

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

**ANÁLISE MULTIVARIADA COM DADOS GENÔMICOS E
TRANSCRIPTÔMICOS PARA PERFIL DE ÁCIDOS
GRAXOS DA CARNE EM BOVINOS NELORE
TERMINADOS EM CONFINAMENTO**

Bianca Ferreira Olivieri

Bióloga

2019

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

**ANÁLISE MULTIVARIADA COM DADOS GENÔMICOS E
TRANSCRIPTÔMICOS PARA PERFIL DE ÁCIDOS
GRAXOS DA CARNE EM BOVINOS NELORE
TERMINADOS EM CONFINAMENTO**

Bianca Ferreira Olivieri

Orientador: Prof. Dr. Fernando Sebastián Baldi Rey

Coorientadores: Dra. Camila Urbano Braz

Dr. Fernando Brito Lopes

Dr. Marcos Vinícius Antunes de Lemos

**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**

2019

O49a

Olivieri, Bianca Ferreira

Análise multivariada com dados genômicos e transcriptômicos para perfil de ácidos graxos da carne em bovinos Nelore terminados em confinamento / Bianca Ferreira Olivieri. -- Jaboticabal, 2019

67 p. : tabs. + 1 CD-ROM

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

Orientador: Fernando Sebastián Baldi Rey

Coorientadora: Camila Urbano Braz

1. Zebu. 2. Genética animal. 3. Sequenciamento de nucleotídeos em larga escala. 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: ANÁLISE MULTIVARIADA COM DADOS GENÔMICOS E TRANSCRIPTÔMICOS PARA PERFIL DE ÁCIDOS GRAXOS DA CARNE EM BOVINOS NELORE TERMINADOS EM CONFINAMENTO

AUTORA: BIANCA FERREIRA OLIVIERI

ORIENTADOR: FERNANDO SEBASTIAN BALDI REY

COORIENTADOR: MARCOS VINICIUS ANTUNES DE LEMOS

COORIENTADORA: CAMILA URBANO BRAZ

COORIENTADOR: FERNANDO BRITO LOPES

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




Prof. Dr. FERNANDO SEBASTIAN BALDI REY
Departamento de Zootecnia / FCAV / Unesp - Jaboticabal



Pesquisadora Dra. NEDENIA BONVINO STAFUZZA
Instituto de Zootecnia / Sertãozinho/SP



Prof. Dr. DANISIO PRADO MUNARI
Departamento de Ciências Exatas / FCAV / Unesp - Jaboticabal



Dr. RAFAEL MEDEIROS DE OLIVEIRA SILVA
ZOETIS Brasil / São Paulo/SP



Pesquisadora Dra. ANA FABRÍCIA BRAGA MAGALHÃES
Instituto de Zootecnia / Sertãozinho/SP

DADOS CURRICULARES DO AUTOR

Bianca Ferreira Olivieri – nascida no dia 22 de janeiro de 1989, na cidade de Guararapes, S.P. A autora possui Licenciatura em Ciências Biológicas pela Universidade Federal de Mato Grosso do Sul – campus Três Lagoas, M.S., concluída em dezembro de 2011. Durante a graduação realizou estágio em Hemoglobinopatias e Talassemias no Laboratório de Genética Médica sob a supervisão da Profa. Dra. Glória Maria Gelle de Oliveira. Entre março 2012 e março de 2013, a autora ministrou aulas no Ensino Fundamental e Médio na Secretária de Ensino Estadual de São Paulo, no município de Guararapes, S.P. Em abril de 2013, autora mudou-se para Sertãozinho, S.P., onde iniciou estágio no Melhoramento Animal no Instituto de Zootecnia/ APTA – Bovinos de Corte, como bolsista CAPES, sob supervisão da pesquisadora Dra. Maria Eugênia Zerlotti Mercadante. Em março de 2014 ingressou na Pós-Graduação, no Mestrado em Genética e Melhoramento Animal na UNESP – Faculdade de Ciências Agrárias e Veterinárias (FCAV), campus Jaboticabal, como bolsista CAPES, sob a orientação do Prof. Dr. Fernando Baldi. Finalizou em julho de 2015 o Mestrado com a dissertação intitulada “Estudo de associação entre polimorfismos de base única com características de eficiência de conversão, consumo e desempenho em bovinos da raça Nelore”. E desde agosto de 2015 é aluna regular de Doutorado no Programa de Genética e Melhoramento Animal na UNESP – FCAV, como bolsista CAPES.

AGRADECIMENTOS

À Deus, pelo direito a vida e por sempre me amparar nos momentos difíceis e pelas bênçãos durante toda minha caminhada acadêmica; e permitir a realização de mais esse sonho.

Ao meu orientador, prof. Dr. Fernando Baldi. Por todo ensinamento, oportunidades, incentivo, pelo grande exemplo de competência e produtividade e toda contribuição ao mundo científico.

Aos meus coorientadores Dra. Camila Urbano, Dr. Marcos Vinicius Antunes de Lemos e Dr. Fernando Brito, agradeço por todo auxílio e suporte para realização desse trabalho. E principalmente Dr. Camila por todos ensinamentos e paciência todo esse tempo.

Aos membros da banca examinadora da qualificação: Marcos Lemos, Nedenia Bonvino, Priscila Arrigucci e Mariana Berton por toda contribuição e disponibilidade.

Aos membros da banca examinadora da defesa: Danisio Munari, Nedenia Bonvino, Ana Fabrícia Magalhães e Rafael Medeiros por toda contribuição e disponibilidade.

Aos meus pais, por toda motivação, apoio e amor durante minha trajetória. Nada teria sentido sem vocês ao meu lado.

Ao Vinícius, pelo companheirismo, amor e toda paciência.

À toda minha família e amigos que me acompanharam durante todo período.

Aos meus queridos hermanos Fabieli, Mariana, Rafael, Marcos, Sabrina K, Elisa, Sabrina A., Juan, João e Gustavo muito obrigada pela partilha, companheirismo e auxílio.

Aos meus grandes amigos Ana Paula, Laiza, André, Laura, Cleisy e Naiara obrigada pelos anos de amizade compartilhada.

Aos Professores e todos os membros do Programa de Pós-Graduação em Genética e Melhoramento Animal agradeço por toda acolhida, suporte e ensinamentos.

À toda equipe da UNESP – FCAV Jaboticabal pelo suporte e toda estrutura oferecida.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Muito obrigada!

SUMÁRIO

RESUMO.....	iv
ABSTRACT.....	vi
LISTA DE ABREVIACÕES.....	vii
CAPÍTULO 1 – CONSIDERAÇÕES GERAIS.....	1
INTRODUÇÃO.....	1
REVISÃO DE LITERATURA.....	1
Ácidos graxos na carne bovina.....	3
Variabilidade genética do perfil de ácidos graxos na carne.....	5
Estudos de transcriptômica aplicados ao melhoramento do perfil de ácidos graxos.....	7
Análise de componentes principais.....	8
Estruturas Causais.....	10
OBJETIVOS.....	11
REFERÊNCIAS BIBLIOGRÁFICAS.....	11
CAPÍTULO 2 – GENES DIFERENCIALMENTE EXPRESSOS IDENTIFICADOS COM VALORES EXTREMOS DE COMPONENTES PRINCIPAIS PARA PERFIL DE ÁCIDOS GRAXOS NA CARNE BOVINA DA RAÇA NELORE.....	21
RESUMO.....	21
ABSTRACT.....	22
INTRODUÇÃO.....	23
MATERIAL E MÉTODOS.....	24
Animais e fenótipos.....	24
Perfil de ácidos graxos.....	25
Quantificação de RNA (RNA-seq).....	26
Alinhamento das sequências.....	26
Análise de Componentes Principais (PCA) e genes diferencialmente expressos (GDE).....	27

Análise de Enriquecimento	27
RESULTADOS E DISCUSSÃO	28
CONCLUSÃO	40
REFERÊNCIAS BIBLIOGRÁFICAS	41
CAPÍTULO 3 – INVESTIGAÇÃO DE REDES CAUSAIS PARA PERFIL DE ÁCIDOS GRAXOS EM BOVINOS NELORE TERMINADOS EM CONFINAMENTO.....	46
RESUMO.....	46
ABSTRACT.....	47
INTRODUÇÃO	48
MATERIAL E MÉTODOS	49
Dados Fenotípicos	49
Animais e informações sobre manejo	49
Formação dos lotes e procedimento no abate	50
Determinação do Perfil de Ácidos Graxos em carne	50
Dados Genômicos.....	52
Genotipagem dos animais.....	52
Dados Transcriptômicos	52
Extração de RNA total.....	52
Verificação da concentração e integridade do RNA extraído.....	53
Alinhamento das sequências.....	54
Treinamento de Redes Bayesianas	54
RESULTADOS E DISCUSSÃO	56
CONCLUSÃO	60
REFERÊNCIAS BIBLIOGRÁFICAS.....	60
ANÁLISE MULTIVARIADA COM DADOS GENÔMICOS E TRANSCRIPTÔMICOS PARA PERFIL DE ÁCIDOS GRAXOS DA CARNE EM BOVINOS NELORE TERMINADOS EM CONFINAMENTO	

RESUMO – A compreensão de processos regulatórios e organização molecular dos organismos vivos progrediram consideravelmente na última década. As metodologias também evoluíram com o sequenciamento de DNA e RNA e de ferramentas genômicas permitindo a análise de centenas ou milhares de genes, proteínas ou metabólitos. O uso simultâneo dessas informações auxilia na obtenção de informações relevantes sobre as variáveis que envolvem as variações fenotípicas de características de interesse. O objetivo do presente estudo foi integrar dados fenotípicos, genotípicos e transcriptômicos em busca de aprimoramento sobre os mecanismos genéticos e metabólicos que determinam o perfil de ácidos graxos na carne de bovinos Nelore, a fim de contribuir para o melhoramento da composição de ácidos graxos da carne. Foram utilizados machos da raça Nelore terminados em confinamento, abatidos com média de idade 24 meses. Amostras do músculo *L. thoracis*, entre a 12^a a 13^a costela foram coletadas para as análises de perfil de ácidos graxos, extração de RNA e de DNA. Os resultados foram apresentados nos capítulos 2 e 3. No capítulo 2, o objetivo foi identificar genes diferencialmente expressos pelo método RNA-seq e perfil de ácidos graxos no músculo *L. thoracis* com uso de componentes principais (*principal components*: PC). Foram selecionados dois grupos de 10 animais, os quais possuíam valores de PC1 e PC2 extremos (Alto x Baixo) para os grupos somatórios de ácidos graxos (AG): ácidos graxos saturados (AGS), ácidos graxos monoinsaturados (AGMI), ácidos graxos poli-insaturados (AGPI), ômega 3 (n-3) e ômega 6 (n-6) para análise de genes diferencialmente expressos - GDE ($p < 0,10$). Para o grupo AGS foram encontrados 33 GDE para PC1 e nenhum GDE para PC2; para AGMI foram 56 GDE em PC1 e 10 GDE para PC2; AGPI apresentou apenas dois GDE para PC1 e nenhum para PC2; para ômega 3 nenhum GDE foi encontrado para PC1 e nenhum para PC2; para ômega 6 sete GDE para PC1 e três GDE para PC2. A análise de enriquecimento funcional foi realizada por meio do programa DAVID v. 6.8. Dentre os GDE, foram identificados genes relacionados a metabolismo energético (*UCP3*), adipogênese (*MBNL1*), metabolismo ou catabolismo de gordura (*MVD*, *PCYT2*, *LIPE* e *EPHX2*), que desempenham um papel na biologia lipídica celular em bovinos da raça Nelore. Esses resultados contribuem para o conhecimento dos mecanismos biológicos envolvidos na expressão perfil de ácidos graxos na carne de animais Nelore. No capítulo 3, o objetivo foi utilizar dados fenotípicos, genotípicos e transcriptômicos para investigar redes causais para perfil de ácidos graxos na carne de animais da raça Nelore, para buscar genes diferencialmente expressos (GDE) relacionados com as características em estudos. Quatro regiões genômicas (eQTLs) mostraram associações significativas ($p < 0,05$) com sete fenótipos, C16:0, C16:1, C18:0, C18:1c9, C18:2c, CLA, C22:6, e também com a expressão de 30 genes, incluindo *SREBF1*, *MDE9*, *ACSF2*, relacionados a metabolismo e catabolismo de lipídeos; e *ABCC3*, *ADCY1*, *OGDH*, metabolismo energético. O estudo apresentou relação causal entre a expressão do perfil de ácidos graxos e os processos metabólicos que envolvem a regulação de tal expressão, demonstraram efeitos causais entre os GDE encontrados. Os resultados encontrados no presente estudo contribuíram para o conhecimento dos mecanismos biológicos envolvidos no perfil de ácidos graxos na carne de animais Nelore. Pode ser aplicado no futuro para utilizar GDE encontrados como marcadores genéticos em bovinos Nelore para selecionar a composição lipídica que contribua para melhor nutrição e a saúde do consumidor.

Palavras – chave: *Bos taurus indicus*, genótipo, gordura intramuscular, vias metabólicas, RNA-seq

MULTIVARIATE ANALYSIS WITH GENOMIC AND TRANSCRIPTOMIC DATA FOR MEAT FATTY ACID PROFILE IN NELORE CATTLE FINISHED IN FEEDLOT

ABSTRACT – The understanding of regulatory processes and molecular organization of organisms has progressed considerably in the last 10 years. The methodologies also evolved with the sequencing of DNA, RNA and genomic tools allowing the analysis a lot of genes, proteins or metabolites. The simultaneous use of this information should help to obtain relevant information about the variables that result the phenotypic variations of traits of interest. The objective of the present study was to integrate phenotypic, genotypic and transcriptomic studies in order to clarify the genetic and metabolic mechanisms that determine the fatty acid profile in Nelore beef, in order to contribute to the improvement of the fatty acid composition of the meat. Nelore males were used in feedlot, coming from farms that integrate three breeding programs and slaughtered with an average of 24 months. Samples of the *L. thoracis* muscle between the 12th to 13th rib were collected for analysis of fatty acid profile, RNA and DNA extraction. The results were presented in chapters 2 and 3. In chapter 2, the objective was to identify genes differentially expressed by RNA-seq method and fatty acid profile in the *L. thoracis* muscle with the use of Principal Components (PC). Two groups of 10 animals were selected, which had PC1 and PC2 extreme values (High x Low) for the fatty acids (FA) groups: saturated fatty acids (SFA), monounsaturated fatty acids (MUFA), polyunsaturated fatty acids (PUFA), omega 3 (n-3) and omega 6 (n-6)) for analysis of differentially expressed genes - GDE ($p < 0.10$). The SFA group, 33 DEG for PC1 and no DEG for PC2 were found; for MUFA were found 56 DEG for PC1 and 10 DEG for PC2; PUFA presented only two DEG for PC1 and none for PC2; for omega-3 no DEG found for PC1 and none for PC2; already for Omega 6, seven DEG for PC1 and three DEG for PC2 were found. Functional enrichment analysis was performed by DAVID v. 6.8. Among the DEG, genes about the energy metabolism (*UCP3*), adipogenesis (*MBNL1*), fat metabolism or catabolism (*MVD*, *PCYT2*, *LIPE* and *EPHX2*) have been identified that play a role in cellular lipid Nelore breed. In chapter 3, the objective was to use phenotypic, genotypic and transcriptomic data to investigate causal networks for the profile of fatty acids in the meat of Nelore animals. Four genomic regions showed significant associations ($p < 0.05$) with seven phenotypes (C16: 0, C16: 1, C18: 0, C18: 1c9, C18: 2, CLA, C22: 6) and also with expression of 30 genes, including *SREBF1*, *MDE9*, *ACSF2*, related to lipid metabolism and catabolism; and *ABCC3*, *ADCY1*, *OGDH*, energy metabolism. The present study showed a causal relationship between the fatty acid profile expression and the metabolic processes that involve the regulation of such expression, demonstrated causal effects among the found DEG. The results should contributed to the knowledge of the biological mechanisms involved in the fatty acid profile in beef cattle. In the future, it may be to use DEG found as genetic markers in Nelore cattle to select the lipid composition that contributes to better nutrition and consumer health.

Keywords: *Bos Taurus indicus*, genotype, intramuscular fat, metabolic pathways, RNA – seq

LISTA DE ABREVIACÕES

AG – Ácidos graxos

AGS – Ácidos graxos saturados

AGI – Ácidos graxos insaturados

AGMI – ácidos graxos monoinsaturados

AGPI – Ácidos graxos poliinsaturados

HDL – *high density lipoprotein* – lipoproteína de alta densidade

LDL – *low density lipoprotein* - lipoproteína de baixa densidade

CLA – *conjugated linoleic acid* – ácido linoleico conjugado

NGS - *new generation sequencing* - sequenciamento de nova geração

FDR – *false discovery rate* – taxa de falsos positivos

GDE – genes diferencialmente expressos

PCA – *principal componentes analysis* – análise de componentes principais

PCs – *principal components* - componentes principais

LD – *linkage disequilibrium* – desequilíbrio de ligação

CAPÍTULO 1 – CONSIDERAÇÕES GERAIS

INTRODUÇÃO

O crescimento socioeconômico da população aumentou o consumo de alimentos de origem animal, o que demanda a pecuária buscar por novas estratégias de produção animal, a fim de incrementar sua produção. No ano 2017, em decorrência dos investimentos e melhorias na produção de bovinos, o Brasil conquistou a posição de segundo maior produtor mundial de carne bovina (USDA, 2017). A demanda por carnes com características organolépticas de alta qualidade e mais saudáveis exigem o desenvolvimento de estratégias de melhoramento e ferramentas que possibilitem melhorar a qualidade da carne (Cánovas et al., 2014; Carrillo et al., 2016).

A carne bovina é um alimento altamente nutritivo, fonte de proteínas, lipídios, vitaminas do complexo B e minerais essenciais. O tipo de gordura e a proporção de ácidos graxos na carne ganham importância, pois afetam características que demonstram a qualidade do produto e a saúde do consumidor. A gordura pode apresentar composição indesejável, com alta porcentagem de ácidos graxos saturados (AGS). O consumo excessivo de gordura *trans* e de AGS aumenta os níveis séricos de colesterol e lipoproteínas de baixa densidade (LDL) na corrente sanguínea, ocasionando o aumento de doenças coronárias, que são consideradas a maior causa de morte mundial (Katan et al., 1994; Wang et al., 2016). Por outro lado, os ácidos graxos insaturados (AGI), quando consumidos de maneira equilibrada, trazem benefícios à saúde humana, uma vez que podem diminuir os níveis de colesterol e LDL e aumentar as lipoproteínas de alta densidade (HDL) (Tapiero et al., 2002).

Metodologias de sequenciamento de DNA e RNA tornaram-se relevantes para explicar os mecanismos relacionados a características quantitativas complexas, tais como o perfil de ácidos graxos da carne, cuja expressão é resultado dos efeitos genéticos e ambientais somado a interação entre eles (De Smet et al., 2004). Além disso, informações genômicas e transcriptômicas somadas à informação fenotípica dos animais aumentam o poder de predição e garantem a seleção de animais superiores para as características alvo de seleção.

REVISÃO DE LITERATURA

O Brasil se encontra na posição de segundo maior produtor mundial de carne bovina, produzindo anualmente aproximadamente 10 milhões de toneladas de carne (USDA, 2017). A fim de aumentar e melhorar a produtividade para atender mercado interno e externo e suas respectivas exigências, busca-se em programas de melhoramento genético selecionar animais superiores para características de interesse econômico. Em bovinos de corte, as características associadas a qualidade da carne são importantes para o fornecimento de um produto final que atenda as expectativas do consumidor. Além do valor nutricional, que depende da composição da carne (Dekkers, 2012), a maciez, suculência e sabor são atributos apreciados pelos consumidores (Maltin et al., 2003). As características de qualidade da carne, como sabor e maciez, são influenciadas pelo quantidade e tipo de gordura presente na carne, conseqüentemente pela gordura intramuscular (Wood et al., 1999).

