i>	chr5	-	38.48	46.92	8.44	0.024
N	cassette exon	<i>PDLIM5</i>	chr6	-	96.37	97.84	1.47	0.005
N	cassette exon	<i>PDLIM5</i>	chr6	-	93.03	96.15	3.12	0.009
N	cassette exon	<i>PDLIM5</i>	chr6	-	1.7	1.05	-0.65	0.024
N	cassette exon	<i>PEBP4</i>	chr8	-	90.74	95.46	4.72	0.001
N	cassette exon	<i>PEBP4</i>	chr8	-	83.33	85.66	2.33	0.024
N	alternative 3' splice site	<i>PERM1</i>	chr16	+	12.09	6.25	-5.84	0.007
N	alternative 3' splice site	<i>PFKFB4</i>	chr22	+	93.62	95.61	1.98	0.046
N	alternative 3' splice site	<i>PIP4K2B</i>	chr19	-	100	92.7	-7.3	0.026
N	alternative first exon	<i>PLXNB2</i>	chr5	-	95.24	96.3	1.06	0.048
N	alternative 3' splice site	<i>POLR1D</i>	chr12	-	8.89	6.25	-2.64	0.038
K	cassette exon	<i>POLR2E</i>	chr2	+	98.15	100	1.85	0.027

N	coordinate cassette exon	<i>POLR2E</i>	chr7	-	97.62	100	2.38	0.040
N	alternative 3' splice site	<i>PPFIA4</i>	chr16	+	89	92.73	3.73	0.004
K	cassette exon	<i>PPFIBP1</i>	chr5	-	77.46	63.33	-14.13	0.022
N	alternative 5' splice site	<i>PPP2R2A</i>	chr8	+	2.55	0	-2.55	0.006
N	cassette exon	<i>PSMD1</i>	chr12	-	99.09	100	0.91	0.008
K	cassette exon	<i>PSMD14</i>	chr2	-	5.41	2.9	-2.52	0.017
N	alternative first exon	<i>PTP4A3</i>	chr14	-	4.72	0	-4.72	0.013
N	alternative first exon	<i>PTP4A3</i>	chr14	-	83.76	90.48	6.72	0.024
N	cassette exon	<i>RAB18</i>	chr13	-	100	100	0	0.042
K	alternative 3' splice site	<i>RAD23A</i>	chr7	-	23.65	26.17	2.52	0.045
N	alternative 5' splice site	<i>RAD9A</i>	chr29	-	4.76	3.39	-1.37	0.047
N	alternative 5' splice site	<i>RAMP1</i>	chr3	+	96.77	100	3.23	0.042
N	alternative 3' splice site	<i>RAPGEF1</i>	chr11	-	95.24	100	4.76	0.008
N	alternative 3' splice site	<i>RAPGEF1</i>	chr11	-	3.95	0	-3.95	0.013
K	alternative 5' splice site	<i>RB1CC1</i>	chr14	-	4.65	0	-4.65	0.035
N	alternative 3' splice site	<i>RBL2</i>	chr18	+	87.5	100	12.5	0.000
N	alternative 3' splice site	<i>RBM24</i>	chr23	-	3.42	2	-1.42	0.018
N	alternative 3' splice site	<i>RBMX</i>	chrX	-	94.89	100	5.11	0.049
N	coordinate cassette exon	<i>RCHY1</i>	chr6	-	100	100	0	0.022
N	alternative 5' splice site	<i>RFNG</i>	chr19	+	7.02	9.52	2.5	0.034
N	alternative 5' splice site	<i>RNASE4</i>	chr10	-	97.37	100	2.63	0.010
N	alternative 5' splice site	<i>RNF7</i>	chr1	-	2.55	0	-2.55	0.017
N	alternative 5' splice site	<i>RNF7</i>	chr1	-	7.12	0	-7.12	0.042
N	alternative 3' splice site	<i>RPL3L</i>	chr25	-	94.91	96.75	1.84	0.043
N	alternative 5' splice site	<i>RPLP1</i>	chr10	+	95.13	97.67	2.54	0.022
N	alternative 5' splice site	<i>RPLP1</i>	chr10	+	4.87	2.33	-2.54	0.022
N	alternative 5' splice site	<i>RPS26</i>	chr5	-	96.7	97.85	1.15	0.012
N	cassette exon	<i>RRAGD</i>	chr9	+	88.24	93.88	5.64	0.041

