

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

Jennifer Jessica Bruscadin

“Variantes reguladoras associadas à expressão alelo-específica (aseQTL) em músculo de Nelore (*Bos indicus*)”

São Carlos

2020

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

Jennifer Jessica Bruscadin

“Variantes reguladoras associadas à expressão alelo-específica (aseQTL) em músculo de Nelore (*Bos indicus*)”

Dissertação apresentada ao Programa de Pós-Graduação em Genética e Evolução do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos para obtenção do título de mestre em Genética e Evolução.

Orientação: Dra. Luciana Correia de Almeida Regitano

Co-orientação: Dra. Marcela Maria de Souza

São Carlos

2020



UNIVERSIDADE FEDERAL DE SÃO CARLOS

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

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado da candidata Jennifer Jessica Bruscadin, realizada em 27/02/2020:

Profa. Dra. Luciana Correia de Almeida Regitano
EMBRAPA

Prof. Dr. Marcos Roberto Chiaratti
UFSCar

Prof. Dr. Juliano Coelho da Silveira
USP

Dedico este trabalho aos meus pais, Luiz e Luzia, por todo o incentivo e apoio em todos os momentos.

AGRADECIMENTOS

Agradeço à Dra. Luciana Correia de Almeida Regitano pela oportunidade de fazer parte do grupo de pesquisa, pela orientação e aconselhamento sobre o projeto e sobre a vida acadêmica.

Agradeço à Dra. Marcela Maria de Souza, pela orientação e colaboração durante todo o desenvolvimento do trabalho.

Agradeço à Embrapa Pecuária Sudeste, por oferecer a infraestrutura necessária para a execução do trabalho, bem como aos funcionários pelo tratamento sempre cordial e colaborativo.

Agradeço ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular e à Universidade Federal de São Carlos pelo auxílio financeiro a atividades, estímulo à divulgação científica e oferecimento de disciplinas que contribuíram com os conhecimentos necessários para a execução desse projeto.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de mestrado.

Agradeço à FAPESP e ao CNPq pelo auxílio financeiro necessário para a obtenção de todos os dados utilizados neste trabalho.

Agradeço a todos os colegas do grupo de pesquisa pela troca de idéias, colaborações e companheirismo.

Agradeço aos meus pais, Luiz Fernando Bruscadin e Luzia de Fátima Jacometo Bruscadin, minha avó Therezinha Cipolini Jacometo e ao meu namorado William Uliana Albarice pelo suporte emocional, apoio e paciência nos momentos mais difíceis, nos quais muitas vezes estive distante.

Agradeço às minhas amigas Letícia Cristina Xavier e Leticia Carvalho de Alcântara, bem como a todos os meus verdadeiros amigos, por terem acompanhado minha trajetória e me estimulado a não desistir de ser pesquisadora apesar das dificuldades.

Agradeço ao meu antigo grupo de pesquisa, “Biologia da Socialidade”, em especial ao Dr. Angel Roberto Barchuk e ao Dr. Delcio Eustaquio de Paula Junior, por contribuírem com o meu conhecimento e estimularem o meu interesse nas áreas de genética e biologia molecular.

RESUMO

O Brasil é um grande exportador de carne, o que faz o melhoramento da raça Nelore, majoritariamente presente no rebanho brasileiro, impactar diretamente nesse setor. Ferramentas genômicas podem identificar genes e SNPs associados a fenótipos de interesse agropecuário, mas a eficiência da seleção depende da expressão dos alelos selecionados nos animais. Estudos prévios identificaram 820 SNPs apresentando expressão alelo-específica (ASE, do inglês *allele-specific expression*) no músculo *Longissimus thoracis* de Nelore, para os quais o presente estudo realizou a busca de SNPs reguladores associados ao padrão de expressão alélica, ou aseQTLs (do inglês *allele-specific expression quantitative trait loci*). Para isso, foram realizadas varreduras dentro de janelas de 1Mb em relação ao SNP com ASE (SNP ASE) e, posteriormente, os genótipos candidatos dessa janela foram relacionados ao desequilíbrio alélico no SNP ASE. Para obter o perfil de metilação dos aseQTLs, as posições candidatas foram sobrepostas a dados de metilação obtidos por RRBS (Do inglês *Reduced Representation Bisulfite Sequencing*) de 12 animais da mesma população. As sequências flanqueadoras dos aseQTLs foram utilizadas na predição da presença alelo-específica de sítios de ligação de fatores de transcrição (SLFTs) e de miRNAs. Foram identificados 1.134 aseQTLs, associados a 126 SNPs ASE. Dezesesseis aseQTLs correspondiam a regiões sujeitas à metilação na nossa população, 215 potencialmente afetavam os sítios de ligação de fatores de transcrição expressos em músculo bovino e 162 afetavam sítios de ligação de miRNA expressos em músculo da nossa população. A integração dos aseQTLs a estudos de associação obtidos previamente na mesma população mostrou um enriquecimento para características de qualidade de carne, tais como maciez, área de olho de lombo e gordura intramuscular. AseQTLs associados ao gene *CMYA5* apresentaram intenso desequilíbrio de ligação (DL), múltiplas sobreposições a fenótipos de qualidade de carne e predição de sítios de ligação para miRNAs, indicando potenciais sítios reguladores da expressão desse gene. No geral, os resultados indicam que a ASE pode ser resultante da regulação exercida por múltiplos aseQTLs e sofrer modulação por mecanismos epigenéticos e *cis*-reguladores. Os aseQTLs descritos apresentaram grande influência do DL. Essas variantes podem interferir na expressão de genes importantes para fenótipos que afetam, entre outros, características sensoriais da carne, fatores imprescindíveis para a aceitação do consumidor.