O perfil dos ácidos graxos (AG) presente na gordura intramuscular é influenciado por fatores ambientais (dieta e manejo) e genéticos (raça e polimorfismos) (De Smet et al., 2004). Segundo a Associação Brasileira de Criadores de Zebu (ABCZ), estima-se que 80% dos bovinos brasileiros tenham influência genética de raças zebuínas (*Bos taurus indicus*), onde a raça Nelore compõe o maior rebanho bovino comercial no Brasil (Ferraz; Felício, 2010). Quando comparados a animais de raças taurinas e cruzados (*Bos taurus taurus* x *Bos taurus indicus*), as raças zebuínas apresentam tecido adiposo com menor proporção de ácidos graxos saturados (AGS) (Huerta-Leidenz et al., 1993; Pitchford et al., 2002; Menezes et al., 2009; Rossato et al., 2010, Bressan et al., 2011; Sevane et al., 2014; Cesar et al., 2014), os quais estão associados ao aumento do risco de doenças cardíacas (Katan et al., 1994; Wang et al., 2016; Benjamin et al., 2019)

Os ácidos graxos podem ser benéficos para a saúde humana, tais como os ácidos graxos poli-insaturados (AGPI), que se encontram em menores concentrações que os AGS (Metz et al., 2009; Feitosa et al., 2016). Os AGPI estão relacionados a redução das concentrações de lipoproteínas de baixa densidade (“colesterol ruim”) na corrente sanguínea, e podem reduzir o risco da ocorrência de outras doenças, como o câncer de mama (Binukumar; Mathew, 2005). Estudos comprovaram alterações na proporção e composição dos AG, mediante estratégias na alimentação em bovinos (Rossato et al., 2010; Costa et al., 2012; Costa, 2013; Barducci et al., 2016; Bressan

et al., 2016), possibilitando a obtenção de carne mais saudável para o consumidor (Ladeira 2014).

Os avanços tecnológicos na biologia molecular durante os últimos anos possibilitaram aprimorar o entendimento dos mecanismos que envolvem as características complexas. Tendo em vista a importância dos AG na saúde humana, estudos relacionados à processos moleculares do metabolismo de gordura tornaram-se cruciais na produção de carne de alta qualidade, e conseqüentemente mais saudáveis (Huang et al., 2017). O sequenciamento de nova geração (*new generation sequencing* - NGS) fornece informações para uma visão mais detalhada sobre os sistemas biológicos com relação a DNA/RNA e, além de caracterização e quantificação das “ômicas” (genômica, transcriptômica e epigenômica), permite explorar a relação entre a diversidade genética e fenotípica (Sharma et al., 2017). A exploração das redes gênicas, vias biológicas e regulatórias, relacionadas com os ácidos graxos da carne bovina fornecerão embasamento para melhorar a seleção dos bovinos.

Ácidos graxos na carne bovina

A carne bovina é um alimento altamente nutritivo, fonte de micronutrientes e minerais essenciais a alimentação humana (ferro, selênio, zinco e vitamina B12). A carne é composta de músculo, tecido adiposo, osso, tecido conjuntivo, nervos e vasos sanguíneos. O que atrai mais atenção dos nutricionistas são os músculos e tecido adiposo (Wood et al., 2017), pois as propriedades nutricionais e qualitativas da carne estão relacionadas ao teor de gordura intramuscular (IMF – do inglês *intramuscular fat*) e a sua composição de ácidos graxos (AG) (Wood et al., 2008). O perfil de ácidos graxos da IMF é relevante para a saúde humana, uma vez que esse tipo de gordura não pode ser removida ou extraída antes do consumo da carne (De Smet et al., 2004). Para padrões internacionais da qualidade de carne, a quantidade de IMF depositada no músculo *Longissimus* (marmorização) é dada como preditor de palatabilidade e é o principal determinante do valor da carcaça (Ferraz; Felício, 2010).

O perfil dos AG influencia diretamente os aspectos sensoriais da carne como sabor, suculência, maciez, aroma e o valor nutricional final do produto. Existe

interesse em manipular a composição da carne bovina por ser uma rica fonte de proteína, gordura e ácidos graxos essenciais (não produzidos pelo organismo, adquirido na alimentação). No entanto, possui alto teor de ácidos graxos saturados (AGS), que são indesejados para a saúde do consumidor. Os AGS estão associados a ocorrência de doenças cardiovasculares e obesidade, ocasionadas pelo aumento nos níveis de lipoproteínas de baixa densidade (LDL) e alto colesterol no plasma sanguíneo. Entretanto, o consumo de carne com maior proporção de ácidos graxos monoinsaturados (AGMI), como o ácido oleico (C18:1, n9), e poli-insaturados (AGPI) como os ácidos linoleico (C18:2n6), linolênico (C18:3n3) e o linoleico conjugado (CLA) podem trazer efeitos benéficos para a saúde humana (Siciliano et al., 2013). Os ácidos graxos insaturados (AGI) citados oferecem proteção ao sistema cardiovascular, pois diminuem os níveis séricos de colesterol e aumentam as lipoproteínas de alta densidade, como o ácido graxo linoleico conjugado (CLA) que pode aumentar a atividade imunoestimulante, antimutagênica e antioxidante (Ip, 1997). Estudo realizado por Zong et al. (2016) sobre consumo dos AGS predominantes na carne bovina (láurico - C12:0, mirístico - C14:0, palmítico - C16:0 e oleico - C18:0), identificou associação dos AGS ao risco de ocorrência de doenças cardiocoronárias, o qual é significativamente menor quando substituída a soma destes quatro AGS por gorduras poliinsaturadas.

Como a maioria das características de interesse econômico, a formação e composição de gordura em bovinos e outros ruminantes é um processo complexo e multifatorial (Ladeira et al., 2016). Estudos tem sido realizados para caracterizar e quantificar AG em diferentes raças bovinas, a fim de agregar informações que tragam melhorias genéticas para composição de AG em bovinos de corte. A maioria dos estudos foram realizados em raças taurinas (Warren et al., 2008; Nuernberg et al., 2005; Realini et al., 2004; Cecchinato et al., 2012; Ishii et al., 2013; Saatchi et al., 2013; Zhu et al., 2017). Poucos estudos foram realizados na premissa de estimar a variabilidade genética para a composição de AG da carne em animais de raças zebuínas, em especial a raça Nelore (Cesar et al., 2014; Aboujaoude et al., 2016; Feitosa et al., 2016; Lemos et al., 2016). Além disso, bovinos da raça Nelore e demais raças zebuínas tendem a ser produzidas no sistema a pasto no Brasil, o que influencia a composição química da carne e resulta na menor ocorrência de AGS, menor

porcentagem de gordura e proporciona um perfil de AG mais benéfico à saúde humana, com maiores concentrações dos ácidos graxos esteárico e linolênico (Realini et al., 2004; Ladeira; Oliveira, 2006)

Variabilidade genética do perfil de ácidos graxos na carne

Compreender as bases genéticas de características que se expressam sob efeito de fatores genéticos e ambientais, além de buscar por informações e resultados que contribuem o entendimento sobre o que modula o metabolismo, tem sido os objetivos de muitos pesquisadores (Ritchie et al., 2015; Aboujaoude et al., 2016; Feitosa et al., 2016; Cesar et al., 2016; Lemos et al., 2016; Berton et al, 2016). Diante da dificuldade de seleção para características complexas, como o perfil de ácidos graxos, são realizadas avaliações genéticas a fim de estimar valores genéticos para as características de interesse, e então os animais são selecionados de acordo com os seus valores genéticos estimados.

Em relação a características de qualidade da carne, estudos moleculares e fisiológicos têm sido realizados, principalmente para deposição de gordura intramuscular. Entretanto, os mecanismos que regulam o perfil de ácidos graxos em ruminantes têm muito a ser investigado. Mudanças no perfil de ácidos graxos na carne bovina podem ser provocadas por estratégias na alimentação. Entretanto, os fatores genéticos que afetam a composição dos ácidos graxos em bovinos têm sido menos investigados, apesar dos estudos apontarem para diferenças entre raças quanto a composição de ácidos graxos (Rossato et al., 2009; Dinh et al., 2010; Sevane et al., 2014; Lemos et al., 2018).

A avaliação fenotípica dos AG é de difícil mensuração e onerosa, e na maioria dos casos é avaliada somente após o abate dos animais. Neste sentido, a seleção genômica, bem como a identificação de genes candidatos e seus mecanismos de ação, vias de regulação e redes gênicas envolvidas no perfil de ácidos graxos na carne bovina fornecerão subsídios para aprimorar a seleção. Estudos realizados por Inoue et al. (2011, 2017) e Nogi et al. (2011) estimaram parâmetros genéticos para perfil de ácidos graxos em bovinos da raça Wagyu. As estimativas de herdabilidade relatadas nestes trabalhos variaram entre 0,01 a 0,82. Cesar et al. (2014) estimaram parâmetros

genéticos para o perfil de ácidos graxos da IMF do músculo *Longissimus thoracis* em 386 bovinos da raça Nelore. Os autores descreveram que estimativas de herdabilidade foram de baixas ($<0,10$) a moderadas (0,29) para os AG individuais. Aboujaoude et al. (2016) ao estimarem parâmetros genéticos para perfil de AG da IMF do músculo *Longissimus thoracis* em 963 touros da raça Nelore terminados em confinamento, obtiveram herdabilidades entre 0,01 a 0,35. Esses resultados sugerem mudanças na composição de ácidos graxos em bovinos por meio de estratégias como seleção. Esta informação é importante para os programas de melhoramento que visam aprimorar a composição de ácidos graxos na carne de bovinos.

A partir da publicação da primeira versão do genoma bovino em 2009 (Lsik et al., 2009), o desenvolvimento de ferramentas genômicas para pesquisas genéticas em bovinos trouxe a possibilidade de uma nova maneira de seleção animal. Estudos com uso de *microarrays* de marcadores moleculares do tipo polimorfismo de nucleotídeo único (SNP) possibilitaram investigações em nível molecular e posteriormente a associação de polimorfismos no DNA com as características de interesse. Dentre os marcadores genéticos, o SNP é amplamente utilizado, pois apresenta menores taxas de mutação, é de fácil genotipagem e são amplamente distribuídos no genoma (Youngerman et al., 2004).

Com o sequenciamento do genoma bovino tornou-se possível a identificação de SNP. Os estudos de associação genômica ampla (GWAS) possibilitaram a busca por regiões genômicas, redes gênicas, e informações de algumas vias metabólicas associadas às características complexas de interesse econômico (Matukumalli et al., 2009). O sequenciamento de DNA e RNA permite a identificação de potenciais biomarcadores potenciais para prever possíveis informações genéticas que os animais poderão compartilhar com sua descendência.

Pesquisas tem indicado regiões gênicas associadas a características de qualidade da carne. Em estudo de associação genômica ampla (GWAS), com animais Nelore, Cesar et al. (2014) encontraram diversas regiões genômicas com genes associados à deposição e a composição de gordura intramuscular. Lemos et al. (2016) avaliaram o perfil de AG em bovinos da raça Nelore, e identificaram 115 diferentes regiões genômicas, que explicavam mais de 1% da variância genética aditiva para os 22 ácidos graxos estudados. Entre as centenas de genes candidatos

identificados por Lemos et al. (2016); alguns como *PLCH1*, *ELOVL5*, *ESSRG*, *PCYT1A* e da família *ABC* (*ABC5*, *ABC6* e *ABC10*) são genes associados ao metabolismo lipídico e energético.

O perfil de expressão gênica pode ser utilizado para identificar genes e isoformas diferencialmente expressas relacionadas a redes que controlam características quantitativas. O perfil de expressão gênica permite ainda direcionar caminhos sobre como funcionam os mecanismos moleculares responsáveis pelas variações fenotípicas (Tizioto et al., 2015). Desta forma, localizar no genoma regiões que estejam influenciando a composição desses AG auxiliará na seleção de animais que favorecerão a produção de alimentos com menores teores de gorduras indesejáveis. Posto isto, fica clara a necessidade de desenvolver estratégias que tragam melhorias à composição de AG e suas proporções na carne bovina por meio da seleção de animais com os genótipos desejáveis.

Estudos de transcriptômica aplicados ao melhoramento do perfil de ácidos graxos

Desde a descoberta do papel do RNA como intermediário-chave entre o genoma e o proteoma, o sequenciamento de RNA (RNA-Seq) tornou-se protocolo padrão e fundamental na investigação do transcriptoma e tem sido aplicado para estudo de diversos organismos e características (Li; Jiang, 2012). A análise do RNA-Seq possibilita determinar o nível de mRNA de um organismo para estimar seu perfil de expressão gênica em um determinado momento (Carrilo et al., 2016). O sucesso de tal metodologia reside no fato de que em um único estudo, há a possibilidade de busca pela expressão gênica e sua quantificação em larga escala.

Estudos apresentados demonstraram que o uso da técnica de RNA-seq tem sido vantajoso na investigação de expressões gênicas, aprimoramento de mecanismos genéticos regulatórios, futura seleção de animais a partir da expressão de genes específicos. Lee et al. (2013) investigaram a diferença da expressão específica de deposição de gordura de três tecidos adiposos diferentes em bovinos e encontraram mais de cinco mil genes diferencialmente expressos (GDE). Li et al. (2015) realizaram estudo de associação de geômica ampla (GWAS) a partir de dados transcriptômicos da parede ruminal de bovinos Angus, submetidos a diferentes

manejos alimentares (capim x grãos) e identificaram 342 GDE e vias metabólicas que podem afetar o desenvolvimento e a qualidade da carne. Carrilo et al. (2016) analisaram a relação da qualidade e composição da carne com o tipo de alimentação fornecida aos animais Angus na fase de terminação. Utilizando perfis de transcriptoma, os mesmos autores, reconheceram 241 genes GDE, e entre as vias encontradas a fosforilação oxidativa que está diretamente relacionada a qualidade de carne. Berton et al. (2016), também com o uso de RNAseq e avaliação do perfil de ácidos graxos em bovinos Nelore, criados em confinamento, identificaram mais de mil DEG. Estes, relacionados com metabolismo lipídicos e de ácidos graxos; como os genes: *ACSM3* e *ACSS1*, precursores de síntese de AG; e *DGAT2* que atua na deposição de gordura saturada no tecido adiposo.

Ramayo-Caldas et al. (2012) relataram 55 DEG em dois grupos extremos de perfil de ácidos graxos em suínos cruzados (Iberian x Landrace); dentre estes, genes que afetam o metabolismo de lipídios e ácidos graxos, como *ABCG8*, *CYP2C9*, *CYP2C19*, *APO*, *CYP7A1*, *MTMR7*, *ME3*. Buchanan et al. (2013) avaliaram expressão gênica em novilhas em diferentes sistemas de terminação (pasto e confinamento) para gordura intramuscular e perfil de ácidos graxos; e encontraram genes envolvidos no metabolismo de ácidos graxos como o *DGAT2*, *FASN*, *PPAR γ* , *GPAM*, e *ADIPOQ*. Os estudos citados identificaram genes que apresentam expressão diferencial, portanto podem fornecer marcadores genéticos para a seleção para o perfil de ácidos graxos desejáveis em futuros programas de seleção.

Análise de componentes principais

Conjuntos de dados de grande dimensão são frequentes nas mais diversas áreas, e isto pode ser uma dificuldade para interpretação desses dados. A análise de componentes principais (*Principal component analysis - PCA*) (Pearson, 1901; Jolliffe, 1986) torna-se uma alternativa a fim de reduzir a dimensão dos dados sem permitir que ocorra perda de informações relevantes, preservando o máximo possível da variabilidade. O PCA permite quantificar os principais fatores que contribuem para a variabilidade de determinada característica e a partir de variáveis correlacionadas, cria-se novas variáveis (componentes principais) que não são correlacionados.

Essa metodologia tem sido utilizada em estudos de melhoramento genético para reduzir o tamanho da matriz de covariância genética aditiva quando utilizados modelos multicausal (Pinto et al., 2006; Buzanska et al., 2013; Boligon et al., 2016). Em estudos de RNA-seq, Ramayo-Caldas et al. (2012) e Merino et al. (2016) realizaram PCA para controle de qualidade dos dados, a fim de obter representação de menor proporção, e que descrevesse a maior variação possível do conjunto de variáveis originais. A fim de entender as correlações entre as características, e então reduzir o número de características analisadas, Pinto et al. (2006) aplicaram o PCA para características de performance e carcaça de frangos, para fins reprodutivos e genéticos.

A PCA tem sido utilizada para elucidar as relações estruturais entre diferentes características e eliminar informações redundantes de uma população. Assim, pode-se usar essa técnica multivariada para reduzir o tamanho de um conjunto de variáveis originais, sem perder informações importantes, mantendo a maior representação de variação presente no conjunto de dados iniciais (Yeung e Ruzzo, 2001). Isso é possível através da transformação em um novo conjunto de variáveis, os componentes principais (PCs), que não estão correlacionados, e que são ordenados de modo que os primeiros retêm a maior parte da variação presente em todas as variáveis originais (Jolliffe, 1986). Estudos de PCA com dados de microarranjos permitem visualizar a versatilidade do agrupamento por componentes principais, pois são utilizados a fim de avaliar as relações entre os arranjos ou para identificar genes com perfis de expressão semelhantes.

Ramayo-Caldas et al. (2012) com dados de RNA sequenciado e perfil de ácidos graxos da IMF em suínos, realizaram PCA e os dois primeiros componentes principais explicaram aproximadamente 50% de toda variação fenotípica das características em questão. Estudos de PCA com dados genômicos transcriptômicos ainda são escassos na literatura. Contudo, a elevada quantidade de informações geradas em pesquisas desse gênero e a complexidade de análise devido demanda computacional, o uso da PCA torna-se útil e eficaz, quando reduz o número de variáveis e sem grandes perdas de informações e da variabilidade.

Estruturas Causais

As relações entre as características fenotípicas de um animal são comumente estudadas pela genética quantitativa mediante utilização de modelos estatísticos, como os modelos multicaracterísticas. Ao utilizar os modelos multicaracterísticas em programas de melhoramento animal, consideram-se as correlações genéticas entre as características, e assim consegue-se aumentar a acurácia da avaliação. Ao estimar parâmetros genéticos simultaneamente para características correlacionadas, informações mais confiáveis, facilita as decisões para as seleções. Embora tais modelos possam ser utilizados satisfatoriamente para inferir como são os prováveis eventos, eles não são estáveis o suficiente para prever como as probabilidades mudariam por resultado de intervenções externas (ambientais). Modelos de Equação Estrutural (MEE) são usados para estudar relações recursivas e simultâneas entre fenótipos em sistemas multivariados. Nos MEEs uma característica pode ser considerada como um preditor de outra, fornecendo umnexo causal entre elas (Rosa et al., 2011).

A construção de redes permite integrar dados fenotípicos com informação genotípica e perfil transcricional, com potencial para descobrir grupos de genes envolvidos na regulação da característica desejada e a uma rede de interações entre os efeitos de genes. Nos últimos anos, as redes bayesianas têm sido utilizadas em muitos campos, desde a análise do desempenho do serviço médico (Acid et al., 2004), a análise da expressão gênica (Friedman et al. 2000), prognóstico do câncer de mama e epidemiologia (Nagl et al., 2008). As redes bayesiana são modelos gráficos para relacionamentos probabilísticos entre um conjunto de variáveis. Na última década, a rede bayesiana tornou-se uma representação popular para a codificação de conhecimento especializado incerto em sistemas especialistas (Heckerman et al., 1995a). Mais recentemente, pesquisadores desenvolveram métodos para aprender redes bayesianas a partir de dados. As técnicas que foram desenvolvidas são novas e continuam evoluindo, mas demonstraram ser notavelmente eficazes para alguns problemas de análise de dados.