N	coordinate cassette exon	<i>RRAGD</i>	chr9	+	93.1	97.03	3.93	0.027
N	cassette exon	<i>RYK</i>	chr1	+	100	100	0	0.027
N	cassette exon	<i>SAR1A</i>	chr28	-	21.32	11.23	-10.09	0.014
N	alternative 3' splice site	<i>SESN1</i>	chr9	+	1.33	0	-1.33	0.010
N	alternative 3' splice site	<i>SFT2D1</i>	chr13	+	15.85	19.59	3.74	0.031
N	alternative 5' splice site	<i>SFT2D1</i>	chr9	-	2.7	0	-2.7	0.042
N	alternative 3' splice site	<i>SLC16A1</i>	chr3	+	94.67	100	5.33	0.024
K	alternative first exon	<i>SLC29A1</i>	chr23	+	99.09	100	0.91	0.009
N	alternative 3' splice site	<i>SLC29A2</i>	chr29	-	2.54	0	-2.54	0.028
N	cassette exon	<i>SLC38A2</i>	chr5	+	100	100	0	0.047
K	cassette exon	<i>SMYD1</i>	chr11	-	51.43	60.87	9.44	0.013
K	cassette exon	<i>SORBS1</i>	chr26	-	97.47	100	2.53	0.047
K	alternative 5' splice site	<i>SPEG</i>	chr2	+	95.45	97	1.55	0.037
K	cassette exon	<i>SRRM2</i>	chr25	-	47.87	33.52	-14.35	0.048
N	cassette exon	<i>ST3GAL3</i>	chr3	-	98.53	100	1.47	0.007
K	alternative 3' splice site	<i>ST6GALN</i>	chr11	-	2.22	0	-2.22	0.005
N	coordinate cassette exon	<i>STAC3</i>	chr5	+	97.13	94.62	-2.51	0.038
K	alternative 3' splice site	<i>STK11</i>	chr7	-	95.92	98.73	2.81	0.011
N	cassette exon	<i>STRN3</i>	chr21	-	100	100	0	0.046
K	alternative last exon	<i>STX4</i>	chr25	+	96.77	95	-1.77	0.049
K	cassette exon	<i>SYNPO2L</i>	chr29	+	83.97	75	-8.97	0.020
N	alternative 5' splice site	<i>TACC2</i>	chr26	+	55.83	50	-5.83	0.036
N	cassette exon	<i>TAX1BP1</i>	chr4	-	96.3	100	3.7	0.004
K	coordinate cassette exon	<i>TAX1BP1</i>	chr4	-	96.43	100	3.57	0.007
K	cassette exon	<i>TFDP2</i>	chr1	+	24.72	20	-4.72	0.009
K	cassette exon	<i>THRAP3</i>	chr5	+	17.31	13.58	-3.73	0.011
N	alternative 3' splice site	<i>THYN1</i>	chr15	-	1.92	0	-1.92	0.009
K	alternative first exon	<i>THYN1</i>	chr15	-	2.48	0	-2.48	0.015