Palavras-chave: Qualidade da carne, genômica, epigenética, mecanismos de regulação

ABSTRACT

Brazil is a significant exporter of beef, so the improvement of Nelore, a breed mostly present in the Brazilian herd, directly impact this sector. Genomic tools can identify SNPs associated with interesting livestock phenotypes, but the selection efficiency depends on the expression of the selected alleles in the animals. Previous studies have identified 820 SNPs presenting allele-specific expression (ASE) in Nelore's *Longissimus thoracis* muscle, for which we searched for regulatory SNPs associated with this allelic expression pattern, or aseQTLs (allele-specific expression quantitative trait loci). To this end, scans were performed within 1Mb windows from the SNP with ASE (ASE SNP), and subsequently, the candidate genotypes of that window were related to the allelic imbalance in the ASE SNP. For obtaining the methylation profile of aseQTLs, the candidate positions were overlapped with RRBS (Reduced Representation Bisulfite Sequencing) methylation data from 12 animals of the same population. The flanking sequences of aseQTLs were used to predict the allele-specific presence of transcription factor binding sites (TFBSs) and miRNAs. We identified 1,134 aseQTLs associated with 126 ASE SNPs. Sixteen aseQTLs were methylated in our population, 215 affected TFBSs expressed in bovine muscle, and 162 aseQTLs affected miRNA binding sites. The integration of aseQTLs with previous association studies performed in the same population showed enrichment for meat quality traits, such as tenderness, rib eye area, and intramuscular fat. AseQTLs associated with the *CMYA5* gene showed high linkage disequilibrium (LD), multiple overlaps with meat quality phenotypes and prediction of binding sites for miRNAs, thus indicating potential regulatory sites for this gene. In general, the results indicate that ASE may be the result of regulation exercised by multiple aseQTLs and undergo modulation by epigenetic and *cis*-regulatory mechanisms. The aseQTLs described had a significant influence of LD. These variants may impact the expression of genes functionally important for phenotypes that affect, among others, meat sensory characteristics, factors that are essential for consumer acceptance.

Keywords: Meat quality, genomics, epigenetics, regulatory mechanisms

LISTA DE ILUSTRAÇÕES

Figura 1. Subtipos de expressão alelo-específica (ASE).....	17
Figura 2. Esquema de um <i>locus</i> quantitativo associado à expressão alelo-específica (aseQTL).....	32
Figura 3. Distribuição de SNPs com expressão alelo-específica (SNPs ASE) e <i>loci</i> quantitativos associados à expressão alelo-específica (aseQTLs) de acordo com a sua posição genômica, mecanismos de regulação preditos e sobreposições com QTLs obtidos na mesma população.....	37
Figura 4. Resultados gerais da análise de associação de SNPs à expressão alelo-específica (ASE).....	38
Figura 5. Distribuição dos <i>loci</i> quantitativos associados à expressão alelo-específica (aseQTLs) em relação à sua distância aos sítios de início de transcrição (TSSs) dos genes com expressão alelo-específica (genes ASE, em rosa) e em relação à sua distância aos SNPs com expressão alelo-específica (SNPs ASE) associados (em azul), em Kb.....	39
Figura 6. Maior bloco de desequilíbrio de ligação (DL) entre os <i>loci</i> quantitativos associados à expressão alelo-específica (aseQTLs).....	40
Figura 7. Principais blocos de desequilíbrio de ligação (DL) entre os SNPs com expressão alelo-específica (SNPs ASE).....	41
Figura 8. Anotação funcional dos <i>loci</i> quantitativos associados à expressão alelo-específica (aseQTLs).....	43
Figura 9. Porcentagem de metilação dos <i>loci</i> quantitativos associados à expressão alelo-específica (aseQTLs) em 12 animais Nelore.....	45
Figura 10. Síntese esquemática dos resultados desse trabalho.....	47
Figura 11. <i>Heatmap</i> da distribuição de <i>loci</i> quantitativos associados à expressão alelo-específica (aseQTLs) de acordo com os genes com expressão alelo-específica (genes ASE) associados e as regiões reguladoras preditas nesse trabalho.....	48

LISTA DE TABELAS

- Tabela 1** - Arquivos com os dados de entrada para utilização no projeto.....30
- Tabela 2** - *Loci* quantitativos associados à expressão alelo-específica (aseQTLs) integrando com QTLs de características de qualidade de carne identificados em músculo de Nelore.....42
- Tabela 3** - Genótipos e taxa de metilação dos *loci* quantitativos associados à expressão alelo-específica (aseQTLs) rs136787973 e rs137789368, associados ao SNP com expressão alelo-específica (SNP ASE) rs42719199, com os respectivos genótipos e valores de desequilíbrio da razão alélica (DRA) em três animais da população.....55