A alta dimensionalidade dos conjuntos de dados comuns nesses domínios levou ao desenvolvimento de vários algoritmos de aprendizado focados na redução

da complexidade computacional, auxiliando a compreensão do comportamento de características complexas.

Morota et al. (2012) avaliaram o desequilíbrio de ligação (LD) em bovinos da raça Holandesa usando uma rede bayesiana. Os autores detectaram marcadores condicionalmente dependentes, tais resultados confirmam que as relações LD são de natureza multivariada. Peñagaricano et al. (2015) explorando em suínos mestiços associações marginais entre genótipos, fenótipos e dados transcriptômicos; e utilizando construção de redes causais, encontrou uma região genômica que apresentava associações significativas entre três fenótipos (peso do músculo lombar, gordura subcutânea na 10ª costela e porcentagem média de gordura intramuscular) com 7 genes expressos.

As características de produção em bovinos são quantitativas, envolvem múltiplos genes e suas interações. Ao explorar redes gênicas por meio de redes bayesianas pode ajudar a identificar grupos de genes diferencialmente expressos (GDE) e regiões reguladoras para os ácidos graxos. Melhorar a compreensão sobre redes gênicas, vias biológicas e regulatórias envolvidas no perfil de ácidos graxos da carne bovina, fornecerá mais conhecimento para definir novas estratégias de seleção e melhorar a gestão para seleção de bovinos.

OBJETIVOS

O objetivo deste estudo foi (i) buscar representatividade nos componentes principais (PC) afim de substituir as variáveis fenotípicas (ácidos graxos); e então buscar genes diferencialmente expressos (GDE) com uso de RNA-seq para perfil de ácidos graxos no músculo *Longissimus thoracis*, (ii) investigar redes causais (*Bayesian Network*) para obter informações que possam auxiliar no entendimento dos mecanismos genéticos e metabólicos relacionados ao perfil de AG em bovinos da raça Nelore.

REFERÊNCIAS BIBLIOGRÁFICAS

Aboujaoude C, Pereira ASC, Feitosa FLB, Lemos MVA, Chiaia HLJ, Berton MP, Peripolli E, Silva RMO, Ferrinho AM, Mueller LF, Olivieri BF, Albuquerque LG, Oliveira HN, Tonhati H, Espigolan R, Tonussi R, Gordo DM, Magalhães AFB, Baldi F (2016) Genetic parameters for fatty acids in intramuscular fat from feedlot-finished Nelore carcasses. **Animal Production Science** 58(2):1-14.

Acid S, De Campos LM, Fernández-Luna JM, Rodríguez S, Rodríguez JM, Salcedo JL (2004) A comparison of learning algorithms for Bayesian networks: a case study based on data from an emergency medical service. **Artificial Intelligence in Medicine** 30:215–232.

Barducci RS, Franzói MCS, Sarti LMN, Millen DD, Putarov TC, Perdigão A, Martins CL, Arrigoni MDB (2016) Perfil de ácidos graxos e características da carne de bovinos Nelore confinados com diferentes fontes lipídicas protegidas. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia** 68(1):233–242.

Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, Chamberlain AM, Chang AR, Cheng S, Das SR, Delling FN, Djousse L, Elkind MSV, Ferguson JF, Fornage M, Jordan LC, Khan SS, Kissela BM, Knutson KL, Kwan TW, Lackland DT, Lewis TT, Lichtman JH, Longenecker CT, Loop MS, Lutsey PL, Martin SS, Matsushita K, Moran AE, Mussolino ME, O'Flaherty M, Pandey A, Perak AM, Rosamond WD, Roth GA, Sampson UKA, Satou GM, Schroeder EB, Shah SH, Spartano NL, Stokes A, Tirschwell DL, Tsao CW, Turakhia MP, VanWagner LB, Wilkins JT, Wong SS, Virani SS (2019) Heart disease and stroke statistics – 2019 update: A report from the American Heart Association. **Circulation** 139:e56-e528.

Berton MP, Fonseca LFS, Gimenez DFJ, Utembergue BL, Cesar ASM, Coutinho LL, Lemos MVA, Aboujaoude C, Pereira ASC, Silva RMO, Stafuzza NB, Feitosa FLB, Chiaia HLJ, Olivieiri BF, Peripolli E, Tonussi RL, Gordo DM, Espigolan R, Ferrinho AM, Mueller LF, Albuquerque LG, Oliveira HN, Duckett S, Baldi F (2016) Gene expression profile of intramuscular muscle in Nelore cattle with extreme values of fatty acid. **BMC Genomics** 17:972.

Binukumar B, Mathew A (2005) Dietary fat and risk of breast cancer. **World Journal of Surgical Oncology** 7:1–7.

Bolignon AA, Vicente IS, Vaz RZ, Campos GS, Souza FRP, Carvalheiro R, Albuquerque LG (2016) Principal component analysis of breeding values for growth and reproductive traits and genetic association with adult size in beef cattle. **Journal of Animal Science** 94(12):5014–5022.

Bouroche JM e Saporta G. Análise de dados. Rio de Janeiro: Zahar, 1982.

Bressan MC, Rossato LV, Rodrigues EC, Alves SP, Bessa RJB e Ramos EM. (2011) Genotype × environment interactions for fatty acid profiles in *Bos indicus* and *Bos taurus* finished on pasture or grain. **Journal of Animal Science** 89:221–232.

Bressan MC, Rodrigues EC, De Paula ML, Ramos EM, Portugal PV, Silva JS, Bessa RB, Gama LT (2016) Differences in intramuscular fatty acid profiles among *Bos indicus* and crossbred *Bos taurus* × *Bos indicus* bulls finished on pasture or with concentrate feed in Brazil. **Italian Journal of Animal Science** 15:10–21.

Buchanan JW, Garmyn AJ, Hilton GG, Vanoverbeke DL, Duan Q, Beitz DC, Mateescu RG (2013) Comparison of gene expression and fatty acid profiles in concentrate and forage finished beef. **Journal of Animal Science** 91(1):1–9.

Buzanskas ME, Savegnago RP, Grossi DA, Venturini GC, Queiroz SA, Silva LOC, Torres Júnior RAA, Munari DP, Alencar MM Genetic parameter estimates and principal component analysis of breeding values of reproduction and growth traits in female Canchim cattle. **Reproduction, Fertility and Development** 25(5):775–781.

Cánovas A, Reverter A, DeAtley KL, Ashley RL, Colgrave ML, Fortes MRS, Islas-Trejo A, Lehnert S, Porto-Neto L, Rincón G, Silver GA, Snelling WM, Medrano JF, Thomas MG (2014) Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. **PLoS ONE** 9(7):e102551.

Carrillo JA, He Y, Li Y, Liu J, Erdman RA, Sonstegard TS, Song J (2016) Integrated metabolomics and transcriptome analyses reveal finishing forage affects metabolic pathways related to beef quality and animal welfare. **Scientific Reports** 6:25948.

Cecchinato A, De Marchi M, Penasa M, Casellas J, Schiavon S, Bittante G (2012) Genetic analysis of beef fatty acid composition predicted by near-infrared spectroscopy. **Journal of Animal Science** 90:429–438.

Cesar ASM, Regitano LCA, Mourão GB, Tullio RR, Lanna DPD, Nassu RT, Mudado MA, Oliveira PSN, Nascimento ML, Chaves AS, Alencar MM, Sonstegard TS, Garrick DJ, Reecy JM, Coutinho LL (2014) Genome-wide association study for intramuscular fat deposition and composition in Nelore cattle. **BMC Genetics** 15:39.

Cesar ASM, Regitano LCA, Poleti MD, Andrade SCS, Tizioto PC, Oliveira PSN, Felício AM, Do Nascimento ML, Chaves AS, Lanna DPD, Tullio RR, Nassu RT, Koltjes JE,

Fritz-Waters E, Mourão GB, Zerlotini-Neto A, Reecy JM, Coutinho LL (2016) Differences in the skeletal muscle transcriptome profile associated with extreme values of fatty acids content. **BMC Genomics** 17:961.

Costa ASH, Lopes PA, Estevão M, Martins SV, Alves SP, Pinto RMA, Pissarra H, Correia JJ, Pinho M, Fontes CMGA, Prates JAM (2012) Constrasting cellularity and fatty acid composition in fat depots from Alentejana and Barrosã bovine breeds fed high and low forage diets. **International Journal of Biological Sciences** 8(2):214-227.

Costa ASH (2013) **Genetic and metabolic regulation of fatty acid deposition in autochthonous bovine breeds with distinct genetic background**. 196 f. Tese (Doutorado em Ciências Veterinárias) – Universidade de Lisboa.

De Smet S, Raes K, Demeyer D (2004) Meat fatty acid composition as affected by fatness and genetic factors: a review. **Animal Research** 53:81–98.

Dekkers JCM (2012) Application of Genomics Tools to Animal Breeding. **Current Genomics** 13:207-212.

Dinh TTN, Blanton Jr. JR, Riley DG, Chase Jr. CC, Coleman SW, Phillips WA, Brooks JC, Miller MF, Thompson LD (2010) Intramuscular fat and fatty acid composition of longissimus muscle from divergent pure breeds of cattle. **Journal of Animal Science** 88(2):756–766.

Feitosa FLB, Olivieri BF, Aboujaoude C, Pereira ASC, Lemos MVA, Chiaia HLJ, Berton MP, Peripolli E, Ferrinho AM, Mueller LF, Mazalli MR, Albuquerque LG, Oliveira HN, Tonhati H, Espigolan R, Tonussi RL, Silva RMO, Gordo DGM, Magalhães AFB, Aguilar I, Baldi F (2016) Genetic correlation estimates between beef fatty acid profile with meat and carcass traits in Nelore cattle finished in feedlot, **Journal of Applied Genetics** 58:123-132.

Ferraz JBS, Felício PE (2010) Production systems: an example from Brazil. **Meat Science** 84:238-243.

Friedman N, Linial M, Nachman I, Pe'er D (2000) Using Bayesian Networks to Analyze Expression Data. **Journal of Computational Biology** 7(3-4):601–620.

Heckerman, D. (1998). **A Tutorial on Learning with Bayesian Networks. Learning in Graphical Models** 301–354. doi:10.1007/978-94-011-5014-9_11

Huang TY, Zheng D, Houmard JA, Brault JJ, Hickner RC, Cortright RN (2017) Overexpression of PGC-1 α increases peroxisomal activity and mitochondrial fatty acid oxidation in human primary myotubes. **American Journal of Physiology Endocrinology and Metabolism** 312(4):E253-E263.

Huerta-Leidenz NO, Cross HR, Savell JW, Lunt DK, Baker JF, Pelton LS, Smith B. (1993) Comparison of the fatty acid composition of subcutaneous adipose tissue from mature Brahman and Hereford cows. **Journal of Animal Science** 71:625-630.

Inoue K, Kobayashi M, Shoji N, Kato K (2011) Genetic parameters for fatty acid composition and feed efficiency traits in Japanese Black cattle. **Animal** 5:987–994.

Inoue N, Shoji N, Honda T, Oyama K (2017) Genetic relationships between meat quality traits and fatty acid composition in Japanese black cattle. **Animal Science Journal** 88(1):11-18.

Ip C (1997) Review of the effects of trans fatty acids, oleic acid, n-3 polyunsaturated fatty acids, and conjugated linoleic acid on mammary carcinogenesis in animals. **American Journal of Clinical Nutrition** 66:1523-1529.

Ishii A, Yamaji K, Uemoto Y, Sasago N, Kobayashi E, Kobayashi N, Matsushashi T, Maruyama S, Matsumoto H, Sasazaki S, Mannen H (2013) Genome-wide association study for fatty acid composition in Japanese Black cattle. **Animal Science Journal** 84(10):675-682.

Jolliffe IT (1986) Principal component analysis. New York: Springer, 488p.

Katan MB, Zoock PM, Mensink RP. (1994) Effects of fats and fatty acids on blood lipids in humans: an overview. **American Journal for Clinical Nutrition** 60:1017-1022.

Ladeira MM, Oliveira RL (2006) Estratégias nutricionais para melhoria da carcaça bovina. In: II SIMBOI - SIMPÓSIO SOBRE DESAFIOS E NOVAS TECNOLOGIAS NA BOVINOCULTURA DE CORTE. **Anais...** Brasília: Faculdades Integradas – UPIS.

Ladeira MM, Santarosa LC, Chizzotti ML, Ramos EM, Neto ORM, Oliveira DM, Carvalho JRR, Lopes LS, Ribeiro JS (2014) Fatty acid profile, color and lipid oxidation of meat from young bulls fed ground soybean or rumen protected fat with or without monensin. **Meat Science** 96(1):597–605.

Ladeira MM, Schoonmaker JP, Gionbelli MP, Dias JCO, Gionbelli TRS, Carvalho JRR e Teixeira PD (2016) Nutrigenomics and Beef Quality: A Review about Lipogenesis. **International Journal of Molecular Sciences** 17:918.

Lee HJ, Jang M, Kim H, Kwak W, Park W, Hwang JY, Lee CK, Jang GW, Park MN, Kim HC, Jeong JY, Seo KS, Kim H, Cho S, Lee BY (2013) Comparative Transcriptome Analysis of Adipose Tissues Reveals that ECM-Receptor Interaction Is Involved in the Depot-Specific Adipogenesis in Cattle. **PLoS ONE** 8(6):e66267.

Lemos MVA, Chiaia HLJ, Berton MP, Feitosa FLB, Aboujaoude C, Camargo GMF, Pereira ASC, Albuquerque LG, Ferrinho AM, Mueller LF, Mazalli MR, Furlan JJM, Carneiro R, Gordo DGM, Tonussi R, Espigolan R, Silva RMO, Oliveira HN, Duckett S, Aguilar I, Baldi F (2016) Genome-wide association between single nucleotide polymorphisms with beef fatty acid profile in Nelore cattle using the single step procedure. **BMC Genomics** 17:213.

Lemos MVA, Peripolli E, Berton MP, Feitosa FLB, Olivieri BF, Stafuzza NB, Tonussi RL, Kluska S, Chiaia HLJ, Mueller L, Ferrinho AM, Prereira ASC, Oliveira HN, Albuquerque LG, Baldi F (2018) Association study between copy number variation and beef fatty acid profile of Nelore cattle. **Journal of Applied Genetics** 59(2):203-223.

Li W, Jiang T (2012) Transcriptome assembly and isoform expression level estimation from biased RNA-Seq reads. **Bioinformatics** 28(22):2914-2921.

Li Y, Carrillo JA, Ding Y, He Y, Zhao C, Zan L, Song J (2015) Ruminant Transcriptomic Analysis of Grass-Fed and Grain-Fed Angus Beef Cattle. **PLoS ONE** 10(7):e0134067.

Lsik CG, Tellam RL, Worley KC (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. **Science** 324:522–528.

Maltin C, Balcerzak D, Tilley R, Delday MY (2003) Determinants of meat quality: Tenderness. **The Proceedings of the Nutrition Society** 62(2):337-47.

Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, O'Connell J, Moore SS, Smith TPL, Sonstegard TS, Van Tassell CP (2009) Development and characterization of a high density SNP genotyping assay for cattle. **PLoS ONE** 4(4):e5350.

Menezes LFG, Restle J, Brondani IL, Kozloski GV, Deschamps F, Sachet RH (2009) Perfil de ácidos graxos na carne de novilhos Charolês e Nelore puros e de gerações avançadas do cruzamento rotativo, terminados em confinamento. **Ciência Rural** 39(8):2478–2484.

Merino GA, Fresno C, Netto F, Netto ED, Pratto L, Fernández EA (2016) The impact of quality control in RNA-seq experiments. **Journal of Physics: Conference Series** 705(1):012003.

Metz PAM, Menezes LFG, Santos AP, Brondani IL, Restle J, Lanna DPD (2009) Perfil de ácidos graxos na carne de novilhos de diferentes idades e grupos genéticos terminados em confinamento. **Revista Brasileira de Zootecnia** 38(3):523-531.

Morota G, Valente BD, Rosa GJM, Weigel KA, Gianola D (2012) An assessment of linkage disequilibrium in Holstein cattle using a Bayesian network. **Journal of Animal Breeding and Genetics** 129(6):474-487.

Nagl S, Williams M, Williamson J (2008) Objective Bayesian nets for systems modelling and prognosis in breast cancer. In.: Holmes DE (Ed.) **Innovations in Bayesian networks: Theory and applications**. New York: Springer, p.131.

Nogi T, Honda T, Mukai F, Okagaki T, Oyama K (2011) Heritabilities and genetic correlations of fatty acid compositions in longissimus muscle lipid with carcass traits in Japanese Black cattle. **Journal of Animal Science** 89:615–621.

Nuernberg K, Dannenberger D, Nuernberg G, Ender K, Voigt J, Scollan ND, Wood JD, Nute GR, Richardson RI (2005) Effect of a grass-based and a concentrate feeding system on meat quality characteristics and fatty acid composition of *longissimus* muscle in different cattle breeds. **Livestock Production Science** 94(1-2):137-147.

Pearson K (1901) On Lines and Planes of Closest Fit to Systems of Points in Space. **Philosophical Magazine** 2(6):559–572.

Peñagaricano F, Valente BD, Steibel JP, Bates RO, Ernst CW, Khatib H, Rosa GJM. (2015) Exploring causal networks underlying fat deposition and muscularity in pigs through the integration of phenotypic, genotypic and transcriptomic data. **BMC Systems Biology** 9:58.

Pinto LFB, Packer IU, De Melo CMR, Ledur MC, Coutinho LL (2006). Principal components analysis applied to performance and carcass traits in the chicken. **Animal Research** 55:419-425.

Pitchford WS., Deland MPB, Siebert BD, Malau-Aduli AEO, Bottema CDK (2002) Genetic variation in fatness and fatty acid composition of crossbred cattle. **Journal Animal Science** 80:2825–2832.

Ramayo-Caldas Y, Mach N, Esteve-Codina A, Corominas J, Castelló A, Ballester M, Estellé J, Ibáñez-Escriche N, Fernández AI, Pérez-Enciso M, Folch JM (2012) Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. **BMC Genomics** 13(547):1471-2164.

Realini CE, Duckett SK, Brito GW, Dalla Rizza M, De Mattos D (2004) Effect of pasture vs. concentrate feeding with or without antioxidants on carcass characteristics, fatty acid composition, and quality of Uruguayan beef. **Meat Science** 66:567-577.

Ritchie MD, Holzinger ER, Li R, Pendergrass SA, Kim D (2015) Methods of integrating data to uncover genotype-phenotype interactions. **Nature Reviews Genetics** 16(2):85-97.

Rosa GJM, Valente BD, de los Campos G, Wu XL, Gianola D, Silva MA (2011) Inferring causal phenotype networks using structural equation models. **Genetics Selection Evolution** 43:6.

Rossato LV, Bressan MC, Rodrigues EC, Alves MI, Carolino DCM, José R, Bessa B, Paula S, Alves P (2009) Composição lipídica de carne bovina de grupos genéticos taurinos e zebuínos terminados em confinamento. **Revista Brasileira de Zootecnia** 38(9):1841-1846.

Rossato LV, Bressan MC, Rodrigues EC, Gama LT, Bessa RJB, Alves SPA (2010) Parâmetros físico-químicos e perfil de ácidos graxos da carne de bovinos Angus e Nelore terminados em pastagem. **Revista Brasileira de Zootecnia** 39:1127-1134.

Saatchi M, Garrick DJ, Tait Jr RG, Mayes MS, Drewnoski M, Schoonmaker J, Diaz C, Beitz, James MR (2013) Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle. **BMC Genomics** 14:730.