N	alternative 3' splice site	<i>TIAL1</i>	chr26	-	81.97	85.44	3.47	0.029
N	alternative 3' splice site	<i>TIAL1</i>	chr26	-	17.91	14.11	-3.8	0.035
N	cassette exon	<i>TMEM189</i>	chr13	-	33.52	47.87	14.35	0.046
N	cassette exon	<i>TMEM259</i>	chr7	-	100	96.66	-3.34	0.041
N	coordinate cassette exon	<i>TNS1</i>	chr2	-	86.41	97.06	10.66	0.006
K	alternative 5' splice site	<i>TOB2</i>	chr5	-	99.09	100	0.91	0.049
K	alternative 3' splice site	<i>TRIP12</i>	chr2	-	2.04	0	-2.04	0.009
K	cassette exon	<i>TWF2</i>	chr22	+	98.1	97.63	-0.47	0.035
N	alternative 3' splice site	<i>UBAC2</i>	chr12	+	3.13	0	-3.13	0.023
K	cassette exon	<i>UBE2D1</i>	chr26	-	96.55	100	3.45	0.006
K	cassette exon	<i>UBE3A</i>	chr21	-	98.39	100	1.61	0.006
K	alternative 3' splice site	<i>UBE3C</i>	chr4	+	67.06	63.64	-3.42	0.019
N	alternative 5' splice site	<i>UBE3C</i>	chr6	+	78.76	84.31	5.55	0.006
N	alternative 5' splice site	<i>UBE3C</i>	chr6	+	11.54	0	-11.54	0.046
N	alternative 5' splice site	<i>UBL4A</i>	chrX	-	25	23.61	-1.39	0.046
K	alternative 3' splice site	<i>UCHL5</i>	chr16	+	22.81	20	-2.81	0.025
N	alternative 3' splice site	<i>ULK1</i>	chr17	-	47.87	33.52	-14.35	0.033
K	cassette exon	<i>UQCC</i>	chr13	-	100	96	-4	0.045
K	alternative first exon	<i>USP2</i>	chr15	-	82.35	90	7.65	0.016
K	alternative 3' splice site	<i>USP28</i>	chr15	-	23.73	16.03	-7.7	0.035
N	alternative 5' splice site	<i>USP47</i>	chr15	-	97.37	100	2.63	0.013
N	alternative 3' splice site	<i>UXS1</i>	chr11	+	99.09	100	0.91	0.019
N	cassette exon	<i>WIPI1</i>	chr19	+	94.19	100	5.81	0.035
N	alternative 3' splice site	<i>XIRP1</i>	chr22	-	82.84	86.32	3.48	0.050
K	cassette exon	<i>XIRP2</i>	chr2	-	87.75	94.74	6.99	0.020
N	cassette exon	<i>YIF1A</i>	chr29	-	96.49	90	-6.49	0.042
K	cassette exon	<i>ZMYND11</i>	chr13	-	51.15	38.89	-12.26	0.002

Table S3. Biological process GO terms and pathways containing genes with differentially expressed alternatively spliced transcripts in ribeye muscle in Nelore bulls selected to be divergent for ribeye muscle area.

Category	Terms	Genes	p-value
Biological Process	GO:0042787~protein ubiquitination involved in ubiquitin-dependent protein catabolic process	<i>RNF115, UBR5, ANAPC11, RNF181, HECTD1</i>	0.01
	GO:0006936~muscle contraction	<i>MYOM2, MYBPC1, SLMAP</i>	0.02
	GO:0014823~response to activity	<i>SLC25A25, SMTNL1</i>	0.03
	GO:0032091~negative regulation of protein binding	<i>GNL3L, CARM1, PARK7</i>	0.03
	GO:0031647~regulation of protein stability	<i>USP28, GNL3L, ASPH</i>	0.04
	GO:2000021~regulation of ion homeostasis	<i>WNK1, WNK2</i>	0.04
	GO:0090073~positive regulation of protein homodimerization activity	<i>GNL3L, PARK7</i>	0.04
	GO:0033234~negative regulation of protein sumoylation	<i>GNL3L, PARK7</i>	0.04
	GO:0048745~smooth muscle tissue development	<i>ZFAND5, ITGA8</i>	0.05
	GO:0048268~clathrin coat assembly	<i>EPS15, PICALM</i>	0.05
	GO:0008283~cell proliferation	<i>USP28, PICALM, STAT3, TACC2</i>	0.06
	GO:2000637~positive regulation of gene silencing by miRNA	<i>STAT3, FXR1</i>	0.06
GO:0000209~protein polyubiquitination	<i>RNF115, UBR5, RNF181</i>	0.08	
Pathways	bta04921:Oxytocin signaling pathway	<i>ADCY2, PPP3CC, CAMK2B, RCAN1, PIK3R1</i>	0.01
	bta05414:Dilated cardiomyopathy	<i>YL3, ITGA8, ITGA7</i>	0.01
	bta04066:HIF-1 signaling pathway	<i>EGLN1, CAMK2B, STAT3, PIK3R1</i>	0.01
	bta04114:Oocyte meiosis	<i>ADCY2, PPP3CC, CAMK2B, ANAPC11</i>	0.02
	bta04510:Focal adhesion	<i>ITGA8, ITGA7, FLNC, PIK3R1, VCL</i>	0.02
	bta04120:Ubiquitin mediated proteolysis	<i>UBR5, RHOBTB1, ANAPC11, CDC34</i>	0.03
	bta04923:Regulation of lipolysis in adipocytes	<i>ADCY2, MGLL, PIK3R1</i>	0.03
	bta04261:Adrenergic signaling in cardiomyocytes	<i>ADCY2, MYL3, CAMK2B, PIK3R1</i>	0.03
	bta05166:HTLV-I infection	<i>ADCY2, BOLA-DOA, PPP3CC, ANAPC11, PIK3R1</i>	0.04
	bta05410:Hypertrophic cardiomyopathy (HCM)	<i>MYL3, ITGA8, ITGA7</i>	0.06
	bta04062:Chemokine signaling pathway	<i>ADCY2, CXCL12, STAT3, PIK3R1</i>	0.06
	bta04020:Calcium signaling pathway	<i>ADCY2, PLCD4, PPP3CC, CAMK2B</i>	0.07