LISTA DE ABREVIATURAS

AOL	Área de olho de lombo
ASE	Expressão alelo-específica, do inglês <i>allele-specific expression</i>
AseQTL	<i>Locus</i> quantitativo associado à expressão alelo-específica
b*	Coloração amarelada da gordura
CAR	Consumo alimentar residual
CpG	Dinucleotídeo de citosina seguido por guanina
CRA	Capacidade de retenção de água
DL	Desequilíbrio de ligação
DRA	Desequilíbrio da razão alélica
EGS	Espessura de Gordura Subcutânea
eQTLs	<i>Locus</i> quantitativo associado à expressão gênica
FC	Força de cisalhamento 24 horas após o abate
FC14	Força de cisalhamento 14 dias após o abate
FC7	Força de cisalhamento 7 dias após o abate
GDE	Gene diferencialmente expresso
GEBV	Valores genéticos estimados
Gene ASE	Gene com expressão alelo-específica
GWAS	Estudos de associação genômica global
Marmoreio	Gordura intramuscular
L*gordura	Luminosidade da gordura
L*músculo	Luminosidade do músculo
MAF	Frequência alélica mínima
miRNA	Micro RNA
mRNA	RNA mensageiro
PC	Perda por cozimento
QTL	<i>Locus</i> quantitativo associado a uma característica
RRBS	Sequenciamento de representação reduzida de bissulfito, do inglês <i>Reduced Representation Bisulfite Sequencing</i>)
SLFT	Sítio de ligação de fator de transcrição
SNP	Polimorfismo de nucleotídeo único, do inglês <i>single nucleotide polymorphism</i>
SNPs ASE	Polimorfismo de nucleotídeo único com expressão alelo-específica
TSS	Sítio de início de transcrição
VEP	<i>Variant Effect Predictor</i> (Ferramenta de anotação do <i>Ensembl</i>)

SUMÁRIO

1. Introdução	10
2. Revisão Bibliográfica.....	12
2.1 Histórico e biologia do gado Nelore.....	12
2.2 Fenótipos de qualidade de carne	13
2.3 Genômica aplicada ao melhoramento	14
2.4 Expressão alelo-específica	16
2.5 Mecanismos de regulação gênica.....	18
2.6 aseQTLs	23
3. Justificativa	25
4. Hipótese.....	26
5. Objetivos	27
5.1 Objetivo geral.....	27
5.2 Objetivos específicos.....	27
6. Material e métodos	28
6.1 Métodos realizados anteriormente pelo grupo de pesquisa.....	28
6.1.1 Criação dos animais e coleta das amostras.....	28
6.1.2 Extração de DNA e genotipagem.....	28
6.1.3 Extração de RNA e sequenciamento	29
6.1.4 Dados de ASE	30
6.2 Análises realizadas nesse trabalho	31
6.2.1 Análise de associação de variantes à expressão alelo-específica (aseQTL).....	31
6.2.2 Análise de desequilíbrio de ligação.....	33
6.2.3 Predição de mecanismos de regulação	33
6.2.4 Integração de dados.....	34
7. Resultados	36
7.1 aseQTLs	36
7.2 Desequilíbrio de Ligação	39
7.3 Integração de dados.....	41
7.4 Mecanismos de regulação	43
7.4.1 Anotação dos SNPs.....	43
7.4.2 aseQTLs em SLFTs.....	44
7.4.3 aseQTLs com metilação.....	44

7.4.4	aseQTLs em sítios de ligação de miRNAs	45
7.4.5	Integração dos resultados	46
8	Discussão.....	49
8.1	aseQTLs	49
8.2	Integração com dados de GWAS	49
8.3	Influência da distância genômica e do desequilíbrio de ligação	50
8.4	Atividade cis-reguladora dos aseQTLs	52
8.5	SLFT podem ser afetados por aseQTLs	52
8.6	A metilação de DNA pode afetar aseQTLs na sua relação com a ASE	53
8.7	Sítios de ligação pra miRNAs são potencialmente modificados por alelos de aseQTLs	55
8.8	Integração dos resultados e conclusões finais	56
9	Conclusão.....	58

1. Introdução

Abordagens genômicas são utilizadas para se executar múltiplas análises de associação de polimorfismos a determinado fenótipo de interesse, o que é feito através de ferramentas de associação genômica global (GWAS, do inglês *genome-wide association studies*). Segundo Korte e Farlow (2013), esse tipo de abordagem pode identificar novos loci candidatos, pois através da varredura de todo o genoma tais ferramentas encontram polimorfismos de nucleotídeo único (SNPs, do inglês *single nucleotide polymorphisms*) com capacidade de exercer atividades reguladoras direta ou indiretamente, alterando a expressão de genes específicos e conduzindo a alterações fenotípicas que podem afetar características de interesse tais como a eficiência alimentar (HIGGINS et al., 2018), qualidade do leite (MANCINI et al., 2013) e fertilidade bovina (COCHRAN et al., 2013).

Alguns fenótipos de interesse na pecuária são de suma importância para a movimentação do mercado de carnes, pois afetam diretamente a aceitação do consumidor. Dentre esses, os mais relevantes são a cor, a quantidade de gordura (RESURRECCION, 2004), a aparência e a maciez (KILLINGER et al., 2004). Diante disso, buscando os determinantes moleculares envolvidos no parâmetro de maciez, (DE SOUZA et al., em revisão¹) obtiveram o perfil da expressão alelo-específica (ASE, do inglês *allele-specific expression*) de SNPs no músculo de animais da raça Nelore (*Bos indicus*) adultos, sendo grande parte deles contidos em genes responsáveis por funções biológicas importantes para o desenvolvimento do tecido muscular e relacionados à maciez da carne.

Tendo em vista o grande potencial de se encontrar polimorfismos reguladores da expressão gênica no genoma, é de grande relevância a realização de uma abordagem genômica que busque SNPs que possam estar associados à ASE como forma de identificar potenciais mecanismos de regulação. AseQTLs (do inglês *allele-specific expression quantitative trait loci*) são SNPs com regulação de ação em *cis* associados à ASE, que podem estar envolvidos com vários mecanismos reguladores.