Sevane N, Nute G, Sañudo C, Cortes O, Cañon J, Williams JL, Albertí P, Checa ML, Christensen M, Crisá A, Delourme D, Delourme D, Failla S, García D, Gigli S, Hocquette JF, Levéziel H, Mangin B, Marchitelli C, Miranda D, Nute G, Olleta JL, Razzaq N, Renand G, Sañudo C, Valentini A, Williams JL (2014) Muscle lipid composition in bulls from 15 European breeds. **Livestock Science** 160(1):1–11.

Sharma A, Park JE, Chai HH, Jang GW, Lee SH, Lim D (2017) Next generation sequencing in Livestock species – A Review. **Journal of Animal Breeding and Genomics** 1(1):23-30.

Siciliano C, Belsito E, De Marco R, Di Gioia ML, Leggio A, Liguori A (2013) Quantitative determination of fatty acid chain composition in pork meat products by high resolution H NMR spectroscopy. **Food Chemistry** 136(2):546-554.

Tapiero H, Nguyen-Ba G, Couvreur P, Tew KD (2002) Polyunsaturated fatty acids (PUFA) and eicosanoids in human health and pathologies. **Biomedicine & Pharmacotherapy** 56:215-222.

Tizioto PC, Coutinho LL, Decker JE, Schnabel RD, Rosa KO, Oliveira PSN, Souza MM, Mourão GB, Tullio RR, Chaves AS, Lanna DPD, Zerlotini-Neto A, Mudadu MA, Taylor JF, Regitano LCA (2015) Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC Genomics** 16:242.

USDA : United States Department of Agriculture. Livestock and Poultry: World Markets and Trade. Office of Global Analysis, 2017 (April).

Wang Q, Afshin A, Yakoob MY, Singh GM, Rehm CD, Khatibzadeh S, Micha R, Shi P, Mozaffarian D, Global Burden of Diseases Nutrition and Chronic Diseases Expert Group (2016) Impact of nonoptimal intakes of saturated, polyunsaturated, and trans fat on global burdens of coronary heart disease. **Journal of the American Heart Association** 5(1):e002891.

Warren HE, Scollan ND, Enser M, Hughes SI, Richardson RI, Wood JD (2008) Effects of breed and a concentrate or grass silage diet on beef quality in cattle of 3 ages. I: Animal performance, carcass quality and muscle fatty acid composition. **Meat Science** 78:256–69.

Wood JD, Enser M, Fisher AV, Nute GR, Richardson RI, Sheard PR (1999) Manipulating meat quality and composition, **Proceeding Nutrition Society** 58:363–370.

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

Wood J, Enser M (2017) Manipulating the fatty acid composition of meat to improve nutritional value and meat quality. In: Purslow PP (Ed.) **New aspects of meat quality. From genes to ethics**. Duxford, UK: Woodhead Publishing Limited, p. 501-535.

Yeung KY, Ruzzo WL (2001) Principal component analysis for clustering gene expression data. **Bioinformatics** 17(9):763-774.

Youngerman SM, Saxton AM, Oliver SP, Pighetti GM (2004), Association of CXCR2 polymorphism with sub clinical and clinical mastitis in dairy cattle. **Journal of Dairy Science** 87:2442-2448.

Zhu B, Niu H, Zhang W, Wang Z, Liang Y, Guan L, Guo P, Chen Y, Zhang L, Guo Y, Ni H, Gao X, Xu L, Li J (2017) Genome wide association study and genomic prediction for fatty acid composition in Chinese Simmental beef cattle using high density SNP array. **BMC Genomics** 18:464.

Zong G, Li Y, Wanders AJ, Alsema M, Zock PL, Willett WC, Hu FB, Sun Q (2016) Intake of individual saturated fatty acids and risk of coronary heart disease in US men and women: two prospective longitudinal cohort studies. **BMJ** 355:i5796.

CAPÍTULO 2 – GENES DIFERENCIALMENTE EXPRESSOS IDENTIFICADOS COM VALORES EXTREMOS DE COMPONENTES PRINCIPAIS PARA PERFIL DE ÁCIDOS GRAXOS NA CARNE BOVINA DA RAÇA NELORE

RESUMO – O objetivo do presente estudo foi identificar genes diferencialmente expressos pelo método RNA-seq e perfil de ácidos graxos no músculo *Longissimus thoracis* com uso de componentes principais (PC) em animais da raça Nelore terminados em confinamento. Foram coletadas amostras de 48 touros da raça Nelore, abatidos com média de 24 meses terminados em confinamento. Após armazenamentopor 48 horas em câmara fria (0 a 2 ° C, foram coletadas amostras com 2,5 g do músculo *Longissimus thoracis*, entre a 12^a a 13^a costela de cada meia-carcaça esquerda para as análises de perfil de ácidos graxos e extração de RNA para o sequenciamento. Os dados sequenciados gerados por meio da Plataforma HiSeq System foram convertidos ao format *FastQ*, separados por bibliotecas através do *software* Casava, e analisados utilizando o *software* R. Para consistência dos dados, foram utilizados o *FastQC* (version 0.10.1) e *Hisat2* (2.1.0). A análise de componentes principais (PCA) foi realizada com o pacote *prcomp* do *software* R, utilizando apenas informações de perfil de ácidos graxos. Para cada somatória, foram compilados para análise os AG pertencentes ao grupo somatório. Foram considerados os PC1 e PC2. Após a análise, foram selecionados dois grupos de 10 animais cada, que possuíam valores de PC1 e PC2 extremos (*High x Low*) para cada grupo de AG (AGS, AGMI, AGPI, n-3 e n-6) para análise de genes diferencialmente expressos - GDE ($p < 0,10$). A análise diferencial foi realizada com o pacote DESeq2 no *software* R. Para o grupo AGS foram encontrados 33 GDE para PC1 e nenhum GDE para PC2; para AGMI foram 56 GDE em PC1 e 10 GDE para PC2; AGPI apresentou apenas dois GDE para PC1 e nenhum para PC2; para ômega 3 nenhum GDE foi encontrado para PC1 e PC2; já para ômega 6, foram encontrados sete GDE para PC1 e três GDE para PC2. A classificação dos genes quanto à sua função biológica e busca por Vias Metabólicas relacionadas aos GDE foram realizadas utilizando o DAVID v. 6.8. O estudo realizado identificou genes envolvidos com metabolismo de lipídeos e ácidos graxos, lipólise, adipogênese, fornecimento de energia a célula e componentes de membrana celular. Esses resultados contribuem para o conhecimento dos mecanismos biológicos envolvidos no perfil de ácidos graxos na carne de animais da raça Nelore.

Palavras-chave: Bovino de corte, expressão gênica, *Longissimus thoracis*, lipoproteínas, vias metabólicas.

CHAPTER 2 - DIFFERENTIALLY EXPRESSED GENES IDENTIFIED WITH EXTREME VALUES OF PRINCIPAL COMPONENTS FOR PROFILE OF FATTY ACIDS IN NELORE CATTLE

ABSTRACT – The aim of the present study was to identify genes differentially expressed by RNA-seq method and fatty acid profile in the *Longissimus thoracis* muscle with the use of principal components (PC) in Nelore finished feedlot. Samples were collected from 48 Nelore bulls, slaughtered with an average of 24 months in feedlot. After storage for 48 hours in a cold chamber (0 to 2 ° C), 2.5 g thick samples of *L. thoracis* muscle were collected between the 12th to 13th rib of each left half carcass for analysis. Principal component analysis (PCA) was performed with the software package prcomp, using only fatty acid profile information. For each sum, the FA belonging to the summation group were compiled for analysis. PC1 and PC2 were considered. After analysis, two groups of 10 animals each were selected, which had PC1 and PC2 extreme values (High x Low) for each FA group (SFA, MUFA, PUFA, n-3 and n-6) for gene analysis differentially expressed (DEG) ($p < 0.10$). The differential analysis was performed with the DESeq2 package in the R software. For the SFA group, 33 DEG were found for PC1 and no DEG for PC2; for MUFA were 56 DGE in PC1 and 10 DEG for PC2; PUFA presented only 2 DEG for PC1 and none for PC2; for omega-3 no DEG was found for PC1 and none for PC2; already for omega 6, seven DEG for PC1 and three DEG for PC2 were found. The classification of the genes related to their biological function and search for metabolic pathways related to the DEG were performed using DAVID v. 6.8 ". The study identified genes involved with lipid and fatty acid metabolism, lipolysis, adipogenesis, energy supply to the cell and cell membrane component. These results contribute to the knowledge of the biological mechanisms involved in the fatty acid profile in the meat of Nelore.

KEYWORDS: Beef cattle, gene expression, *Longissimus thoracis*, lipoprotein, metabolic pathways.

INTRODUÇÃO

No Brasil, aproximadamente 80% dos animais bovinos (puros e cruzados) destinados a produção de carne possuem influência zebuína (*Bos taurus indicus*) em sua composição genética (Rosa; Menezes, 2016). Comparadas a raças taurinas (*Bos taurus taurus*), as zebuínas apresentam carcaça com menor teor de gordura, baixa porcentagem de gordura intramuscular (*intramuscular fat* - IMF) e menor proporção de ácido graxo saturado (AGS) (Huerta-Leidenz et al., 1993). Com isso, Rossato et al. (2010) constataram que a carne de Nelore é nutricionalmente mais saudável comparada a carne de Angus.

A gordura presente na carne influencia características organolépticas, marmorização, suculência, aroma, sabor e maciez (Wood et al., 1999), que refletem a qualidade da carne, mas pode interferir na saúde do consumidor, uma vez os consumidores têm apresentado maior preocupação com a composição da gordura da carne e conseqüentemente com o perfil de ácidos graxos (AG). O consumo de ácidos graxos saturado (AGS) em excesso estão relacionados à resistência da insulina, obesidade, aumento dos níveis séricos de colesterol e lipoproteína de baixa densidade (LDL) na corrente sanguínea, associados ao aumento da doença coronariana (Wang et al., 2016). Contudo, o consumo adequado de ácidos graxos insaturados pode melhorar o estado de saúde e prevenir doenças. O ômega-3 (ω 3) e o ácido α -linolênico (C18:3n3), possuem associação com redução de colesterol, LDL e triglicerídeos na corrente sanguínea, além de funções imunológicas e anti-inflamatórias (Ip, 1997; Tapiero et al., 2002).

A partir do sequenciamento do RNA (RNA-seq) expandiu o conhecimento de expressão de genes que tem contribuído com mais informações para compreensão de elementos funcionais do genoma. Com o RNA-seq é possível estimar o perfil de expressão gênica do animal, e assim identificar os genes expressos diferencialmente (GDE) ou sem expressão em determinada situação (Carrillo et al., 2016). Ramayo-Caldas et al. (2012) e Buchanan et al. (2013) analisaram por RNA-seq o perfil lipídico em suínos e identificaram genes com expressão diferencial que podem fornecer informações de marcadores genéticos para a seleção de perfil de AG desejáveis em futuros programas de seleção. Berton et al. (2016) por meio de RNA-seq e perfil de ácidos graxos de bovinos Nelore terminados em confinamento, identificaram genes

com diferencial de expressão relacionados com metabolismo lipídicos e de ácidos graxos, como os genes *ACSM3* e *ACSS1* que são precursores de síntese de AG, e *DGAT2* que atua na deposição de gordura saturada no tecido adiposo.

O avanço das metodologias exige que as ferramentas estatísticas acompanhem simultaneamente essa evolução, que visa gerar grandes quantidades de dados com grande exigência computacional. A análise de componentes principais (PCA) é uma análise estatística multivariada, que quantifica as variáveis e busca reduzir a dimensionalidade dos dados com perda mínima de informações e preservação máxima de variabilidade (Jolliffe, 1986). Os componentes principais (PC) que explicam a maior variedade de fenótipos originais são utilizados para análise das características. Nos programas de melhoramento genético, devido densos banco de dados dos rebanhos, os PCs podem ser considerados novas características compostas e aumentaria praticidade para o uso de índices econômicos que raramente são utilizados pela dificuldade de estimativa dos pesos para as diversas características (Boligon et al., 2016).

O objetivo do estudo foi buscar representatividade dos PC nos grupos somatórios (AGS, AGMI, AGPI e ômega-3 e ômega-6 a fim de substituí-los pelos PC. Em seguida, identificar genes diferencialmente expressos com valores extremos de PC para cada grupo somatório com uso de RNA-seq e perfil de ácidos graxos no músculo *Longissimus thoracis* em bovinos da raça Nelore terminados em confinamento.

MATERIAL E MÉTODOS

Animais e fenótipos

As amostras coletadas provêm de 48 touros jovens da raça Nelore, abatidos com idade média de 24 meses e terminados em confinamento. Os animais pertenciam a Fazenda Capivara participante do Programa de Melhoramento Genético Animal Qualitas. Neste programa, a seleção dos animais ocorreu com base em características de crescimento, terminação e precocidade sexual. O critério utilizado para o abate foi que os animais atingissem entre 500 e 550 kg. Os animais foram alimentados em condições de pastejo utilizando forragens *Brachiaria sp.* e *Panicum sp.* e sal mineral *ad libitum*. Após um ano, os animais reprodutores foram selecionados e os demais

mantidos em confinamento por um período de pelo menos 90 dias. Os animais foram abatidos em frigoríficos comerciais, de acordo com os procedimentos do Serviço de Inspeção Federal. Após abate, as carcaças foram armazenadas 48 horas em câmara fria com temperatura entre 0 a 2 ° C. Foram coletadas amostras com 2,5 g (armazenadas em tubos de 50 mL) do músculo *Longissimus thoracis*, entre a 12^a a 13^a costela de cada meia-carcaça esquerda para as análises.

Perfil de ácidos graxos

O perfil de ácidos graxos da carne (AG) foi determinado pelo método de extração de Folch et al. (1957) e os ésteres metílicos foram formados de acordo com Kramer et al. (1997). Os ácidos graxos foram quantificados utilizando cromatografia gasosa (GC-2010 Plus, auto-injetor Shimadzu AOC 20i) com coluna capilar SP-2560 (100 m × 0,25 mm de diâmetro com 0,02 mm de espessura, Supelco, Bellefonte, PA, EUA). Os ácidos graxos foram identificados por comparação do tempo de retenção dos ésteres metílicos das amostras com os padrões de ácidos graxos C4 – C24 (FAME Mix Sigma®), ácido vacênico: C18: 1 t11 (V038-1G, Sigma®), CLA: C18 : 2t10c12 (UC-61M 100 mg), CLA: C18: 2 c9t11 (UC-60M 100 mg) e ácido tricosanóico (Sigma®). Os ácidos graxos foram quantificados pela normalização da área sob a curva de ésteres metílicos usando o software GC Solution Analysis (versão 2.42.00, Shimadzu, Kyoto, Japão). O teor de ácidos graxos foi expresso como a porcentagem em peso (mg / mg) do total de ésteres metílicos de ácidos graxos quantificados. A quantificação do perfil de ácidos graxos na carne foi realizada no Laboratório de Ciência da Carne (LCC) do Departamento de Nutrição e Produção Animal da FMVZ / USP, Pirassunga, SP.

Com base na identificação, 23 AG (18 individuais e 5 grupos somatórias de AG) foram selecionados devido à sua importância em relação à saúde humana. Os AG determinados foram: láurico (C12:0), mirístico (C14:0), palmítico (C16:0), esteárico (C18:0), miristoléico (C14:1), palmitoléico (C16:1), oleico (C18:1 c9), elaídico (C18:1 n9t), vacênico (C18:1 t11), linoleico (C18:2 n6c), CLA-cis (C18:2 c9t11), CLA-trans (C18:2 t10c12), alfa-linolênico (C18:3 n3), gama-linolênico (C18:3 n6), eicosatrienóico (C20:3 n6), eicosapentanóico (C20:5 n3), araquidônico (20:4 n6) e docosahexaenóico (C22:6 n3). O somatória de AG foi calculada com as concentrações individuais de

cada AG selecionado do grupo a que pertence: saturados (C10:0 + C12:0 + C13:0 + C14:0 + C15:0 + C16:0 + C17:0 + C18:0 + C21:0 + C22:0 + C23:0 + C24:0), monoinsaturados (C14:1 + C15:1c10 + C16:1 + C17:1c10 + C18:1t11 + C18:1c9 + C18:1t9 + C18:1n7 + C20:1c11 + C22:1n9 + C24:1), poli-insaturados (C18:2n6 + C18:3n3 + C18:3n6 + C20:3n3 + C20:3n6 + C20:4n6 + C20:5n3 + C22:6n3), ômega 6 (C18:3 n6 + C20:3 n6 + C18:2 n6 + C20:4 n6) e ômega 3 (C18:3 n3 + C20:3 n3 + C22:6 n3 + C20:5 n3).

Quantificação de RNA (RNA-seq)

O RNA total foi extraído utilizando reagent TRIzol® (Life Technologies, Carlsbad, CA, USA) a partir de amostras de 100 mg congeladas em nitrogênio líquido do músculo *Longissimus thoracis*. A integridade do RNA foi verificada pelo Agilent 2100 Bioanalyzer® (Agilent, Santa Clara, CA, EUA), onde apenas amostras com RIN > 8 foram utilizadas. Um total de 2 µg de RNA de cada amostra foi utilizado para a preparação da biblioteca de acordo com o protocolo descrito no kit TruSeq RNA Preparation kit® v2 (Illumina, San Diego, CA, EUA). As bibliotecas resultantes foram quantificadas usando um kit KAPA Library Quantification® (KAPA Biosystems, Foster City, CA, EUA), de acordo com o protocolo de quantificação da biblioteca da Illumina, e então diluídas e reunidas. Três faixas de uma célula de fluxo de sequenciamento, usando o kit TruSeq PE Cluster v3-cBot-HS (Illumina, San Diego, CA, EUA), foram agrupadas e sequenciadas usando Hi Scan SQ (Illumina, San Diego, CA, EUA) com o Kit TruSeq SBS v3-HS (200 ciclos), de acordo com as instruções do fabricante. As análises de sequenciamento foram realizadas no Centro de Genoma da ESALQ, Piracicaba, São Paulo, Brasil e foram realizadas para todos os 48 animais considerados neste estudo.

Alinhamento das sequências

Os dados sequenciados gerados por meio da Plataforma HiSeq System foram convertidos ao format *FastQ*, separados por bibliotecas através do *software* Casava (https://support.illumina.com/sequencing/sequencing_software/casava.html), e analisados utilizando o *software* R. Para consistência dos dados, foram utilizados o

FastQC (version 0.10.1) e *Hisat2* (2.1.0). O controle de qualidade foi realizado por meio do pacote *Trimmomatic* (0.36). Consequentemente, foi utilizado o pacote *Hisat2* (2.1.0) para mapear e o *StringTie* para o alinhamento das *reads* com genoma bovino referência (*Bos taurus* UMD3.1.89), o qual contém 24.616 genes. Após o alinhamento, foi realizada a contagem total das *reads* utilizando *Feature Counts* e foram comparados os valores de FPKM (Fragments Per Kilobase of transcript per Million fragments mapped) dos transcritos em cada amostra. Para cada biblioteca, foi gerado um arquivo com a extensão ".bam", contendo todas as *reads* que se alinharam ao genoma de referência de todas as amostras.

Análise de Componentes Principais (PCA) e genes diferencialmente expressos (GDE)

Análise de componentes principais foi realizada com a função *prcomp* do software R, utilizando as informações de perfil de ácidos graxos. As PCA foram realizadas apenas para os grupos de AG: AGS, AGMI, AGPI, ômega 3 e ômega 6. Para cada grupo de ácidos graxos, os AG foram separados e selecionados de acordo com seu estado de insaturação ou saturação. Foram considerados os dois primeiros PC (PC1 e PC2) para cada um dos grupos somatórios de AG analisados. Após a PCA, foram selecionados dois grupos de 10 animais cada, que possuíam valores de PC1 e PC2 extremos (HIGH x LOW) para cada grupo de AG (AGS, AGMI, AGPI, n-3 e n-6). As informações transcriptômicas foram acrescentadas para identificação dos genes com expressões significativas. Para análise de GDE utilizou como genoma de referência bovino a versão *Bos taurus* UMD3.1.89, o pacote *DESeq2* (1.18.1) do software R (R CORE TEAM; R DEVELOPMENT CORE TEAM, 2013). Foram considerados como genes diferencialmente expressos, os que apresentaram $q < 0,10$.