bta04512:ECM-receptor interaction	<i>CD36, ITGA8, ITGA7</i>	0.07
bta04914:Progesterone-mediated oocyte maturation	<i>ADCY2, ANAPC11, PIK3R1</i>	0.07
bta04024:cAMP signaling pathway	<i>FXD1, ADCY2, CAMK2B, PIK3R1</i>	0.08
bta05205:Proteoglycans in cancer	<i>CAMK2B, FLNC, STAT3, PIK3R1</i>	0.08
bta04922:Glucagon signaling pathway	<i>ADCY2, PPP3CC, CAMK2B</i>	0.08
bta04810:Regulation of actin cytoskeleton	<i>ITGA8, ITGA7, PIK3R1, VCL</i>	0.09
bta04750:Inflammatory mediator regulation of TRP channels	<i>ADCY2, CAMK2B, PIK3R1</i>	0.10

Table S4. Biological process GO terms and pathways containing genes with differentially expressed alternatively spliced transcripts in ribeye muscle of Nelore bulls selected to be divergent for intramuscular fat content

	Terms	Genes	p-value
Biological Process	GO:0042787~protein ubiquitination involved in ubiquitin-dependent protein catabolic process	<i>ARIH2, RNF7, UBE3A, UBE3C, ANAPC11, RCHY1, TRIP12</i>	0.003
	GO:0016579~protein deubiquitination	<i>USP28, COPS5, USP2, USP47, UCHL5</i>	0.004
	GO:0030239~myofibril assembly	<i>MYOZ1, MYOZ3, LMOD3</i>	0.007
	GO:0035914~skeletal muscle cell differentiation	<i>MEF2C, MEF2D, NR4A1, SMYD1</i>	0.019
	GO:0006366~transcription from RNA polymerase II promoter	<i>MEF2C, POLR2E, HNRNPK, NR4A1, NFE2L1, RBMX</i>	0.024
	GO:0043484~regulation of RNA splicing	<i>MBNL1, HNRNPH1, AHNK</i>	0.031
	GO:0001958~endochondral ossification	<i>MEF2C, MEF2D, INPPL1</i>	0.034
	GO:0006511~ubiquitin-dependent protein catabolic process	<i>USP28, USP2, USP47, UCHL5, UBE2D1</i>	0.037
	GO:0023051~regulation of signaling	<i>LMO7, MVP</i>	0.038
	GO:0007274~neuromuscular synaptic transmission	<i>KIF1B, MYLK2, STAC3</i>	0.042
	GO:0006099~tricarboxylic acid cycle	<i>CS, IDH3B, FH</i>	0.042
	GO:0051262~protein tetramerization	<i>UXS1, CUTC, FH</i>	0.058
	GO:2001237~negative regulation of extrinsic apoptotic signaling pathway	<i>ZMYND11, CTTN, RB1CC1</i>	0.062
	GO:0034644~cellular response to UV	<i>KDM1A, USP28, USP47</i>	0.062