O presente trabalho identificou variantes reguladoras responsáveis pela ASE utilizando inicialmente o teste de associação entre variantes e o fenótipo de expressão alelo-específica (aseQTL) (BATTLE et al., 2014). Em seguida, os SNPs estatisticamente relacionados à ASE foram analisados quanto a indicativos de mecanismos reguladores envolvidos, tais como ocorrência e localização de metilação de DNA no SNP, se estes podem modificar sítios de ligação de fatores de transcrição (SLFTs) expressos em músculo bovino ou sítios de ligação de

¹ DE SOUZA, M. M. et al. Allele-specific expression is widespread in *Bos indicus* muscle and affects meat quality candidate genes. **Scientific Reports**, [s.d.].

miRNAs expressos em músculo da mesma população. Depois disso, estes aseQTLs foram combinados com dados de GWAS associados a características de qualidade de carne e de eficiência alimentar obtidos a partir da mesma população experimental, para identificar possíveis aplicações dos conhecimentos acerca da expressão e regulação gênica sobre fenótipos de interesse agropecuário.

2. Revisão Bibliográfica

2.1 Histórico e biologia do gado Nelore

Os bovinos atuais evoluíram a partir dos auroques (*Bos primigenius*). Esse processo pode ter ocorrido em uma população ancestral cujos haplótipos sobreviveram e foram se tornando específicos a determinadas populações, sendo posteriormente recombinados conforme as gerações, formando haplótipos relacionados às raças bovinas (WILLIAMSON et al., 2011). Existem evidências de que os auroques passaram por dois processos de domesticação responsáveis pela divergência entre os dois grandes grupos de bovinos: o gado europeu (*Bos taurus*) e o gado indiano ou zebuino (*Bos indicus*) (LOFTUS et al., 1994).

Os primeiros rebanhos bovinos chegaram ao Brasil após a conquista das Américas, trazidos de Portugal. Somente no século XIX, os primeiros zebras das raças Guzerá, Gir e Nelore foram introduzidos no país (AJMONE-MARSAN et al., 2010). Esses animais se adaptaram bem ao clima tropical, se reproduzindo e formando novas raças, como Tabapuã e Indubrasil (SANTIAGO, 1987). A raça Nelore é predominante no gado brasileiro (SANTANA et al., 2016), fazendo parte de cerca de 80% dos rebanhos nacionais, constituídos por animais puros e oriundos de cruzamentos (DOS SANTOS, 2000).

A partir de 2019, o consumo de carne bovina tem uma projeção de aumentar em taxas de 0,95 ao ano, sendo o Brasil o responsável por 6,2% do aumento do consumo de carne global até 2029 (considerando a carne suína, bovina e de frango) (AGRICULTURAL INTERAGENCY COMMITTEE, 2019). Além disso, a produção de carne brasileira tende a ser maior que o consumo, o que pode alavancar ainda mais a exportação, visto que o país está sempre entre os maiores exportadores de carne bovina (AGRICULTURAL INTERAGENCY COMMITTEE, 2019).

Os animais *B. taurus* e *B. indicus*, apesar de descenderem de um ancestral comum, apresentam diferentes atributos relevantes para o interesse agropecuário. Em relação aos aspectos físicos, a principal diferença é que nos zebras está presente uma corcunda (cupim), a pelagem é mais curta e a pele pouco aderida ao corpo comparada à de animais taurinos (SANTIAGO, 1987). Uma característica importante do Nelore é a sua adaptabilidade a climas tropicais, sendo possuidor de uma considerável termotolerância (BARBOSA et al., 2014; DE MELO COSTA et al., 2018). Além disso, o animal também apresenta resistência a parasitas (BRICARELLO et al., 2007).

Quanto às características de carne, é importante considerar que os animais Nelore produzem carne menos macia quando comparados a determinadas raças taurinas (DUARTE et al., 2013; LAGE et al., 2012; PEREIRA et al., 2015). Pelo estudo da proteômica, autores mostraram que, em músculo de animais Nelore, está presente maior concentração de proteínas relacionadas à contração muscular. O grau de degradação dessas proteínas é diretamente relacionado à maciez da carne (RODRIGUES et al., 2017).

Tendo em vista a participação do Brasil no mercado internacional de carnes, e a boa adaptação e produção do Nelore no país, é necessária a aplicação de programas de melhoramento da carne deste animal visando aumentar o valor do produto final (BROOKS et al., 2000), o que traria impactos positivos na economia brasileira.

2.2 Fenótipos de qualidade de carne

A qualidade da carne é influenciada por diversos fatores, sendo a sua mensuração realizada somente após o abate do animal (TIZIOTO et al., 2013). A maciez é um dos principais fenótipos de qualidade de carne, pois afeta diretamente a aceitação por parte do consumidor (HUFFMAN et al., 1996; MILLER et al., 2001). Logo após o abate, o músculo do bovino se contrai e enrijece, num processo chamado de *rigor mortis*. Durante o tempo de envelhecimento da carne, após o abate do animal, ocorre o processo de amaciamento, atingindo as maiores taxas durante a primeira semana (MARINO et al., 2013). O processo de envelhecimento da carne oferece as condições necessárias para garantir a qualidade da conversão do músculo em carne, dependendo de fatores tais como raça, temperatura, pH, tempo de envelhecimento, tipos de fibras, tamanho do sarcômero e proteólise (BHAT et al., 2018; MARINO et al., 2013). A proteólise é um dos principais fatores que conduzem ao aumento da maciez do músculo. As proteínas desmina e troponina T, por exemplo, possuem funções estruturais no músculo esquelético, e a degradação das mesmas é essencial para o processo de amaciamento da carne (PENNY; DRANSFIELD, 1979; TAYLOR et al., 1995). Diferentes isoformas da enzima proteolítica calpaína degradam proteínas miofibrílicas em resposta a diferentes níveis de cálcio (BHAT et al., 2018). Um dos métodos mais utilizados para mensurar a maciez da carne é através da força de cisalhamento, comumente calculada a partir de um texturizador acoplado a uma lâmina de Warner-Bratzler, sendo esse método quantitativo mais acurado que a análise sensorial (DESTEFANIS et al., 2008). Quanto maior a força de cisalhamento, menos macia é a carne.