Análise de Enriquecimento

A análise de enriquecimento dos genes foi realizada para todos os GDE por grupo somatório de AG utilizando o DAVID 6.8 (*Database for Annotation, Visualization e Integrated Discovery*) (Huang et al, 2009). A classificação dos genes quanto à sua função biológica foi de acordo com a terminologia de Ontologia do Gene (*Gene*

Ontology - GO). O mapeamento das vias metabólicas foi realizado a partir de todos os genes encontrados para o presente estudo, usando informações presentes em KEGG_Pathways.

RESULTADOS E DISCUSSÃO

O alinhamento do material resultado do sequenciamento do RNA com mapa bovino de referência apresentou acima de 92% de cobertura utilizando a versão *Bos taurus* UMD3.1.89. Após quantificar as *reads* já alinhadas, foram identificados 24,616 genes com expressão nos 48 animais. Após o controle de qualidade realizado, foram retirados os genes que não possuíam expressão em todos os animais, portanto restaram 12.214 para as análises de genes diferencialmente expressos (GDE) e 46 animais.

Os resultados encontrados para PC1 e PC2 em análise de componentes principais (PCA) estão apresentados na Tabela 1. Os valores de PC1 e PC2 foram utilizados na separação dos animais em grupos extremos, e assim realizada a análise de genes diferencialmente expressos. Os valores encontrados para o PC1 e PC2, para cada grupo, explicaram mais de 50% da variação fenotípica (PC1 + PC2). Apenas o grupo ω_3 apresentou PC1 + PC2 explicando 95% (Tabela 1).

Como apresentado nos gráficos de PCA (figuras 1 a 5 - PC1 (horizontal) x PC2 (vertical)), a coloração apresentada pelo animal está relacionada com a quantidade de lipídios intramuscular apresentada. A separação em grupos alto (*high*), médio (*medium*) e baixo (*low*) foi com base nos valores de PC apresentados. Nos cinco gráficos, os animais classificados no grupo Alta, pouco animais apresentaram coloração para verde; e no grupo Baixa aconteceu o inverso: a minoria apresentou coloração azul. Portanto, pela análise gráfica, não houve relação entre a porcentagem de IMF com os AG somados, os quais formaram as novas variáveis: os componentes principais. Os animais foram classificados em alto, baixo e médio, de acordo com os

valores dos dois PC com que explicam a maior porcentagem da variância dentro de cada grupo de ácido graxo analisado.

Tabela 1 - Valor explicado por componentes principais 1 e 2 (PC1 e PC2) e quantidade de genes diferencialmente expressos (GDE) para cada grupo de ácido graxo.

Ácido Graxo*	Componente Principal	% explicada pelo PC	GDE	q-valor (%)**
AGS	PC1	47%	33	0,10
	PC2	20%	0	0,10
AGMI	PC1	44%	56	0,10
	PC2	15%	10	0,10
AGPI	PC1	47%	2	0,10
	PC2	24%	0	0,10
ω 3	PC1	70%	0	0,10
	PC2	25%	1	0,10
ω 6	PC1	47%	7	0,10
	PC2	21%	3	0,10

*grupos somatórias - AGS: ácidos graxos saturados; AGMI: ácidos graxos monoinsaturados; AGPI: ácidos graxos poli-insaturados; ω 3: ômega 3; ω 6: ômega 6. **p-valor ajustado por Benjamini-Hochberg.

Foi realizada uma comparação de médias sobre a quantidade de lipídeo apresentada pelos animais de cada grupo classificatório, por meio de uma análise de variância (ANOVA). Com o intuito de comparar as médias entre os grupos (alto, médio e baixo) e dentro do grupo (AG), contrastamos as médias para saber se era significativo a porcentagem de lipídeos apresentada por cada animal. Com o teste t student, obtivemos o resultado que apenas dois grupos apresentaram diferenças estatisticamente significativa sobre as quantidades de lipídeos apresentadas pelos

animais: AGMI/PC2 e o ω 3/PC1. Portanto, apenas nos grupos significativos a coloração do animal (azul ou verde) ilustrada nos gráficos, tem relação com a classificação de grupo que o animal pertence (alto, médio e baixo) (Tabela 2).

Entre as PCA dos grupos de AG, a figura 4 do grupo ômega 3 (ω 3) apresenta o PC1 com maior porcentagem de variância explicada entre todos os PCA e a maior parte dos animais estão na base inferior do gráfico. A distribuição dos animais mais alinhados ao PC1, que explica 70% da variância fenotípica da característica estudada (ω 3), pode ser explicado pela alta correlação apresentada por 3 dos 4 AG individuais que compuseram o grupo 4 (C18:3n3 + C20:3n3 + C22:6n3 + C20:5n3). Após a análise de componentes com os grupos somatórias de AG, foi realizada a busca por GDE em cada grupo para os dois primeiros componentes principais (Tabela 1).

Tabela 2 – Comparação das médias dos grupos e a quantidade de lipídeos.

Ácido Graxo		Alto (\pm DP)*	Médio (\pm DP)*	Baixo (\pm DP)*	F-valor	p-valor
AGS	PC1	0,53 \pm 0,06	0,62 \pm 0,03	0,70 \pm 0,06	2,01	0,14
	PC2	0,59 \pm 0,06	0,64 \pm 0,04	0,60 \pm 0,06	0,33	0,71
AGMI	PC1	0,51 \pm 0,06	0,63 \pm 0,03	0,71 \pm 0,06	2,67	0,08
	PC2	0,52 \pm 0,06 a**	0,61 \pm 0,06 b**	0,75 \pm 0,05 c**	4,04	0,02
AGPI	PC1	0,52 \pm 0,06	0,62 \pm 0,06	0,71 \pm 0,06	2,57	0,09
	PC2	0,58 \pm 0,06	0,62 \pm 0,03	0,66 \pm 0,06	0,31	0,73
ω 3	PC1	0,49 \pm 0,05 a**	0,64 \pm 0,03 b**	0,70 \pm 0,05 c**	3,38	0,04
	PC2	0,62 \pm 0,06	0,66 \pm 0,05	0,51 \pm 0,06	2,28	0,11
ω 6	PC1	0,54 \pm 0,06	0,65 \pm 0,03	0,62 \pm 0,06	1,04	0,36
	PC2	0,70 \pm 0,06	0,58 \pm 0,06	0,58 \pm 0,03	1,23	0,30

*DP: desvio-padrão; **teste de t-student

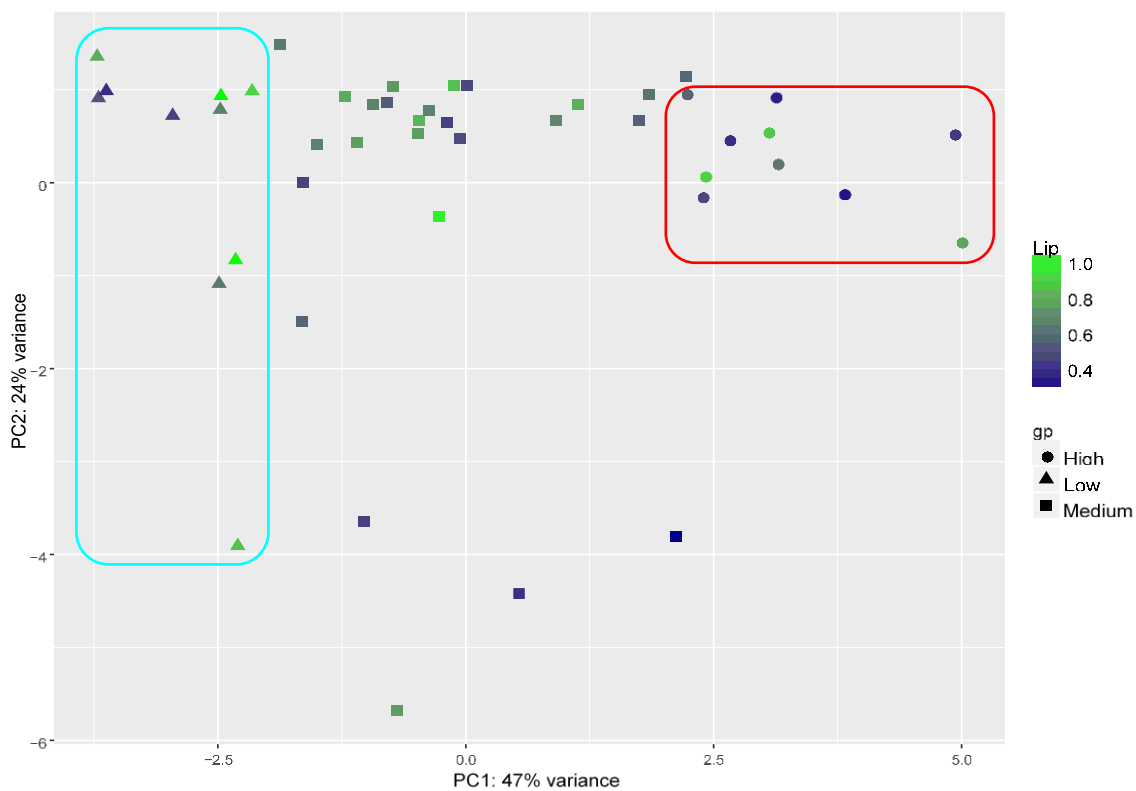


Figura 3 – Gráfico de análise de componente principal para ácidos graxos poli-insaturados para PC1 x PC2

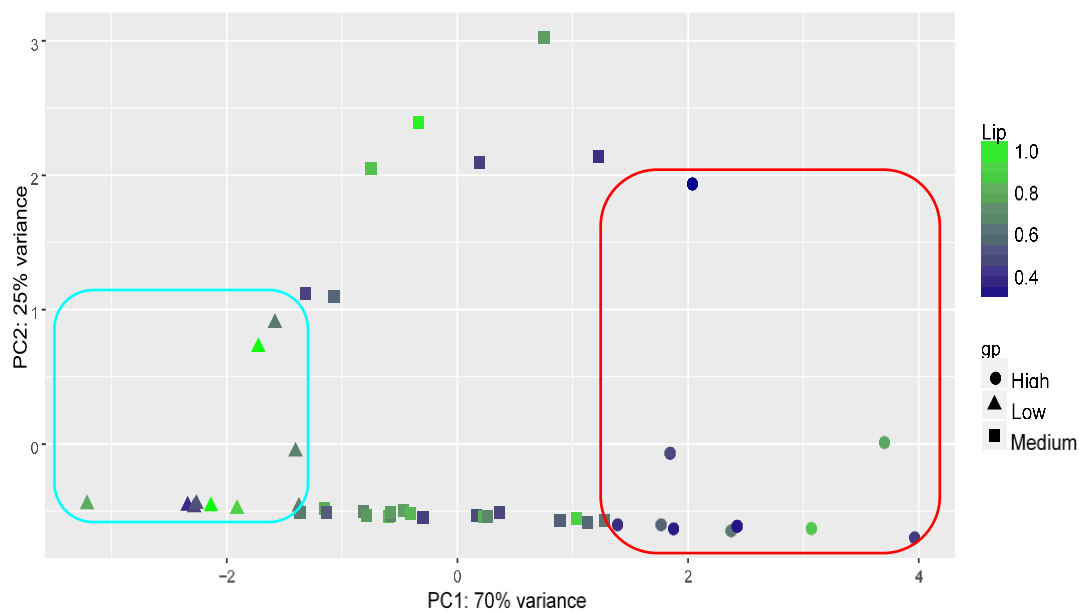


Figura 4 – Gráfico de análise de componente principal para ômega 3 (ω 3) para PC1 x PC2

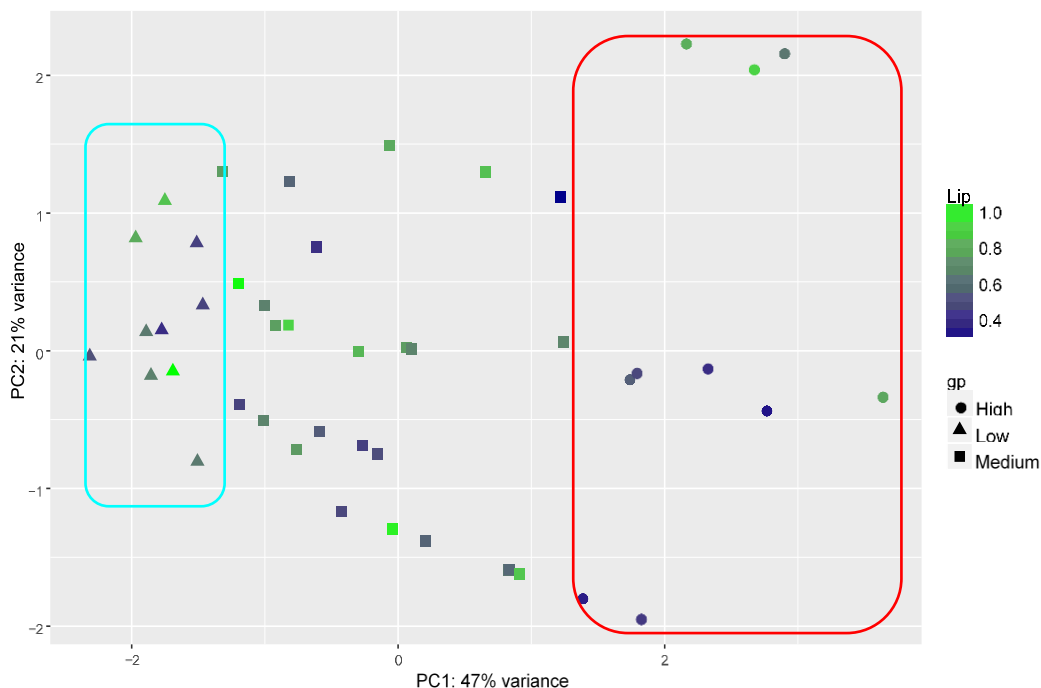


Figura 5 – Gráfico de análise de componente principal para ômega 6 (ω_6) para PC1 x PC2

A análise de GDE foi realizada individualmente para grupo somatório de AG, contrastando os valores de extremos (*HIGH* e *LOW*) que os animais apresentaram para PC nesses grupos. Para o grupo AGS foram identificados GDE apenas no PC1, total de 33 GDE (Tabela 3). O grupo AGMI apresentou 56 GDE no PC1 e 10 para PC2 (Tabelas 4 e 5, respectivamente). O grupo AGPI identificou apenas 2 GDE para o PC1 (Tabela 6). Para o grupo de ômega 3 (ω_3) apenas um GDE identificado para PC1, para ômega 6 (ω_6) 7 GDE para PC1 e 3 GDE para PC2; todos os GDE para ω_3 e ω_6 estão na tabela 5.

Tabela 3 - Genes diferencialmente expressos para componente principal 1 (PC1) da somatória de ácidos graxos saturados.

Genes	log ₂ (fold change)*	Desvio padrão	p-valor	q-valor
<i>STXBP3</i>	-0,5014	0,14	0,0004	0,0930
<i>ATP2B1</i>	-0,4454	0,13	0,0005	0,0988
<i>MPST</i>	0,4488	0,12	0,0003	0,0930
<i>TUBA8</i>	0,4194	0,12	0,0003	0,0930
<i>TMEM140</i>	0,4613	0,12	0,0001	0,0930
<i>ASB10</i>	0,4789	0,13	0,0004	0,0930

<i>EXOC1</i>	-0,3636	0,10	0,0003	0,0930
<i>GRPEL1</i>	0,2656	0,07	0,0002	0,0930
<i>GADDA45G</i>	0,5919	0,17	0,0006	0,0988
<i>TRIM54</i>	0,4735	0,13	0,0002	0,0930
<i>RPS6KA2</i>	0,4535	0,13	0,0006	0,0988
<i>SRSF5</i>	-0,5615	0,16	0,0003	0,0930
<i>CASC4</i>	-0,3993	0,11	0,0004	0,0977
<i>DOCK9</i>	-0,3495	0,10	0,0005	0,0988
<i>SIK3</i>	0,3658	0,10	0,0003	0,0930
<i>RALY</i>	0,3903	0,09	0,0001	0,0625
<i>UBC</i>	0,4579	0,11	0,0001	0,0625
<i>FAM134B</i>	0,5410	0,15	0,0004	0,0977
<i>ASB2</i>	0,4459	0,12	0,0002	0,0930
<i>MVD</i>	0,4067	0,12	0,0006	0,0988
<i>POLR2C</i>	0,2464	0,06	0,0001	0,0625
<i>NAPA</i>	0,3479	0,10	0,0006	0,0988
<i>MED25</i>	0,6188	0,16	0,0001	0,0930
<i>TUBG1</i>	0,3791	0,11	0,0004	0,0930
<i>PCYT2</i>	0,3099	0,09	0,0005	0,0988
<i>NFATC1</i>	0,4728	0,13	0,0003	0,0930
<i>TWF2</i>	0,4786	0,14	0,0006	0,0988
<i>CISH</i>	0,6089	0,17	0,0003	0,0930
<i>RHOA</i>	0,2586	0,07	0,0003	0,0930
<i>PPP2R5D</i>	0,3417	0,10	0,0005	0,0988
<i>PKNOX2</i>	0,4907	0,12	0,0001	0,0625
<i>AIP</i>	0,3909	0,10	0,0001	0,0930
<i>LDB3</i>	0,5073	0,10	0,0001	0,0067

O gene *MVD* (*Mevalonate diphosphate Decarboxylase*) localizado no cromossomo 18, na posição 13 Mb, diferencialmente expresso para o grupo AGS, é um codificador de proteína (Tabela 2). A enzima codificada a partir desse gene,

mevalonato difosfato descarboxilase (MDD), atua no início da biossíntese de colesterol. Em um estudo com *Enterococcus faecalis*, Chen et al. (2017) sugeriram que a MDD como alvo terapêutico para tratamentos de infecções bacterianas resistente a antibióticos, já que MDD é enzima final no Via do Mevalonato, via crucial para o desenvolvimento da bactéria em estudo. No mesmo cromossomo, BTA18, na posição 56 Mb, o gene *MED25* apresentou expressão diferencial diferencialmente expresso também para AGS e sua atividade está relacionada em co-ativação transcricional do complexo PPAR γ , o qual regula o metabolismo lipídico. O complexo PPAR γ é para regulação do metabolismo adiposo da glicose e dos ácidos graxos de cadeia longa (Hauner, 2002; Stienstra et al., 2007). Os PPARs pertencem à uma superfamília de receptores hormonais nucleares e possui três subtipos diferentes já identificados (PPAR γ , PPAR α e PPAR β). Além do metabolismo lipídico, o complexo PPAR regula processos biológicos como reprodução, cicatrização e até carcinogênese (Abbott, 2009).

Localizado no BTA19, região 50 Mb, o gene *PCYT2* apresentou expressão diferencial para o grupo dos AGS. Esse gene tem papel importante na biossíntese de fosfolipídios como a fosfatidiletanolamina, que são precursores em várias vias biológicas em mamíferos (Vance e Tasseva, 2013). Em humanos as fosfatidiletanolaminas são consideradas importante no coração por interromper membranas quando o fluxo sanguíneo está restrito; e ainda desempenha função secretora de lipoproteínas de baixa densidade (LDL) no fígado.