	GO:1904948~midbrain dopaminergic neuron differentiation	<i>RYK, CTNNB1</i>	0.062
	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	<i>MEF2C, COPS5, UBE3A, STRN3, NR4A1, RBMX, CTNNB1, KDM1A, HSF1, THRAP3, ARMCX3, NFE2L1, AKIRIN2</i>	0.065
	GO:0023052~signaling	<i>DLGAP4, TOB2</i>	0.074
	GO:0070125~mitochondrial translational elongation	<i>MRPS18A, MRPS25, TSFM, MRPL33</i>	0.079
	GO:1904262~negative regulation of TORC1 signaling	<i>STK11, SESN1</i>	0.085
	GO:0055008~cardiac muscle tissue morphogenesis	<i>XIRP2, MYLK2</i>	0.097
	GO:0048643~positive regulation of skeletal muscle tissue development	<i>MEF2C, CTNNB1</i>	0.097
Pathways	bta04120:Ubiquitin mediated proteolysis	<i>RNF7, UBE3A, UBE3C, ANAPC11, RCHY1, UBE2D1, TRIP12</i>	0.003
	bta04150:mTOR signaling pathway	<i>EIF4B, ULK1, STK11, RRAGD</i>	0.025
	bta00020:Citrate cycle (TCA cycle)	<i>CS, IDH3B, FH</i>	0.040
	bta03040:Spliceosome	<i>HNRNPK, DDX39B, RBMX, BUD31, HNRNPA1</i>	0.049
	bta05169:Epstein-Barr virus infection	<i>PSMD14, POLR2E, POLR1D, PSMD1, CD58, MAP2K6</i>	0.054
	bta03010:Ribosome	<i>RPS26, MRPS18A, RPLP1, RPL3L, MRPL33</i>	0.055
	bta04151:PI3K-Akt signaling pathway	<i>EIF4B, HSP90AB1, GNB2, RBL2, STK11, COL6A2, NR4A1, PPP2R2A</i>	0.077