O consumidor apresenta preferências quanto à aparência da carne, com critérios que se baseiam na cor da gordura e do músculo (JEREMIAH; CARPENTER; SMITH, 1972). As

diferentes cores de gordura e concentrações de carotenoides ocorrem em resposta à composição de ácidos graxos resultante das dietas dos animais (DUNNE et al., 2009). A composição dos ácidos graxos do tecido adiposo é relacionada à firmeza da gordura, pois esses ácidos graxos possuem pontos de fusão específicos (WOOD et al., 2008). Além disso, o conteúdo de gordura intramuscular forma o marmoreio, que consiste na aparência marmorizada da gordura que entremeia a carne, e é relacionado à suculência da carne, e também afeta a aceitabilidade do produto no mercado (CHENG et al., 2015b; DUNNE et al., 2009; WOOD et al., 2004).

2.3 Genômica aplicada ao melhoramento

Entre as décadas de 70 e 80 foram desenvolvidas técnicas de melhoramento animal nas quais polimorfismos marcadores eram identificados a partir do DNA dos animais (LOURENCO et al., 2017). A abordagem de seleção assistida por marcadores considera que o melhoramento animal pode ser resultante da seleção de pais que possuem os marcadores desejados, o que é eficaz para características afetadas por poucos genes (LOURENCO et al., 2017). Entretanto, quando o programa de melhoramento tem como objetivo características complexas, múltiplos genes podem estar associados a tais fenótipos. Assim, a solução encontrada para aprimorar a acurácia do programa de seleção foi a estimação dos valores genéticos genômicos (GEBV, do inglês *genomic estimated breeding values*) a partir do fenótipo e pedigree de animais amostrados de uma população, junto a informações de SNPs dos seus genótipos, que apresentam maior densidade de marcadores e aos quais é atribuída a maior parte da variação do genoma (LOURENCO et al., 2017; SCHEFERS; WEIGEL, 2012; WIGGANS et al., 2017).

Abordagens GWAS são capazes de identificar novos genes e polimorfismos potencialmente associados às características (SCHEFERS; WEIGEL, 2012). Entretanto, com o aumento da densidade de SNPs a serem testados, tende a ser grande o número de falsos positivos, devido aos múltiplos testes realizados e aos numerosos SNPs resultantes dessas análises (SCHEFERS; WEIGEL, 2012). Esse fator aumenta a probabilidade de erros de tipo I, (falsos positivos) sendo muitas vezes decorrente de associações indiretas devido a desequilíbrio de ligação (DL) (NYHOLT, 2004; SCHEFERS; WEIGEL, 2012), que permite que SNPs proximamente localizados no cromossomo sejam herdados em blocos ao invés de individualmente (INTERNATIONAL HAPMAP CONSORTIUM., 2005). Assim, um único SNP pode não apresentar grande efeito sobre uma característica, mas pode estar em um haplótipo com forte DL e associado ao fenótipo em estudo. Nesse ponto, é importante

considerar que, além da distância física entre SNPs, a deriva também pode contribuir para o DL (TEMPLETON, 2011).

O poder estatístico de testes de associação é diretamente relacionado aos tamanhos amostrais e às frequências alélicas (JOHNSON et al., 2010). Assim, os resultados das associações geralmente devem passar por correções de múltiplos testes, o que diminui consideravelmente o número de SNPs resultantes da associação. Os métodos mais utilizados são os que corrigem o valor p , como a correção de *False Discovery Rate* (FDR; taxa de falsas descobertas ou falsos positivos) (BENJAMINI; HOCHBERG, 1995) e Bonferroni (ROSENTHAL; RUBIN, 1984). Apesar de diminuir o número de falsos positivos, esse tipo de correção pode aumentar a taxa de erros do tipo II, que é quando se aceita uma hipótese nula falsa. A conservativa correção de Bonferroni, por considerar que todas as comparações são independentes e assim ignorando a existência de blocos de DL, pode remover os SNPs verdadeiramente associados ao fenótipo dos resultados significativos (JOHNSON et al., 2010). Testes de permutação também são utilizados para controlar a inflação da taxa de falsos positivos, por meio do cálculo da probabilidade de associações ocorrerem devido ao acaso no conjunto de dados (CHURCHILL; DOERGE, 1994). Quando o objetivo do estudo é testar variantes individualmente, e se deseja evitar ao máximo a perda de dados verdadeiramente positivos, é recomendado utilizar valores de significância FDR menos restritivos (LEE; LEE, 2018).