Tabela 4 - Genes diferencialmente expressos para componente principal 1 (PC1) da somatória de ácidos graxos monoinsaturados (AGMI).

Genes	log ₂ (fold change)*	Desvio padrão	p-valor	q-valor
<i>KLHL34</i>	0,5835	0.16	0.0002	0.0706
<i>CARF</i>	-0,4558	0.13	0.0004	0.0980
ENSBTAG00000037804	-0,5887	0.16	0.0003	0.0779
ENSBTAG00000048210	-0,6577	0.17	0.0001	0.0706
<i>ADGRB2</i>	0,6001	0.16	0.0002	0.0706
<i>MUL1</i>	0,4372	0.10	0.0001	0.0706
<i>ATP1A1</i>	0,6414	0.15	0.0001	0.0706

<i>STXBP3</i>	-0,4998	0.13	0.0001	0.0706
<i>LRRC40</i>	-0,4581	0.12	0.0001	0.0706
<i>MYF5</i>	-0,5736	0.16	0.0004	0.0980
<i>ATP2B1</i>	-0,4283	0.12	0.0004	0.0940
<i>NAGA</i>	0,4856	0.13	0.0002	0.0706
<i>TMEM140</i>	0,4846	0.13	0.0001	0.0706
<i>UBA6</i>	-0,4439	0.12	0.0001	0.0706
<i>NMRK2</i>	0,5451	0.15	0.0004	0.0913
<i>RAD50</i>	-0,6794	0.16	0.0001	0.0706
<i>FCHSD1</i>	0,5718	0.15	0.0002	0.0706
<i>CKMT2</i>	0,4430	0.11	0.0001	0.0706
<i>ZC3H6</i>	-0,5041	0.14	0.0002	0.0706
<i>THNSL2</i>	-0,5670	0.14	0.0001	0.0706
<i>AAK1</i>	0,5208	0.14	0.0002	0.0706
<i>TRIM54</i>	0,5435	0.13	0.0001	0.0706
<i>LYRM2</i>	-0,4443	0.10	0.0001	0.0528
<i>AKAP7</i>	-0,5606	0.15	0.0002	0.0706
<i>TMEM87A</i>	-0,3521	0.09	0.0002	0.0706
<i>CASC4</i>	-0,3898	0.10	0.0001	0.0706
<i>DOCK9</i>	-0,4162	0.10	0.0001	0.0706
<i>UCP3</i>	0,5508	0.16	0.0004	0.0940
<i>P4HA3</i>	0,6047	0.17	0.0003	0.0860
<i>WNT11</i>	0,4301	0.11	0.0002	0.0706
<i>PDE7A</i>	0,5012	0.14	0.0003	0.0779
<i>LRRCC1</i>	-0,5923	0.17	0.0004	0.0940
<i>TRAF5</i>	-0,6073	0.17	0.0003	0.0779
<i>TMEM9</i>	0,4515	0.12	0.0002	0.0706
<i>SLC25A1</i>	0,5405	0.15	0.0003	0.0896
<i>FAM134B</i>	0,5608	0.15	0.0002	0.0706
<i>AEN</i>	0,5157	0.14	0.0002	0.0706
<i>ALPK3</i>	0,5331	0.14	0.0001	0.0706
<i>POLR2C</i>	0,2551	0.07	0.0001	0.0706

<i>LIPE</i>	0,6201	0.13	0.0001	0.0116
<i>GYS1</i>	0,4523	0.12	0.0002	0.0706
<i>TNNI3</i>	0,6126	0.17	0.0003	0.0849
<i>NATD1</i>	0,3791	0.10	0.0002	0.0706
<i>CNTNAP1</i>	-0,5616	0.16	0.0003	0.0849
<i>WIPI1</i>	0,3950	0.11	0.0002	0.0706
<i>BCL2</i>	0,4400	0.12	0.0003	0.0849
<i>ITIH4</i>	-0,5678	0.15	0.0002	0.0706
<i>TWF2</i>	0,4601	0.13	0.0004	0.0896
ENSBTAG00000002960	0,5715	0.16	0.0003	0.0896
<i>AVPI1</i>	0,5872	0.15	0.0001	0.0706
ENSBTAG00000030999	0,5220	0.14	0.0001	0.0706
<i>ANK3</i>	-0,5908	0.15	0.0001	0.0706
<i>LDB3</i>	0,4178	0.12	0.0004	0.0896
<i>ADCY9</i>	0,4952	0.13	0.0001	0.0706
<i>ABHD11</i>	0,4988	0.13	0.0002	0.0706
ENSBTAG00000043557	0,5965	0.17	0.0004	0.0896

Tabela 5 - Genes diferencialmente expressos para componente principal 2 (PC) da somatória de ácidos graxos monoinsaturados (AGMI).

Genes	log2(fold change)*	Desvio padrão	p-valor	q-valor
<i>IFI6</i>	-0,7996	0,18	0,0001	0,0475
<i>RWDD3</i>	-0,6298	0,15	0,0001	0,0475
<i>PTN</i>	-0,7685	0,17	0,0001	0,0475
<i>EPHX2</i>	-0,6916	0,17	0,0001	0,0907
<i>NUSAP1</i>	-0,7660	0,18	0,0001	0,0475
<i>MRC1</i>	-0,7707	0,18	0,0001	0,0504
<i>KLHL21</i>	0,5343	0,13	0,0001	0,0554
ENSBTAG00000045494	-0,7718	0,19	0,0001	0,0554
<i>DMPK</i>	-0,6619	0,17	0,0001	0,0907
<i>EMX2</i>	-0,7378	0,18	0,0001	0,0831

Para AGMI, identificou-se o gene *UCP3* com expressão diferencial para o PC1, ele está localizado no BTA15, na posição 53 Mb. O *UCP3* (*uncoupling protein 3*) codifica proteínas desacopladoras no metabolismo oxidativo da síntese de ATP, produção de calor em vez de armazenamento de energia. Berton et al. (2016) e Cesar et al. (2016) relataram em estudos com animais Nelore, apresentaram o gene *UCP3* diferencialmente expresso para perfil de AG da gordura intramuscular da carne bovina. Van Abeelen et al. (2008) e Salopouro et al. (2009) exploraram a associação do gene *UCP3* com obesidade em humanos. Ambos encontraram associação de algumas variantes no gene com o aumento da capacidade de armazenar energia como gordura, aumento de concentração lipídica na corrente sanguínea e consequentemente o aumento do índice de massa corpórea do indivíduo, expressando assim fenótipo de obesidade.

O gene *LIPE* é codificador de proteína, atua nas vias de degradação de triacilglicerol e no complexo enzimático da AMPK. *LIPE* se encontra na posição BTA18 50 Mb. O gene *SLC25A1* está localizado no cromossomo 17 (72 Mb). Está relacionado as vias de regulação do metabolismo lipídico pelo PPAR γ e pela biossíntese de ácido acil-CoA (biossíntese de ácido graxo). O PPAR γ é um fator de transcrição que atua como um importante regulador do metabolismo lipídico e da homeostase de energia, desempenha um papel fundamental no controle das vias envolvidas na absorção de ácidos graxos, ligação de ácidos graxos, oxidação de ácidos graxos, cetogênese e síntese de carnitina (vitamina B11) (Bogna Grygiel-Górniak, 2014). Nos adipócitos o PPAR γ regula a expressão de outros genes envolvidos no metabolismo de lipídios, como o gene LPL (Tavares et al., 2007). Lemos et al (2018) em estudo de associação de CNV com perfil de ácidos graxos na carne de animais Nelore, identificou dois genes da mesma família *SLC* (*solute carrier family*), o *SLC9A1* e *SLC9C2* associados com AG com mirístico (C14:1) e AGPI, respectivamente. Os genes *LIPE* e *SLC25A1* apresentaram expressões diferenciais para mais de um grupo, o AGMI e ω -6, em ambos no PC1.

Para o PC2 do grupo AGMI, o gene *EPHX2* apresentou-se diferencialmente expresso. *EPHX2* (Epoxide hydrolase 2) que participa do processo de regulação do metabolismo de colesterol, ou seja, participa da via metabólica dos ácidos graxos. O gene *DHRS11* (*Dehydrogenase/Reductase 11*), localizado no BTA19 na posição 12

Mb apresentou expressão diferencial para o ômega 6 (ω 6), o PC1. Entre as vias metabólicas, a biossíntese de hormônio esteroides. Para o PC2 do mesmo grupo de AG (ω 6), o gene *MBNL1* (*Muscleblind like splicing regulator 1*) está presente no cromossomo 1, localizado em 115 Mb com expressão diferencial para o grupo ω 6, no PC2. O *MBNL1* está presente na via metabólica da Adipogênese, onde há processos de diferenciação celular a partir do pré-adipócito para adipócito (células de gordura, reservas de energia).

Tabela 6 - Genes diferencialmente expressos para componente principal 1 (PC1) da somatória de ácidos graxos poli-insaturados.

Genes	log2 (fold change)*	Desvio padrão	p-valor	q-valor
<i>HAUS6</i>	-0.4257	0.10	0,0001	0.0991
<i>LDB3</i>	0.4947	0.11	0,0001	0.0991

* log2(fold change): expressão relativa

Tabela 7 - Genes diferencialmente expressos para componentes principais 1 e 2 (PC1 e PC2) da somatória de ômega 3 (ω 3) e ômega 6 (ω 6).

	Genes	log2 (fold change)*	Desvio padrão	p-valor	q-valor
PC1 ω3	ENSBTAG00000014536	-0.7300	0.16	0.0000	0.0366
	<i>ASB10</i>	0.4747	0.11	0.0001	0.03029
	<i>SLC25A1</i>	0.6607	0.14	0.0001	0.01853
	<i>ALPK3</i>	0.5637	0.13	0.0001	0.03029
PC1 ω6	<i>LIPE</i>	0.5805	0.12	0.0001	0.01853
	ENSBTAG00000038755	-0.5550	0.13	0.0001	0.04478
	<i>TNNI3</i>	0.6954	0.16	0.0001	0.03029
	<i>DHRS11</i>	0.4297	0.10	0.0001	0.03029
	<i>MBNL1</i>	0.5834	0.14	0.0001	0.09317
PC2 ω6	ENSBTAG00000007816	-0.7006	0.17	0.0001	0.09317
	<i>UCK1</i>	-0.3504	0.08	0.0001	0.09317

* log2(fold change): expressão relativa

A análise de enriquecimento relacionadas com a lista total de genes com expressão diferencial do presente estudo identificou quatro vias metabólicas com valores significativos (p-valor<0,05): Via de sinalização cAMP (bta 04024), Sinalização adrenérgica em cariomiócitos (bta 04261), Via de sinalização cGMP-PKG e Secreção

pancreática (04972) (Tabela 7). A via metabólica mais significativa encontrada foi a de sinalização para cAMP, que é sintetizada a partir de ATP). A adenosina 3',5' – cíclica monofosfatase (cAMP) é um nucleotídeo que regula várias funções celulares, incluindo desenvolvimento e diferenciação celular, transcrição gênica e expressão proteica. cAMP regula processos fisiológicos, endócrinos, cardiovasculares, neurais e imunológicos (Yan et al., 2016). A cAMP é um mensageiro secundário na transdução de 66 regulação de genes, sinalização celular e biossíntese de hormônios esteroides (Gonzalez, Ceroni da Silva, 2017). Feitosa (2018) em estudo de associação genômica de haplótipos com ácidos graxos em bovinos da raça Nelore também identificou anotações sobre a cAMP.

Dentre os processos endócrinos, a via de secreção pancreática (bta 04972) uma das vias metabólicas identificadas no presente estudo, está relacionada a insulina, que é secretada pelas células β -pancreáticas em resposta à concentração elevada de glicose plasmática. Quando há uma complicação e a secreção desse líquido pancreático é modificada com vários fatores, na espécie humana torna-se um sério problema de saúde mundial: o diabetes.

Tabela 7 – Enriquecimento funcional da vias metabólicas (KEGG_Pathway) para todos os genes diferencialmente expressos (GDE) ($p < 0.05$).

GO Termo	n°	Genes	p-valor
Bta 04024: Via de sinalização cAMP	7	<i>NFATC1, ATP1A1, RHOA, TNNI3, ATP2B1, ADCY9, LIPE</i>	0.0003
Bta 04261: Sinalização adrenérgica em cariomiócitos	6	<i>TNNI3, ATP1A1, BL2, PPP2R5D, ADCY9, ATP2B1</i>	0.0004
Bta 04022: Via de sinalização cGMP-PKG	5	<i>NFATC1, ATP1A1, RHOA, ATP2B1, ADCY9</i>	0.0060
Bta 04972: Secreção pancreática	4	<i>ATP1B1, RHOA, ATP2B1, ADCY9</i>	0.0097

CONCLUSÃO

O uso de componentes principais substituindo as variáveis originais, não prejudicou a busca por GDE, dada a grande porcentagem da variação fenotípica explicada por dois primeiros componentes. O estudo realizado com dados transcriptômicos para perfil de ácidos graxos na carne, identificou genes envolvidos com metabolismo de lipídeos e ácidos graxos, lipólise, adipogênese, fornecimento de energia a célula e componente de membrana celular. Esses resultados contribuem para o conhecimento dos mecanismos biológicos envolvidos no perfil de ácidos graxos na carne. Demonstram a possibilidade de futuramente utilizar alguns desses GDE como marcadores genéticos em bovinos Nelore para selecionar a composição lipídica que possa contribuir da melhor maneira para nutrição e saúde do consumidor.

REFERÊNCIAS BIBLIOGRÁFICAS

Abbott BD (2009) Review of the expression of peroxisome proliferator-activated receptors alpha (PPAR α), beta (PPAR β), and gamma (PPAR γ) in rodent and human development. **Reproductive Toxicology** 27(3-4):246-257.

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

Benjamini Y, e Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society** 57 (1): 289–300.

Boligon AA, Vicente IS, Vaz RZ, Campos GS, Souza FRP, Carneiro R, Albuquerque LG (2016) Principal component analysis of breeding values for growth and reproductive traits and genetic association with adult size in beef cattle. **Journal of Animal Science** 94(12):5014–5022.

Buchanan JW, Garmyn AJ, Hilton GG, Vanoverbeke DL, Duan Q, Beitz DC, Mateescu RG (2013) Comparison of gene expression and fatty acid profiles in concentrate and forage finished beef. **Journal of Animal Science** 91(1):1–9.

Carrillo JA, He Y, Li Y, Liu J, Erdman RA, Sonstegard TS, Song J (2016) Integrated metabolomics and transcriptome analyses reveal finishing forage affects metabolic pathways related to beef quality and animal welfare. **Scientific Reports** 6:25948.

Cesar ASM, Regitano LCA, Poleti MD, Andrade SCS, Tizioto PC, Oliveira PSN, Felício AM, Do Nascimento ML, Chaves AS, Lanna DPD, Tullio RR, Nassu RT, Koltes JE, Fritz-Waters E, Mourão GB, Zerlotini-Neto A, Reecy JM, Coutinho LL (2016) Differences in the skeletal muscle transcriptome profile associated with extreme values of fatty acids content. **BMC Genomics** 17:961.

Chen CL, Mermoud JC, Paul LN, Steussy CN, Stauffacher CV (2017) Mevalonate 5-diphosphate mediates ATP binding to the mevalonate diphosphate decarboxylase from the bacterial pathogen *Enterococcus faecalis*. **Journal of Biological Chemistry** published online.

Folch J, Lees M, Sloane-Stanley GH (1957) A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues. **Journal of Biology and Chemistry** 226:497–509.

Gonzalez FH, Ceroni da Silva S (2006) Bioquímica hormonal. In: **Introdução à bioquímica hormonal**. 3.ed. Porto Alegre (RS): Editora da UFRGS. 7:251-312.

Grygiel-Górniak B (2014) Peroxisome proliferator-activated receptors and their ligands: nutritional and clinical implications – a review. **Nutrition Journal** 13:17.

Hauer H (2002) The mode of action of thiazolidinediones. **Diabetes Metabolism Research and Reviews** 18(2):S10-5.

Huerta-Leidenz NO, Cross HR, Savell JW, Lunt DK, Baker JF, Pelton LS, Smith B. (1993) Comparison of the fatty acid composition of subcutaneous adipose tissue from mature Brahman and Hereford cows. **Journal of Animal Science**, 71:625-630.

Ip, C. (1997) Review of the effects of trans fatty acids, oleic acid, n-3 polyunsaturated fatty acids, and conjugated linoleic acid on mammary carcinogenesis in animals. **American Journal of Clinical Nutrition** 66:1523-1529.

Jolliffe IT (1986) Principal component analysis. New York: Springer, 488p.

Kramer JKG, Fellner V, Dugan MER (1997) Evaluating acid and base catalysts in the methylation of milk and rumen fatty acids with special emphasis on conjugated dienes and total trans fatty acids. **Lipids** 32:1219–1228.

Long YC, Zierath JR (2006) AMP-activated protein kinase signaling in metabolic regulation. **The Journal of Clinical Investigation** 116(7):1776-1783.

R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Ramayo-Caldas Y, Mach N, Esteve-Codina A, Corominas J, Castelló A, Ballester M, Estellé J, Ibáñez-Escriche N, Fernández AI, Pérez-Enciso M and Folch JM. (2012) Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. **BMC Genomics** 13(547):1471-2164.

Rosa ANF, Menezes GRO (2016) Artigo: Papel do Zebu na pecuária de corte brasileira. Campo Grande: Embrapa: Gado de Corte, 1 p. (Embrapa Gado de Corte. Produção Animal).

Rossato LV, Bressan MC, Rodrigues EC, Gama LT, Bessa RJB, Alves SPA (2010) Parâmetros físico-químicos e perfil de ácidos graxos da carne de bovinos Angus e Nelore terminados em pastagem. **Revista Brasileiro Zootecnia** 39:1127-1134.

Salopuro T, Pulkkinen L, Lindström J, Kolehmainen M, Tolppanen AM, Eriksson JG, Valle TT, Aunola S, Ilanne-Parikka P, Keinänen-Kiukaanniemi S, Tuomilehto J, Laakso M, Uusitupa M (2009) Variation in the *UCP2* and *UCP3* genes associates with abdominal obesity and serum lipids: The Finnish Diabetes Prevention Study. **BMC Medical Genetics** 10:94.

Stienstra R, Duval C, Müller M, Kersten S (2007) PPARs, Obesity, and Inflammation. **PPAR Research** 2007:95974.

Tapiero H, Nguyen-Ba G, Couvreur P, Tew KD (2002) Polyunsaturated fatty acids (PUFA) and eicosanoids in human health and pathologies. **Biomedicine & Pharmacotherapy** 56:215-222.

Tavares W, Hirata MH, Hirata RDC (2007) Receptor ativado por proliferadores de peroxissoma gama ($Ppar_{\gamma}$): estudo molecular na homeostase da glicose, metabolismo de lipídeos e abordagem terapêutica. **Arquivos Brasileiros de Endocrinologia & Metabologia** 51(4).

Van Abeelen AFM, Krom M, Hendriks J, Grobbee DE, Adan RAH, van der Schouw YT (2008) Variations in the uncoupling protein-3 gene are associated with specific obesity phenotypes. **European Journal of Endocrinology** 158(5):669-676.

Vance JE, Tasseva G (2013) Formation and function of phosphatidylserine and phosphatidylethanolamine in mammalian cells. **Biochimica et Biophysica Acta (BBA) – Molecular and Cell Biology of Lipids** 1831(3):543-554.

Wang Q, Afshin A, Yakoob MY, Singh GM, Rehm CD, Khatibzadeh S, Micha R, Shi P, Mozaffarian D, Global Burden of Diseases Nutrition and Chronic Diseases Expert Group (2016) Impact of nonoptimal intakes of saturated, polyunsaturated, and trans fat on global burdens of coronary heart disease. **Journal of the American Heart Association** 5(1):e002891.