Supplementary Figures – S1 and S2

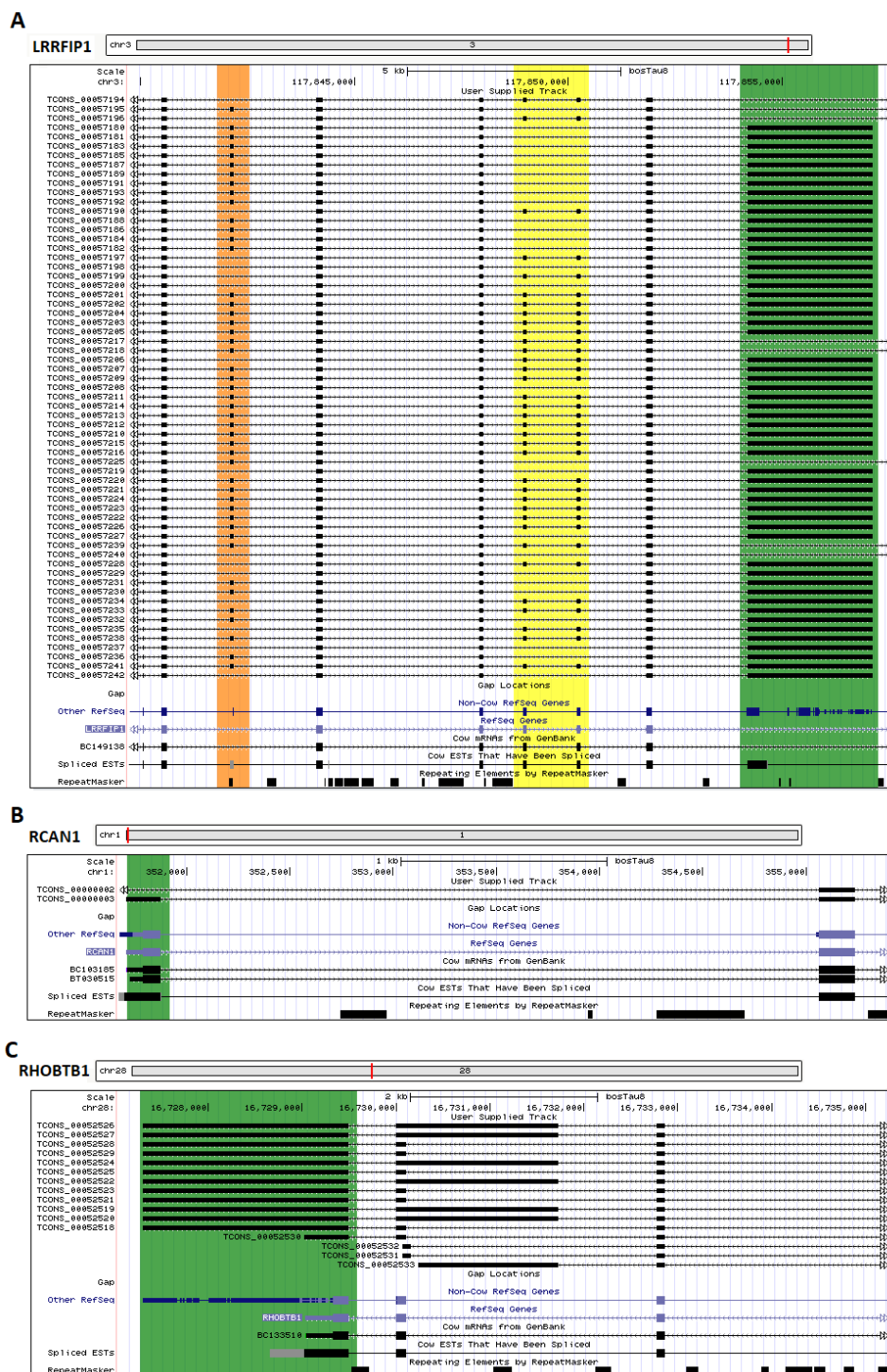


Figure S1. Genome browser visualization of the differentially expressed alternatively spliced transcripts for hub genes associated with ribeye area in Nelore bulls. A) *LRRFIP1* (+ strand), orange highlighted: cassette exon, yellow: coordinate cassette exon, green: alternative last exon; B) *RCAN1* (+ strand), green highlighted: alternative first exon; C) *RHOBTB1* (-strand), green highlighted: alternative 3' splice site.



Figure S2. Genome browser visualization of the differentially expressed alternatively spliced transcripts for the hub genes associated with intramuscular fat content in Nelore bulls. A) *TRIP12* (- strand), green highlighted: alternative 3' splice site; B) *ANXA11* (+ strand), green highlighted: alternative 3' splice site; C) *MAP2K6* (- strand), green highlighted: alternative first exon.

CAPÍTULO 5 – Considerações Finais

De acordo com os resultados encontrados nos capítulos anteriores, no geral, os genes e *splicing* alternativos associados com área de olho de lombo (REA) ou conteúdo de gordura intramuscular (IF), estavam envolvidos com metabolismo muscular, lipídico e sistema imunológico. Ao comparar os resultados dos capítulos 2, 3 e 4 (Figura 1), foram observados:

1) 644 elementos únicos (genes e *splicing* alternativos) foram diferencialmente expressos para ambas as características.

2) 12 genes foram, comumente, diferencialmente expressos em tecido muscular de bovinos Nelore e associados com REA e IF (*RGS2*, *TNC*, *SRXN1*, *GADD45A*, *SPOCK2*, *LOC786565*, *COL18A1*, *BARX2*, *NT5C2*, *SMPDL3A*, *MXRA5* e *TNFRSF12A*);

3) 26 genes que codificam transcritos *spliced* alternativamente (DEAS) foram, comumente, diferencialmente expressos e associados com REA e IF (*SPEG*, *TACC2*, *ST6GALN*, *ANAPC11*, *LOC1008*, *FAM96B*, *APLP2*, *FLNC*, *MYOT*, *HSPE1*, *SLC29A1*, *MLIP*, *C3H1orf*, *SLC29A2*, *CMYA5*, *LRRC2*, *RRAGD*, *MIF4GD*, *LOC1019*, *MPRIP*, *OBSCN*, *HSDL2*, *POLR2E*, *USP28*, *SRRM2* e *HNRNPH3*).

4) Quatro DEGs também foram *spliced* alternativamente (DEAS) e associados com REA (*SPTBN1*, *RCAN1*, *SPECC1* e *MYOM3*).

5) Dois DEGs também foram *spliced* alternativamente (DEAS) e associados com IF (*KIAA1671* e *KLHL30*).