Pode-se utilizar estratégias para diminuir o número de testes, aumentando o poder estatístico. Existem diversas abordagens, como TagSNP, que se baseiam em testar apenas os principais SNPs dentro de blocos de DL (JOHNSON et al., 2010; PATNALA; CLEMENTS; BATRA, 2013; WANG et al., 2017) e a utilização de janelas genômicas, nas quais apenas são associados SNPs em determinada região do genoma, como em regiões promotoras (KONHEIM; WOLFORD, 2003), regiões potencialmente reguladoras (WANG et al., 2018) ou aquelas que compreendem potenciais *cis*-reguladores (CESAR et al., 2018; WATANABE et al., 2017). Estudos ainda mostraram estratégias que priorizam genes e SNPs após a GWAS de acordo com critérios específicos (como o efeito da variante, localização genômica e anotação funcional), que podem agregar aos benefícios oferecidos pelos métodos de associação em prol da descoberta de novos elementos associados a fenótipos complexos (PATNALA; CLEMENTS; BATRA, 2013; XIANG et al., 2019).

Muitos estudos recentes demonstram *loci* de traços quantitativos (QTL, do inglês *quantitative trait loci*) associados a características de interesse agropecuário. Em Nelore, por exemplo, foram identificados QTLs associados a fenótipos de qualidade de carne (TIZIOTO et

al., 2013), ao conteúdo de minerais presente em músculo (TIZIOTO et al., 2015a), ao conteúdo e deposição de marmoreio (CESAR et al., 2014), e QTLs associados a fenótipos de eficiência alimentar (DE OLIVEIRA et al., 2014). Além disso, estudos transcriptômicos também identificaram genes diferencialmente expressos (GDEs) entre grupos de animais divergentes para conteúdo de ácidos graxos (CESAR et al., 2016), marmoreio (CESAR et al., 2015), fenótipos de carcaça (SILVA-VIGNATO et al., 2017), de eficiência alimentar (TIZIOTO et al., 2015b) e concentração mineral (AFONSO et al., 2019).

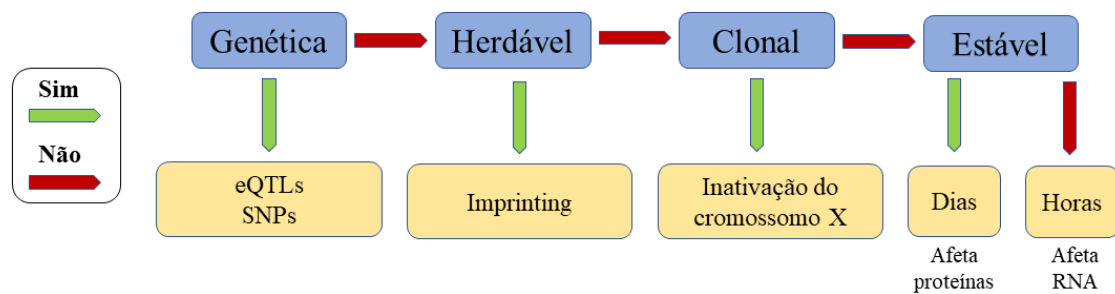
Um QTL pode se apresentar como um potencial eQTL (NICOLAE et al., 2010), capaz de afetar a expressão gênica, sendo um aliado aos resultados de GWAS na busca de SNPs reguladores e genes com funções biológicas essenciais para a característica em estudo, fornecendo informações sobre os processos biológicos envolvidos com os fenótipos complexos (NICOLAE et al., 2010). Em músculo de Nelore foram identificados eQTLs com ações em *cis* e *trans* sobre a expressão gênica, afetando genes envolvidos com fenótipos de marmoreio (CESAR et al., 2018). São considerados *cis*-eQTLs os SNPs que regulam a expressão de genes localizados próximos dessa variante, com menos de 1Mb de distância entre si (CESAR et al., 2018; YAO et al., 2017). Foi também sugerido que a classificação de um *cis*-eQTL estaria relacionada à capacidade do SNP em afetar alelos presentes na mesma fita de DNA (MASON et al., 2018). Então, devido a esse padrão de regulação alelo-específico, *cis*-eQTLs podem estar associados a genes com ASE (KHANSEFID et al., 2018; MASON et al., 2018).

2.4 Expressão alelo-específica

A ASE é um padrão que resulta na expressão desbalanceada entre as duas cópias de um gene em cada par cromossômico (alelos) (YAN et al., 2002). Estudos recentes têm explorado os mecanismos reguladores associados à ASE e ao *imprinting* genômico (BATTLE et al., 2014; GUILLOCHEAU et al., 2019; JIANG et al., 2018; TRIANTAPHYLLOPOULOS; IKONOMOPOULOS; BANNISTER, 2016). O *imprinting* genômico ocorre quando a ASE é guiada por efeito parental, em decorrência de modificações *cis*-reguladoras e epigenéticas, podendo levar à expressão monoalélica de genes essenciais para o desenvolvimento e crescimento de mamíferos (FERGUSON-SMITH, 2001). A ASE pode ser classificada em diferentes subtipos. Se a ASE é resultante de efeitos genéticos, ela pode ser causada por rearranjos de DNA, eQTLs e SNPs. Se é herdável (e se mantém nos alelos provenientes de cada parental), a ASE é afetada pelo *imprinting* genômico. Se é clonal, indica se o evento pode ser herdado mitoticamente da célula mãe para célula filha, causada por inativações no cromossomo

X e expressão monoalélica autossomal. Além disso a ASE pode ter diferentes tempos de estabilidade, afetando proteínas quando é estável por dias e afetando o RNA quando estável por horas (KRAVITZ; GREGG, 2019). Um esquema para essa identificação está presente na Figura 1.