Wood JD, Enser M, Fisher AV, Nute GR, Richardson RI, Sheard PR (1999) Manipulating meat quality and composition, **Proceeding Nutrition Society** 58:363–370.

Yan K, Gao LN, Cui YL, Zhang Y, Zhou X (2016) The cyclic AMP signaling pathway: Exploring targets for successful drug discovery (Review). **Molecular Medicine Reports** 13(5):3715-3723.

ANEXO 1 – Termos de enriquecimento por *Gene Ontology* (GO) com total de genes identificados.

GO Termo	n°	p-valor	Genes
PB GO:0030182 Diferenciação neural	3	0.0279	<i>EMX2, WNT11, NAP</i>
PB GO:0048743 Regulação positiva da fibra muscular esquelética	2	0.0359	<i>BCL2, MYF5</i>
PB GO:0001952 Regulação de célula- matriz de adesão	2	0.0359	<i>BCL2, MYF5</i>

PB GO:0033235 Regulação positiva de proteína sumoilação	2	0.0409	<i>MUL1, RWDD3</i>
PB GO:0006874 Cálcio celular – homeostase iônica	3	0.0426	<i>TNNI3M, ATP2B1, DMPK</i>
CC GO:0005829 citossol	13	0.0056	<i>AKAP7, RHOA, UBA6, STXBP3, AIP, UCK1, TRAF5, MVD, DMPK, EPHX2, BCL2, LIPE, WIPI1</i>
CC GO:0005737 citoplasma	25	0.0195	<i>PKNO2, NFATC1, POLRC2, GYS1, GADD45G, LRRC1, THNSL2, ITIH4, KLHL2, NAGA, AIP, NUSAP1, WNT11, SRSF5, SIK3, ASB10, TUBA8, ANK3, HAUS6, RWDD3, RPS6KA2, CISH, UBC, LIPE,</i>
CC GO:0005813 centrossomo	6	0.0242	<i>LRRC1, MBNL1, KLHL21, TRAF5, HAUS6, TUBG1</i>
CC GO:0005827 microtubo polar	2	0.0253	<i>KLHL21, TUBG1</i>
CC GO:0031965 Membrana nuclear	4	0.0306	<i>DMPK, AEN, BCL2, RPS6KA2</i>
CC GO:0043209 mielina	4	0.0355	<i>ATP1A1, NAPA, CNTNAP1, BCL2</i>
CC GO:0033270 paranode região de áxon	2	0.0451	<i>ANK3, CNTNAP1</i>
FM GO:0004674 Proteína serina/ atividade treonina quinase	5	0.0223	<i>ALPK3, AAK1, DMPK, SIK3, RPS6KA2</i>

*PB=Processo biológico; CC = Componente celular; FM= Função molecular.

CAPÍTULO 3 – INVESTIGAÇÃO DE REDES CAUSAIS PARA PERFIL DE ÁCIDOS GRAXOS EM BOVINOS NELORE TERMINADOS EM CONFINAMENTO

RESUMO – Identificar as relações causais entre atividades gênicas e variações fenotípicas de características de interesse, permite elucidar mecanismos regulatórios de um sistema biológico. As redes bayesianas são um tipo de modelo gráfico probabilístico que calcula probabilidades através da inferência bayesiana, e assim obter a causalidade indicada por arestas em um gráfico direcionado. Assim, objetivo do presente estudo foi explorar relação causal entre genes e fenótipos por meio de redes bayesianas para perfil ácidos graxos na carne bovina de animais da raça Nelore. Foram coletadas amostras do músculo *Longissimus thoracis*, entre a 12^a a 13^a costela de cada meia-carcaça esquerda para as análises. O perfil de ácidos graxos da carne (AG) foi determinado pelo método de extração de Folch et al. (1957) e os ésteres metílicos foram formados de acordo com Kramer et al. (1997). Para a genotipagem dos 967 animais foi utilizado um painel com mais de 777.000 SNP do BovineSNP BeadChip (High-Density Bovine BeadChip). Genótipos perdidos foram imputados usando o software FImpute, v. 2.2. Os dados sequenciados gerados por meio da Plataforma HiSeq System foram convertidos ao format *FastQ*, separados por bibliotecas através do *software* Casava por, e analisados utilizando o *software* R. Para consistência dos dados, foram utilizados o *FastQC* (version 0.10.1) e *Hisat2* (2.1.0). Foram exploradas associações entre genótipos com fenotípicas (pQTL) e genótipos com dados expressão (eQTL). Após as regiões significativas co-localizadas encontradas, a rede gene-fenótipo foi construída utilizando algoritmo de aprendizagem estrutural causal - IAMB. Quatro regiões genômicas mostraram associações significativas (p-valor<0,05) com sete características, C16:0, C16:1, C18:0, C18:1c9, C18:2, CLA, C22:6, e também com a expressão de 30 genes, incluindo *SREBF1*, *MDE9*, *ACSF2*, relacionados a metabolismo e catabolismo de lipídeos; *ABCC3*, *ADCY1*, *OGDH*, metabolismo energético. O estudo esclarece a relação causal entre a expressão do perfil de ácidos graxos e os processos metabólicos que envolvem tal expressão. O presente estudo é primeiro com redes causais e AG em animais da raça Nelore. E pode ser aplicado para desvendar redes causais com fenótipos complexos subjacentes em outras espécies.

Palavras-chave: *Bos indicus*, genes, metabolismo lipídico, redes bayesianas, RNA-Seq

CHAPTER 3 - INVESTIGATION OF CAUSAL NETWORKS FOR FATTY ACID PROFILE IN NELORE CATTLE FINISHED IN FEEDLOT

ABSTRACT – Identifying the causal relationships between gene activities and phenotypic variations of traits of interest allows to elucidate the regulatory mechanisms of a biological system. Bayesian networks are a kind of probabilistic graph model that calculates probabilities through Bayesian inference, and thus obtain the indicated causality by edges in a directed graph. Thus, the objective of the present study was to explore the causal relationship between genes and phenotypes by means of Bayesian networks to profile fatty acids in bovine meat of Nelore breed. Samples of the *Longissimus thoracis* muscle were collected between the 12th to 13th ribs of each half carcass left for analysis. The fatty acid profile of the meat (FA) was determined by the extraction method of Folch et al. (1957) and the methyl esters were formed according to Kramer et al. (1997). For the genotyping of the 967 animals a panel with more than 777,000 SNPs of the BovineSNP BeadChip (High-Density Bovine BeadChip) was used. Lost genotypes were imputed using the FImpute software, v. 2.2. The sequencing data generated through the HiSeq System Platform were converted to the FastQ format, separated by libraries through the Casava software, and analyzed using the R software. For data consistency, FastQC (version 0.10.1) and Hisat2 (2.1.0). Associations between genotypes with phenotypes (pQTL) and genotypes with expression data (eQTL) were explored. After the significant co-located regions found, the gene-phenotype network was constructed using learning causal structural with algorithm - IAMB. Four genomic regions showed significant associations (p -value <0.05) with seven traits, C16: 0, C16: 1, C18: 0, C18: 1c9, C18: 2, CLA, C22: 6, and also with expression of 30 genes, including *SREBF1*, *MDE9*, *ACSF2*, related to lipid metabolism and catabolism; *ABCC3*, *ADCY1*, *OGDH*, energy metabolism. The study clarifies the causal relationship between the expression of the fatty acid profile and the metabolic processes that involve such expression. The present study is first with causal and AG networks in Nelore animals. And it can be applied to unravel causal networks with complex underlying phenotypes in other species.

Keywords: *Bos taurus indicus*, bayesian network, genes, lipid metabolism, RNA-Seq

INTRODUÇÃO

O conhecimento de efeitos causais, que são mais informativos do que as correlações, possibilita desvendar elos funcionais entre as partes múltiplas e interdependentes que controlam um sistema biológico. O entendimento causal pode ainda permitir a previsão de como as partes de um sistema podem se comportar em resposta a intervenções específicas. Para a pecuária, entender os mecanismos de causa e efeito que fundamentam as inter-relações entre fatores ambientais, práticas de manejo, fisiologia animal e resultados de desempenho é, sem dúvida, um pré-requisito para a tomada efetiva e eficiente de decisões no contexto de um sistema integrado de produção (Bello et al, 2018).

Embora estudos de associações genômicas possam ser usados para fornecer evidências sobre o modo como as características fenotípicas e a expressão se conectam, na maioria das vezes essas conexões são exploradas apenas para conectar as variáveis, mas sem direção causal (Peñagaricano et al, 2015). Em estudo de características complexas é importante descobrir não apenas as conexões, mas as direções dessas relações entre as variáveis em estudo.

Os modelos de equações estruturais (SEM: *Structural Equation Model*) é uma metodologia estatística cada vez mais utilizado em estudos científicos na última década (Rosa et al, 2011). A principal razão da propagação desse método estatístico é o fato de calcular as relações diretas ou indiretas entre variáveis causais com apenas um modelo (Civelek, 2018). Afirmações causais sobre a relação entre informações fenotípicas, genômicas e transcriptômicas são justificadas pela lei mendeliana da multiplicação aleatória de alelos, que ocorre durante a meiose e o efeito unidirecional do genótipo na expressão gênica e no fenótipo (Rosa et al, 2011; Chen, 2012).

O conhecimento biológico da atuação dos genes e funcionamento dos processos metabólicos auxiliam significativamente a interpretação dos dados de expressão gênica. Redes causais construídas a partir de relações individuais são particularmente adequadas para essa tarefa, uma vez que criam hipóteses mecanicistas que explicam as mudanças de expressão observadas em conjuntos de dados.

As redes bayesianas são um tipo de modelo gráfico probabilístico que utiliza a inferência bayesiana para cálculos de probabilidade. As redes bayesianas visam calcular a causalidade, representando a dependência condicional por arestas em um gráfico direcionado (Heckerman, 1998; Friedman et al 2000; Scutari, 2011). A identificação de genes que controlam metabolismo dos ácidos graxos da carne, auxilia no entendimento do mecanismo das redes reguladoras e os mecanismos fisiológicos por trás. O objetivo do atual estudo foi investigar por meio de redes bayesianas os efeitos causais entre perfil de ácidos graxos, genótipos e dados expressão na carne bovina de animais Nelore.

MATERIAL E MÉTODOS

Dados Fenotípicos

Animais e informações sobre manejo

Foram coletadas amostras de aproximadamente 967 para realização perfil de ácido graxo. Todos machos da raça Nelore terminados em confinamento, com idade média de 24 meses, provenientes de 8 fazendas que integram três programas de melhoramento genético (DeltaGen, CRV PAINT e Nelore Qualitas). A avaliação dos animais nestas fazendas é realizada para várias características de importância econômica como, pesos em várias idades, escores visuais para características morfológicas como a conformação da carcaça, sua musculosidade e precocidade. O banco de dados fenotípicos e de genealogia completos com as informações constantes dos arquivos zootécnicos destes programas foi disponibilizado. O critério utilizado nas fazendas para o abate dos animais foram o peso (500 a 550 kg.) e grau de acabamento.

Os períodos de nascimento dos bezerros concentraram-se de agosto a outubro e novembro a janeiro, e os mesmos foram mantidos com suas mães até os sete meses de idade. Os animais foram criados em sistema de pastagem, com lotação variando de 1,2 a 1,6 UA/ha, utilizando forrageiras do gênero *Brachiaria sp* e *Panicum sp*, com sal mineral à vontade. Após o sobrano, os animais que não foram selecionados para reprodutores, foram destinados para o abate e permaneceram no confinamento por

um período de no mínimo 90 dias. Foram utilizados volumosos de alta qualidade como silagem de planta inteira de milho ou sorgo. Como concentrado energético foi utilizado grãos de milho e/ou sorgo, e grãos de soja, farelo de soja, ou girassol como concentrado proteico. A composição da dieta foi obtida em cada confinamento em virtude das possíveis alterações que a dieta possa causar sobre a composição dos ácidos graxos. Contudo, o efeito da dieta ou manejo no confinamento foi ajustado no modelo, uma vez que este efeito foi considerado na formação dos grupos de contemporâneos.

Formação dos lotes e procedimento no abate

Para a formação dos lotes de abate foram utilizados dados dos animais em relação à fazenda de nascimento, ano e época de nascimento, e manejo alimentar desde o nascimento até o abate.

Os animais foram abatidos em função do peso e acabamento pré-estabelecido em frigoríficos sob inspeção Federal, de acordo com o procedimento padrão comercial. Durante o resfriamento das carcaças dos animais selecionados, mínimo de 24 horas pós abate, entre 0 e 2°C, foram retiradas amostras de 2,5 cm de espessura do músculo *Longissimus* na região da 12^a - 13^a costelas, das meias-carcaças esquerdas para a determinação do perfil de ácidos graxos.

Determinação do Perfil de Ácidos Graxos em carne

A quantificação do perfil de ácidos graxos da carne foi realizada no Laboratório de Ciência da Carne pertencente ao Departamento de Nutrição e Produção Animal da FMVZ/USP. O perfil de ácidos graxos da carne foi determinado pelo método de extração de Folch et al. (1957). Foi utilizada uma amostra de 2,5 g do centro do músculo *Longissimus thoracis*, acondicionada em um tubo Falcon de 50 mL. Os lipídeos foram extraídos por homogeneização da amostra com uma solução de clorofórmio e metanol 2:1. Em seguida os lipídeos foram isolados após a adição de solução de NaCl a 1,5% e centrifugação por 20 min a 2400 rpm. Os lipídeos isolados

foram transferidos para um tubo de ensaio com tampa e secos com nitrogênio (N₂) em bloco digestor a 50 °C.

A gordura separada foi metilada e os ésteres metílicos foram formados de acordo com Kramer et al. (1997). No tubo de ensaio com os lipídeos isolados e secos com N₂ (extração), foi adicionado metóxido de sódio e levado ao banho-maria por 10 minutos a 50 °C. Após o banho-maria foi resfriado em gelo por 5 minutos. Adicionou-se ácido clorídrico metanólico e levado ao banho-maria por 10 minutos a 80 °C. Foi retirado do banho-maria e resfriado em temperatura ambiente por 15 minutos. Adicionou-se hexano, carbonato de potássio e ficou em espera no escuro por 30 minutos. Retirou-se o sobrenadante e foi acondicionado em outro tubo de ensaio com sulfato de sódio. Os tubos foram resfriados em gelo por 30 minutos. Posteriormente foi retirado o sobrenadante e colocado em vial com padrão interno que foram secos com N₂ e depois diluídos em hexano para após fazer a leitura em cromatógrafo gasoso. Foi utilizado padrão interno C23:0 (tricosanóico) a fim de corrigir as perdas durante o processo de metilação dos ácidos graxos. Os ácidos graxos foram quantificados por cromatografia gasosa (GC-2010 Plus - Shimadzu, auto-injetor AOC 20i), usando coluna capilar SP-2560 (100 m × 0,25 mm de diâmetro com 0,02 mm de espessura, Supelco, Bellefonte, PA). A temperatura inicial foi de 70 °C, com aquecimento progressivo (13 °C/min) até chegar a 175 °C, mantendo por 27 minutos. Em seguida, um novo aumento de 4 °C/minuto foi iniciado até 215 °C, mantendo durante 31 minutos. Hidrogênio (H₂) foi utilizado como gás de arraste com fluxo de 40 cm³/s. Durante o processo de identificação foram utilizados os seguintes padrões: standard C4-C24 de ácidos graxos (F.A.M.E. mix Sigma®), C8-C28 (F.A.M.E. mix Sigma®), ácido vacênico C18:1 trans-11 (V038-1G, Sigma®), C18 CLA:2 trans-10, cis-12 (UC-61M 100mg), CLA e C18:2 cis-9, trans-11 (UC-60M 100mg), (NU-CHEK-PREP EUA ®) e ácido tricosanóico (Sigma®) para identificação dos ácidos graxos que foram formados durante a biohidrogenação de ácidos graxos insaturados.

No presente estudo apenas sete ácidos graxos individuais foram considerados: palmítico (C16:0), palmitoleico (C16:1), esteárico (C18:0), oleico (C18:1 cis-9), linoleico (C18:2 cis-6), ácido linoleico conjugado (CLA) e docosahexaenóico (C22:6). Estes AGs foram escolhidos devido a sua maior importância e impacto em relação à saúde humana.

Dados Genômicos

Genotipagem dos animais

Para a genotipagem dos 967 animais foi utilizado um painel com mais de 777.000 SNP do BovineSNP BeadChip (High-Density Bovine BeadChip). Genótipos perdidos foram imputados usando o software FImpute, v. 2.2 (Sargolzaei et al., 2014). Foi realizada a limpeza e o controle de qualidade dos marcadores SNPs, excluindo-se aqueles com posição genômica desconhecida, marcadores monomórficos além dos que apresentaram MAF (frequência de alelos menores) menor que 0,05. Também foram excluídos das análises marcadores SNPs que apresentaram taxa de atribuição dos genótipos (Call Freq) menor a 93%. Ainda serão retirados marcadores que apresentarem intensidade média do cluster baixa.

Dados Transcriptômicos

Extração de RNA total

A extração do RNA total foi realizada em 78 amostras, que foram processadas no Laboratório de Bioquímica e Biologia Molecular (LBM) do Departamento de Tecnologia da Faculdade de Ciências Agrárias e Veterinárias da UNESP de Jaboticabal, SP. Para isso, foram utilizadas, em média, 50 mg de tecido muscular previamente estocado em RNA holder (BioAgency, São Paulo, SP, Brasil) e o kit para extração de RNA RNeasy Tissue Mini Kit (Qiagen, Valencia, CA, EUA), como descrito abaixo:

As amostras foram retiradas do freezer a -80 °C, acondicionadas em gelo, pesadas, e então foi realizada a ruptura e homogeneização do tecido com 1 ml da solução Qiazol Lysis Reagent no homogenizador de tecidos ultraturrax T50 (IKA). Cada amostra foi homogeneizada três vezes durante 30 segundos e nos intervalos, também de 30 segundos, foram mantidas no gelo. Após esta etapa foi adicionado 200µl de clorofórmio e as amostras permaneceram durante 15 minutos à temperatura

ambiente. Logo após, as amostras foram centrifugadas a 12000 x g por 15 minutos a 4°C.

Formaram-se então três fases, sendo a superior, incolor e aquosa, na qual estava contido o RNA, que foi transferida para um novo tubo. Foram adicionados 600µl de etanol 70% e a amostra foi homogeneizada em vortex. Em seguida, esta solução foi transferida a uma coluna e centrifugada a 8000 x g por 15 segundos para a eliminação dos resíduos. Foram adicionados à membrana, o mix contendo 10 µl de DNase I (1500 kunitz units) e 70 µl de tampão RDD, fornecidos pelo Kit RNase-Free DNase set (Qiagen, Valencia, CA, EUA) e as amostras foram incubadas à temperatura ambiente por 15 minutos. Posteriormente, foram adicionados 350 µl de tampão RW1 e as amostras foram centrifugadas novamente a 8000 x g por 15 segundos para eliminação de resíduos. A coluna foi lavada com 500 µl do tampão RPE por duas vezes e o RNA ligado à coluna foi eluído em 30 µl de água livre de RNase por centrifugação a 8000 x g, durante um minuto. As amostras foram novamente armazenadas em freezer a -80°C.

Verificação da concentração e integridade do RNA extraído

A determinação da concentração e pureza do RNA foi realizada com leitura da absorbância no espectrofotômetro NanoDrop 1000 Spectrophotometer® (Thermo Fisher Scientific, Santa Clara, CA, EUA, 2007) nos comprimentos de onda de 260nm, 280nm e 230nm. Foi utilizado o valor de absorbância a 260 nm para a determinação da concentração do RNA (ng/µl), as razões A260nm/A280nm para avaliar a contaminação por proteínas e A260nm/230nm para avaliar as contaminações por outros compostos, como sais, polissacarídeos e compostos orgânicos como fenol. Os valores de referência para estas razões são de 1,8 a 2 e maior ou igual a 2, respectivamente.