6) Oito DEGs associados com REA também foram *spliced* alternativamente (DEAS) e associados com IF (*ATL2*, *XIRP1*, *LMOD3*, *UXS1*, *DUSP26*, *COL6A2*, *MYH7B* e *SLC16A1*).

7) Um DEG associado com IF também foi *spliced* alternativamente (DEAS) e associado com REA (*RNF115*).

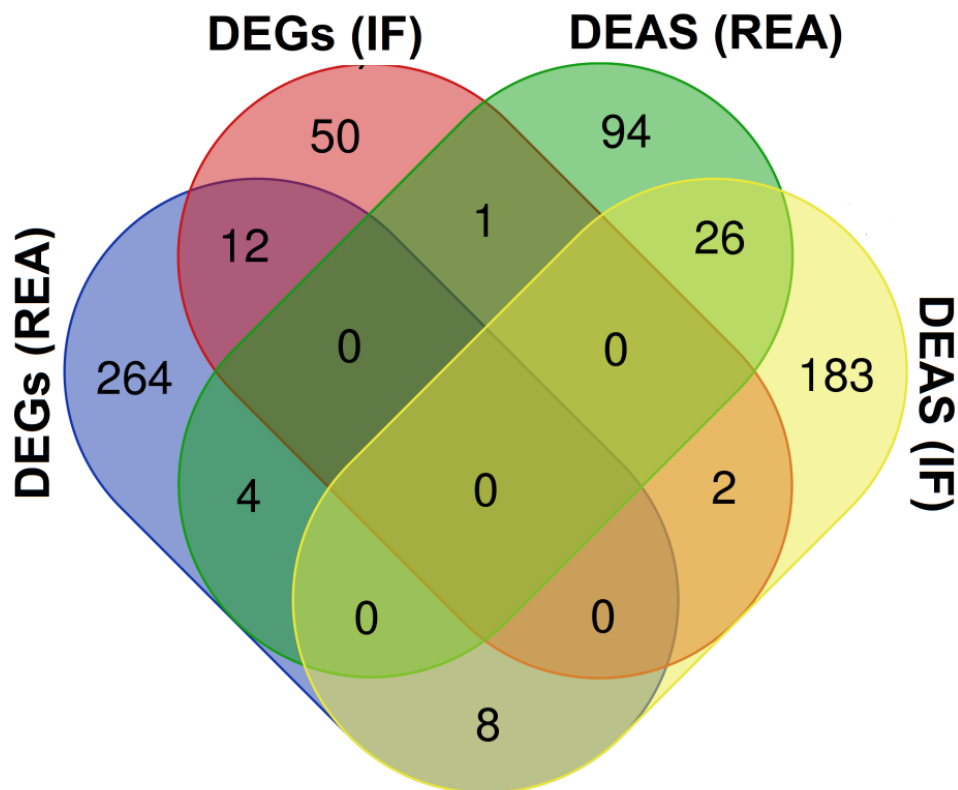


Figura 1. Venn diagram. Genes diferencialmente expressos (DEGs) e genes que codificam transcritos *spliced* alternativamente que foram diferencialmente expressos (DEAS) em animais da raça Nelore com fenótipos divergentes para área de olho de lombo (REA) e conteúdo de gordura intramuscular (IF).

Em síntese geral, pode-se concluir que as características avaliadas (área de olho de lombo e conteúdo de gordura intramuscular) em animais da raça Nelore podem sofrer ou não influência dos mesmos genes e *splicing* alternativos. No entanto, pode haver um diferente nível de expressão e regulação deles. As interações genéticas mostraram que os DEGs (Capítulo 2 e 3) e DEAS (Capítulo 4) encontrados para REA e IF participam de diferentes processos biológicos e vias, mas que de certa forma, estão intimamente relacionados ao metabolismo muscular e lipídico, o que pode futuramente, ser considerados “biomarcadores”

para as características estudadas. Estes resultados fornecem uma base para pesquisas adicionais sobre as funções específicas de genes candidatos e genes que codificam transcritos *spliced* alternativamente que foram diferencialmente expressos em animais da raça Nelore com fenótipos divergentes para área de olho de lombo e conteúdo de gordura intramuscular.