Figura 1 – Subtipos de expressão alelo-específica (ASE).



De acordo com o esquema para a classificação da ASE, quando a ASE é resultado de mecanismos genéticos, pode ser influenciada por eQTLs e SNPs. Quando é causada por mecanismos não-genéticos e o efeito é herdável do parental para a prole, pode ser decorrente de efeito de Imprinting, e quando o padrão da ASE se dá pela clonagem mitótica da célula-mãe para a célula filha, pode ser decorrente de inativação do cromossomo X. A estabilidade da ASE não-genética pode afetar proteínas, se estável por dias, ou afetar o RNA, se estável por poucas horas.

Adaptado de: KRAVITZ; GREGG. (2019)

Em bovinos, foi demonstrado que a ASE é tecido-específica (CHAMBERLAIN et al., 2015), e o padrão de ASE para SNPs presentes em *chips* comerciais foi recentemente descrito em músculo de bovinos Limousin (GUILLOCHEAU et al., 2019) e Nelore (DE SOUZA et al., em revisão)¹, indicando que a ASE está amplamente distribuída em SNPs do músculo bovino (SNPs ASE) e pode afetar genes com funções biológicas relacionadas ao desenvolvimento muscular e fenótipos de qualidade da carne. Genes ASE (genes com ASE ou aqui chamados os genes que contêm os SNPs ASE) podem afetar características de interesse na agropecuária, como a doença de Marek em galinhas (MEYDAN; YILDIZ, 2018), metabolismo de lipídeos e adipogênese em suínos (STACHOWIAK; FLISIKOWSKI, 2019; STACHOWIAK; SZCZERBAL; FLISIKOWSKI, 2018). Assim é demonstrado o potencial de se estudar a ASE no contexto da genética animal.

¹ DE SOUZA, M. M. et al. Allele-specific expression is widespread in *Bos indicus* muscle and affects meat quality candidate genes. **Scientific Reports**, [s.d.].

Abordagens genômicas de ASE comparam os níveis de expressão gênica entre os dois alelos de indivíduos heterozigotos (DELBARE; CLARK, 2018), podendo colaborar com o estudo de variantes reguladoras com ação em *cis* (DELBARE; CLARK, 2018; PASTINEN, 2010) e eQTLs (ZOU et al., 2018), que possuem mecanismos de regulação envolvidos com esse padrão. Esses fatores podem aprimorar programas de melhoramento animal por aumentar a acurácia de modelos preditivos (CHENG et al., 2015a).

2.5 Mecanismos de regulação gênica

O processo de transcrição do mRNA é altamente regulado. Em resumo, a ativação dos promotores dos genes por meio da ligação de fatores de transcrição (FTs) em seus sítios forma o complexo de pré-iniciação, que recruta também a RNA polimerase II (LUSE, 2013; NOGALES; LOUDER; HE, 2017; YAO, 2017). Essa enzima é responsável pela síntese da nova cadeia de RNA pela adição de nucleotídeos na extremidade 3' do RNA nascente (KWAK; LIS, 2013), sendo interrompida a transcrição após o reconhecimento de uma sequência de terminação, onde logo em seguida são liberados o RNA sintetizado, a RNA polimerase II e o DNA molde (ARIMBASSERI; RIJAL; MARAIA, 2013).

Insulators são elementos reguladores que delimitam na cromatina as regiões onde pode existir a influência de outros mecanismos reguladores (BUSHEY; DORMAN; CORCES, 2009), podendo bloquear os *enhancers*, responsáveis pelo recrutamento de FTs e regulação distal, não permitindo a comunicação deles com os promotores de transcrição (GASZNER; FELSENFELD, 2006; SPITZ, 2016). Esses fatos sugerem que os *insulators* e *enhancers* estão ativamente envolvidos na organização nuclear e regulação de expressão gênica (ONG; CORCES, 2011). Assim, o processo de ativação destes elementos deve ser controlado, o que acontece com a compactação e descompactação da fibra da cromatina, que limita o acesso aos *enhancers* e outros elementos reguladores, para garantir que a transcrição, replicação, reparo e recombinação de DNA ocorram corretamente (CLAPIER; CAIRNS, 2009). Elementos *cis*-reguladores também podem controlar o *splicing* alternativo do pré-mRNA (LOPEZ, 1998), o que permite a transcrição de diferentes isoformas, promovendo a diversificação do transcriptoma e proteoma do organismo (LEE; RIO, 2015).

Além dos mecanismos genéticos, existem mecanismos epigenéticos, que são aqueles que refletem a herança e regulação de padrões de expressão gênica que não são relacionados à sequência do DNA, mas sim à conformação da cromatina (ALLIS; JENUWEIN, 2016). A cromatina é caracterizada por DNA complexado a histonas organizado em nucleossomos, de

forma que cada nucleossomo seja formado por quatro pares de histonas (um octâmero) que englobam 147 bp de DNA (ALLIS; JENUWEIN, 2016). Quando a cromatina está condensada, é identificada como heterocromatina, uma conformação que dificulta a ativação da transcrição. Por outro lado, a eucromatina é caracterizada por uma conformação menos densa, que facilita o acesso a iniciadores de transcrição.