A qualidade do RNA total foi verificada através do aparelho Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, EUA, 2009), e do kit Agilent RNA 6000 Nano Chips (Agilent, Santa Clara, CA, EUA). Além disso, a concentração do RNA extraído foi medida por meio de Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA, EUA, 2010). Resultados iguais ou superiores a 7 foram considerados satisfatórios. Foi confirmada

a ausência de contaminação por DNA genômico nas amostras através do aparelho Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA, EUA, 2010), e o kit de quantificação de DNA, Qubit dsDNA Assay Kit (Invitrogen, Carlsbad, CA, EUA). Além disso, utilizando o mesmo aparelho e o kit Qubit RNA Assay, o RNA total foi quantificado e então, uma fração deste a uma concentração de 2µg foi retirada para que pudesse ser enviado para o sequenciamento.

Alinhamento das sequências

Os dados sequenciados gerados por meio da Plataforma HiSeq System foram convertidos ao format *FastQ*, separados por bibliotecas (dados multiplexos) através do software Casava por https://support.illumina.com/sequencing/sequencing_software/casava.html, e analisados utilizando o software R. Para consistência dos dados foram utilizados o *FastQC* (version 0.10.1) e *HiSAT2* (2.1.0). O controle de qualidade foi realizado através do pacote *Trimmomatic* (0.36). Consequentemente, foi utilizado o pacote *Hisat2* (2.1.0) (Pertea et al., 2016) para mapear e realizar o alinhamento das *reads* com genoma bovino de referência (*Bos taurus* UMD3.1.89), qual continha 24.616 genes avaliados pelo banco de dados do NCBI. Após obter um mapeamento gerado pelo *HiSAT2*, as amostras foram submetidas ao pacote *StringTie*, para quantificação dos transcritos (Pertea et al., 2016). Após a quantificação, foi realizada a contagem total das *reads* utilizando *Feature Counts*. Para cada biblioteca, foi gerado um arquivo com a extensão ".bam", contendo todas as *reads* que se alinharam ao genoma de referência de todas as amostras.

Para a identificação de genes diferencialmente expressos foi utilizado o pacote DESeq2 (versão 1.22.1) no Software R para identificar os genes ou regiões do genoma diferencialmente expressos (Love et al., 2014). Para a identificação de regiões diferencialmente expressas foi considerado um valor de q (*p-value adjust*) menor que 0,10.

Treinamento de Redes Bayesianas

A construção da rede bayesiana iniciou-se com os mapeamentos utilizando tais informações: fenotípicas (967 animais), genotípica (967 animais) e transcriptômicos (78 animais). Foram realizadas duas análises genômicas completa. A primeira foi um mapeamento de QTL (*Quantitative trait loci*) clássico, integrando dados fenotípicos e genotípicos (*pQTL – phenotype Quantitative trait loci*); e a segunda, realizamos um mapeamento de QTL de expressão (*eQTL – expression Quantitative trait loci*) integrando informações transcriptômicas com genotípicas.

Para o mapeamento pQTL, foi realizada a partir de um modelo linear com teste de Wald (Wald, 19434) utilizando a matriz G, para cada ácido graxo individualmente. Foram selecionados os genes que estavam localizados em uma janela de até 1 Mb dos pQTLs mais significativos (menor p-valor) para cada ácidos graxos em estudo.

O mapeamento de eQTL foi realizado utilizando modelo linear misto utilizando GCTA-loci com matriz de G foram usados para explicar a população de estrutura incluindo correção de múltiplos testes com taxa de descoberta falsa (FDR) de 5% (Benjamini & Hochberg, 1995). Incluímos a idade ao abate, como covariáveis valores de PCA para considerar os fatores de variação nos dados de expressão gênica e a data de seqüenciamento como efeito fixo. Os eQTL locais (*cis-eQTL putativo*) foram identificados quando os SNPs significativos estavam localizados até de 1 Mb do gene associado. Após a localizar as regiões compartilhadas pelo mapeamento, iniciou-se a construção da rede causal.

A rede Bayesiana (BN) é um modelo probabilístico em que um gráfico G acíclico dirigido é usado para definir as dependências estocásticas quantificados por uma distribuição de probabilidade (Pearl 1988; Koller and Friedman 2009). As variáveis $X = \{X_i\}$ sob investigação no contexto incluem T características (X_{t_1}, \dots, X_{t_T}) e S SNPs (X_{s_1}, \dots, X_{s_S}), cada um dos quais estão associados a um nó no gráfico G. Os arcos entre os nós representam dependências estocásticas diretas e determinam como a *distribuição global* de X se decompõe em um Conjunto de Distribuições Locais,

$$P(\mathbf{X}) = \prod P(X_i | \Pi_{X_i}),$$

(1) Para cada variável X_i , dependendo somente de progenitores Π_{X_i} . Este modelo de representação pode capturar associações diretas e indiretas entre SNPs e fenótipos e associações entre SNPs devido à ligação e estrutura da população.

Uma aplicação da rede bayesiana, é na formação de estruturais causais. As estruturas causais são representadas usando modelos gráficos. Os modelos gráficos são compostos de duas partes: um conjunto V de variáveis aleatórias que descrevem as quantidades de interesse e um gráfico $G = (V, E)$ em que cada vértice $v \in V$ é chamado nó, e cada aresta $e \in E$, também chamada arco ou *link*, é usado para expressar a estrutura de dependência dos dados, ie, o conjunto de relações de dependência entre as variáveis em V . O algoritmo IAMB (*Incremental association Markov Blanket*/baseada no algoritmo de detecção de coberturas de Markov – MB: Markov Blanket), que é baseado em um esquema de seleção de duas fases (Tsamardinos et al, 2003). A cobertura de Markov de um atributo T , denotado por $MB(T)$, é um conjunto mínimo de atributos (ou características, variáveis, os usamos alternadamente) condicionados, nos quais todos os outros atributos são probabilisticamente independentes do alvo T (Definição 1), assim, conhecer os valores do $MB(T)$ é suficiente para determinar a distribuição de probabilidade de T e os valores de todos os demais atributos tornam-se supérfluos. Obviamente, só podemos usar atributos no $MB(T)$ em vez de todos os atributos para classificação ótima. Além disso, sob certas condições (fidelidade a uma Rede Bayesiana) o $MB(T)$ é o subconjunto onde estão os nós do alvo T na rede Bayesiana.

Foi avaliada a estabilidade da estabilidade da rede, deixando de fora uma observação por vez do conjunto de dados, avaliamos cada aresta da rede original em termos de presença (variável binária; presença ou ausência na rede reamostrada) e a direção de aresta (três resultados possíveis; mesma direção que a seta original, direção oposta ou arco não direcionado). Todos estes as análises foram realizadas usando o pacote *bnlearn* (Scutari, 2010) implementado na linguagem do software R.

RESULTADOS E DISCUSSÃO

A primeira fase deste estudo foi realizar uma associação genômica, utilizando informações fenotípicas com genotípicas (mapeamento pQTL). Os fenótipos foram sete ácidos graxos (AG) individuais: palmítico (C16:0), palmitoléico (C16:1), esteárico (C18:0), oleico (C18:1c9), linoleico (C18:2c6), CLA e e docosaheptaenóico (C22:6). Foram identificados, em média, para as sete características 403.000 pQTLs. A segunda

fase do estudo foi realizar o mapeamento a partir de dados de expressão e fenótipos. Nessa fase foram identificados 10026 eQTLs. A fase seguinte, o objetivo era identificar as regiões no genoma que os dois mapeamentos compartilhavam. Dos 7 AG apenas quatro apresentaram regiões co-localizadas ($FDR < 0,05$), que foram: C18:0, C18:2c6, CLA e C22:6. Dentre as quatro regiões compartilhadas, genes que estivessem próximos até 1 Mb seriam identificados. No total, 30 genes foram identificados em associação com as 7 características e os 4 eQTLs.

Com os resultados obtidos nas etapas anteriores, construímos a rede causal (Figura 1). A avaliação da rede a fim de decifrar ligações causais potenciais foi utilizado o algoritmo IAMB (algoritmo baseado em restrição eficiente baseado no algoritmo de indução causal) em conjunto com o teste Z de Fisher para avaliar independência ($\alpha = 0,05$) foram utilizados para inferir as relações envolvendo estas 41 variáveis (fenotípicas, genotípica e de expressão) ajustadas. De informações prévias, os eQTLs receberam a ligação para seu gene em associação direta (ex: eQTL7 \rightarrow *KLHL25*).

A estabilidade da rede foi avaliada usando reamostragem Jackknife. Em cada iteração, a rede foi inferida a partir de um novo conjunto de dados que foi criado pela remoção de um animal de cada vez do conjunto de dados original. A estrutura dessa nova rede foi então comparada com a estrutura original. Em particular, avaliamos a estabilidade de cada elo (presença ou ausência) e também a estabilidade da direção do elo (Figura 2)

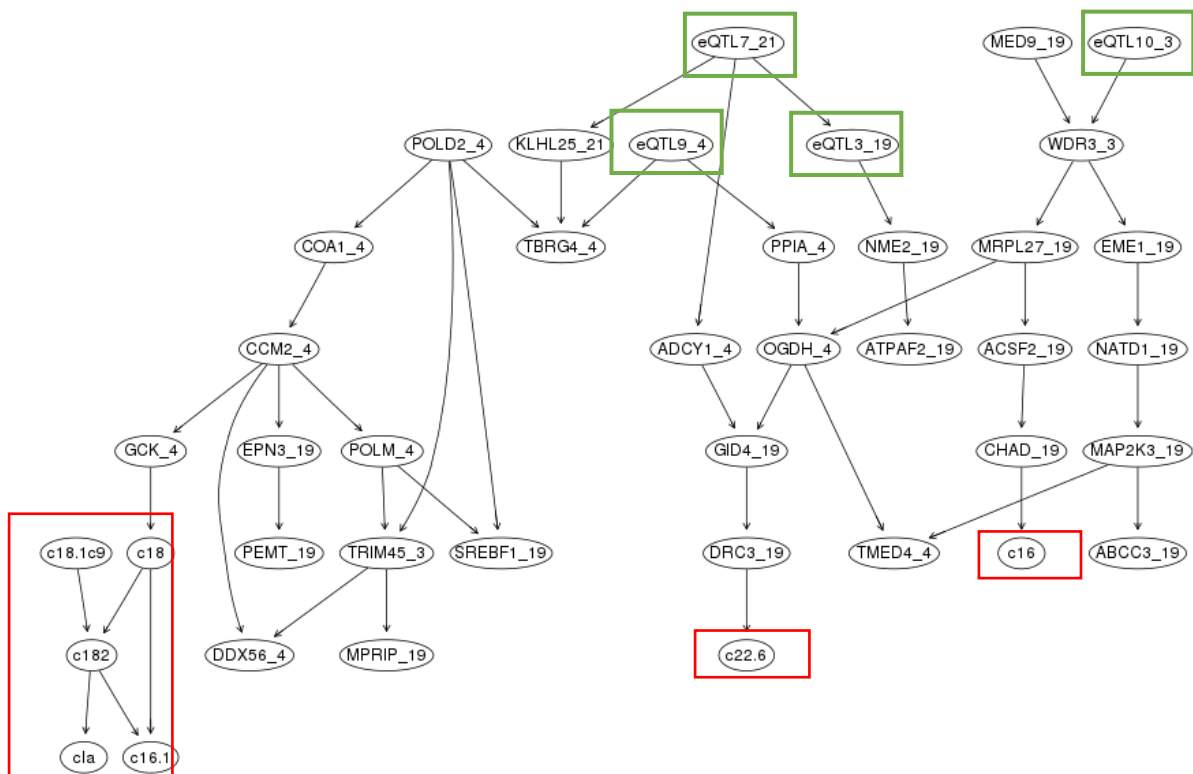


Figura 1 - Rede causal integrando informações genéticas, fenotípicas (vermelho) e transcriptômico (verde).

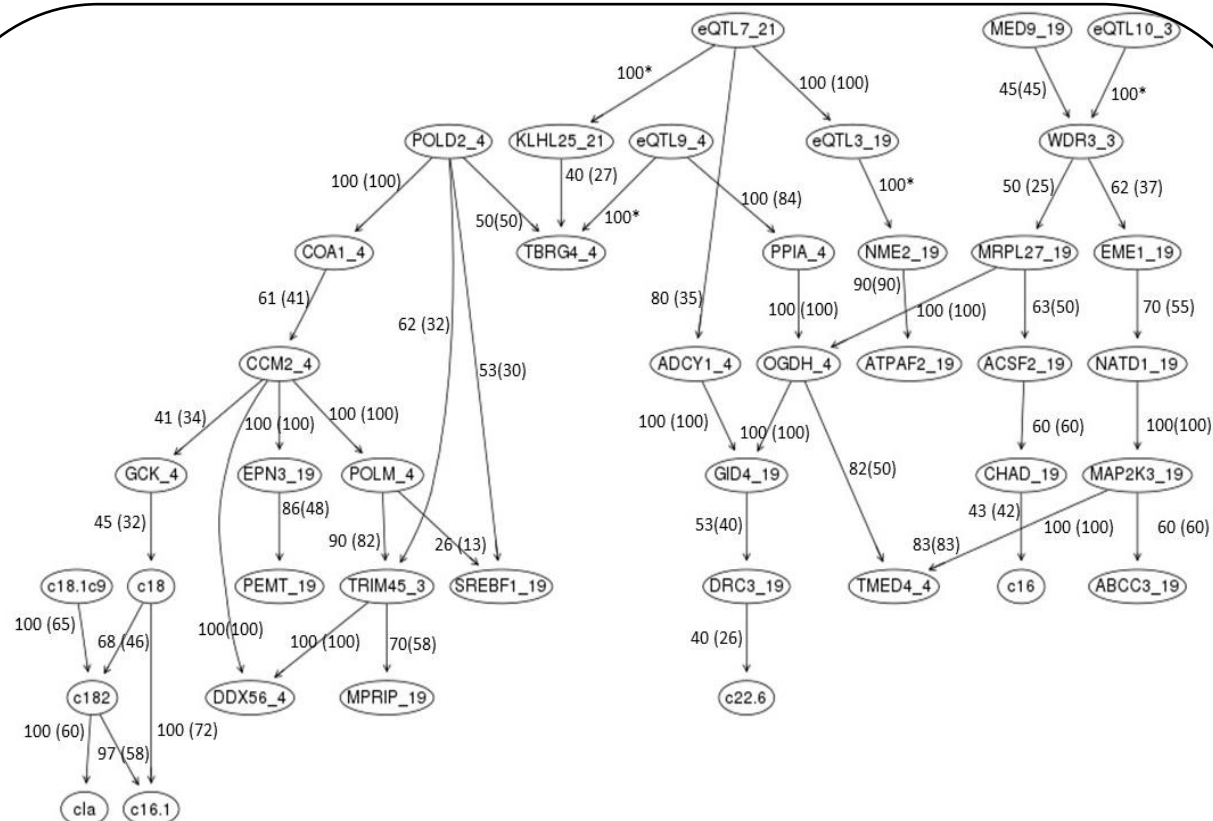


Figura 2 - Avaliação da estabilidade da rede usando o reamostragem Jackknife. Os resultados são expressos como frequência (porcentagem) que um dado arco foi apresentado (com a mesma direção) nas redes reamostradas.

No total, foram identificados 30 genes nas regiões gênicas compartilhadas. Os quais apresentaram funções relacionadas a metabolismo energético (*OGDH*), sistema imunológico (*ABCC3*, *MAPK2K3*), funções estruturais (*MYO1G*) e com metabolismo ou catabolismo de gordura, o gene *SREBF1*.

O gene *SREBF1* (Sterol Regulatory Element Binding Transcription Factor 1) é um dos principais fatores de transcrição envolvidos no metabolismo lipídico e reguladores esterol ligados às proteínas (SREBP -1c). *SREBF1* está relacionado a fator de transcrição, regulação de biossíntese de colesterol e relacionado ao complexo PPAR α (*Regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha*). O PPAR α é um fator de transcrição que atua como um importante regulador do metabolismo lipídico e da homeostase de energia, desempenha um papel

fundamental no controle das vias envolvidas na absorção de ácidos graxos, ligação de ácidos graxos, oxidação de ácidos graxos, cetogênese e síntese de carnitina (vitamina B11) (Bogna Grygiel-Górniak, 2014).

CONCLUSÃO

O estudo realizado é um dos pioneiros em redes bayesianas com AG em animais da raça Nelore em busca de respostas para relações causais. Pode ser aplicado para desvendar redes causais com fenótipos complexos subjacentes em outras espécies. O procedimento que foi realizado em várias etapas para inferir a rede causal integrando dados fenotípicos, genotípicos e transcriptômicos demonstrou que é possível decifrar relações entre variáveis de características complexas e esclarecer relações causais entre variáveis que influenciam as características complexas

REFERÊNCIAS BIBLIOGRÁFICAS

Bello NM, Ferreira VC, Gianola D, & Rosa, GJM. (2018). Conceptual framework for investigating causal effects from observational data in livestock. **Journal of Animal Science**.

Benjamini Y, e Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society** 57 (1): 289–300.

Civelek, ME (2018) "Essentials of Structural Equation Modeling". **Zea E-Books**. 64.

Chen L. (2012) Using eQTLs to reconstruct gene regulatory networks. In: Quantitative Trait Loci (QTL): New York, NY: **Humana Press** 871:175–189.

Friedman N, Linial M, Nachman I, & Peter D. (2000). Using Bayesian networks to analyze expression data. **Proceedings of the Fourth Annual International Conference on Computational Molecular Biology**

Folch J, Lees M, Sloane-Stanley GH (1957) A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues. **Journal of Biology and Chemistry** 226:497–509.

Heckerman, D. (1998). A Tutorial on Learning with Bayesian Networks. **Learning in Graphical Models**, 301–354.

Kramer JKG, Fellner V, Dugan MER (1997) Evaluating acid and base catalysts in the methylation of milk and rumen fatty acids with special emphasis on conjugated dienes and total trans fatty acids. **Lipids** 32:1219–1228.

Love MI, Wolfgang H e Simon A. (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. **Genome Biology** 15:550.

Peñagaricano F, Valente BD, Steibel JP, Bates RO, Ernst CW, Khatib H, Rosa GJM. (2015) Exploring causal networks underlying fat deposition and muscularity in pigs through the integration of phenotypic, genotypic and transcriptomic data. **BMC Systems Biology** 9:58.

R Development Core Team: R: A language and environment for statistical computing. (2011) In. Vienna, Austria: R Foundation for Statistical Computing.

Rosa GJM, Valente BD, De Los Campos G, Wu XL, Gianola D, Silva MA. (2011) Inferring causal phenotype networks using structural equation models. **Genetics Selection Evolution** 43:6.

Sargolzaei M, Chesnaus JP, Schenkel F. (2014) A new approach for efficient genotype imputation using information from relatives. **Methodology** 15:478.

Scutari M. (2010) Learning bayesian networks with the bnlearn R package. **Journal of Statistical Software**. 35(3):1–22

Tsamardinos I, Aliferis CF, Statnikov A: Algorithms for Large Scale Markov Blanket Discovery. (2003) In: **Proceedings of the 16th International Florida Artificial Intelligence Research Society Conference**. vol. 2003: Menlo Park, California: AAAI Press; 2003:376–381. 26.

Wald A. (1943) Tests of statistical hypothesis concerning several parameters when the number of observations is large. **Transactions of the American Mathematical Society**, 54, 426–482.

Sargolzaei M, Chesnaus JP, Schenkel F. (2014). A new approach for efficient genotype imputation using information from relatives. **Methodology**, 15:478.