Dentre os principais mecanismos epigenéticos estão a metilação de DNA, onde a metilação de citosinas pode levar à repressão da expressão (TIAN, 2014), modificações histônicas, como a acetilação, associadas à ativação da expressão gênica (VERDIN; OTT, 2015), remodelamento nucleossômico que, através da alteração da interação entre o DNA e a histona, promove a acessibilidade do DNA a FTs (BECKER; WORKMAN, 2013) e RNAs não codificantes, que podem guiar complexos modificadores de cromatina para locais genômicos específicos (KHALIL et al., 2009; MERCER; MATTICK, 2013). Além da regulação que atua sobre a região promotora dos genes, o mRNA também sofre a influência de mecanismos pós-transcricionais. Um mecanismo pós-transcricional com variados modos de ação é mediado por miRNAs, que reprimem a tradução e degradam o mRNA (CAI et al., 2009; JONAS; IZAURRALDE, 2015). Nas próximas subseções serão abordados alguns dos mecanismos amplamente relacionados com a regulação da ASE.

Fatores de Transcrição

A regulação gênica depende da ação conjunta de vários mecanismos, sendo os FTs um dos principais reguladores do início da transcrição (SPITZ; FURLONG, 2012). Os FTs geralmente reconhecem pequenas seqüências de DNA degeneradas, de 6 a 12 pb de comprimento, o que indica a influência de outros fatores sobre a afinidade de um FT pelo seu sítio (SPITZ; FURLONG, 2012). A ligação de um FT aos elementos regulatórios, *enhancers* e promotores, depende de fatores como a sinalização extracelular, a presença de diferentes sítios de ligação para múltiplos FTs e o posicionamento do nucleossomo, pois histonas e FTs podem competir pelo acesso ao DNA (SPITZ; FURLONG, 2012; YAO, 2017).

A RNA Polimerase II se liga aos fatores de transcrição TFIIB, TFIID, TFIIE, TFIIIF e TFIIH e ao DNA do promotor do gene cuja transcrição será iniciada, formando o complexo de pré-iniciação (LUSE, 2013), momento em que os FTs se apresentam na sua forma funcional (REESE, 2003; SAINSBURY; BERNECKY; CRAMER, 2015). Na presença de nucleosídeos trifosfatos, é formada uma “bolha de transcrição” e o complexo promotor aberto, onde a fita

molde de DNA passa próximo ao sítio ativo da RNA Polimerase II, e a fita de RNA complementar é sintetizada a partir do DNA (SAINSBURY; BERNECKY; CRAMER, 2015).

A presença de sítios de ligação de fatores de transcrição (SLFTs) pode promover ou suprimir a transcrição no respectivo alelo (SERRE et al., 2008). Assim, se a região reguladora de apenas uma das cópias do gene possui afinidade com o FT, essa determinada região apresenta ação em *cis* sobre a expressão do alelo, o que, em um *locus* heterozigoto poderia ser observado pelo padrão de ASE (KNOWLES et al., 2017).

Similarmente ao impacto da presença de SLFTs em diferentes alelos, a ASE pode ser regulada pela metilação alelo-específica do DNA.

Metilação de DNA

A metilação reflete na ligação de um grupo metil na região 5' do DNA, processo esse mediado por proteínas metiltransferases, embora aconteça a metilação em outras regiões genômicas, com sentido biológico ainda desconhecido (EDWARDS et al., 2017). Em mamíferos, normalmente esse processo ocorre em ilhas CpG (EDWARDS et al., 2017), que se caracterizam por regiões do DNA com agrupamentos de dinucleotídeos de citosina seguidos por guanina.

A metilação de DNA é uma modificação epigenética capaz de suprimir a transcrição quando as ilhas CpG estão densamente metiladas e permitir a transcrição quando houverem poucos eventos de metilação isolados naquela região (TIAN, 2014). As ilhas CpG estão presentes mais comumente nos promotores, sendo a densidade de CpG proporcional aos níveis de metilação, pois regiões com pouco conteúdo CpG são parcialmente metiladas, muitas vezes com níveis mínimos incapazes de regular a expressão gênica. Ilhas CpG próximas aos TSSs são normalmente mais curtas e possuem menor probabilidade de serem metiladas do que ilhas distantes dos TSSs (EDWARDS et al., 2010).

O acesso das metiltransferases ao DNA pode ser limitado pela conformação da cromatina, que é determinada pela disposição dos nucleossomos (NILLER; DEMCSÁK; MINAROVITS, 2018). Proteínas de ligação ao grupo metil interagem com a metiltransferase DNMT1, podendo se complexar com histonas desacetilases e silenciar o promotor (NILLER; DEMCSÁK; MINAROVITS, 2018). As histonas desacetilases, junto à DNMT1, removem os grupos acetil dos resíduos de lisina e arginina das histonas, fator responsável pela condensação da cromatina, que resulta em uma conformação transcricionalmente inativa (QIN; LEONHARDT; PICHLER, 2011).

Entretanto, apesar de ser amplamente verificada a associação da metilação de DNA com a repressão da expressão gênica, ainda existem controvérsias quanto à sua relação causa e consequência. Por exemplo, quando uma célula passa por uma infecção, ocorre intensa demetilação em sítios CpG que se sobrepõem a *enhancers*, porém posteriormente à ativação transcricional decorrente da resposta imune (PACIS et al., 2019). Nesse caso, a diminuição brusca dos níveis de metilação é mediada pela ligação de FTs, ocorrendo não como uma causadora do processo de ativação da expressão gênica, mas como uma consequência *downstream* à ligação de FTs na região reguladora (PACIS et al., 2019). Foi demonstrado que as ilhas CpG da região promotora são normalmente protegidas da metilação, sendo a ligação de FTs um dos mecanismos responsáveis por esse padrão (EDWARDS et al., 2010; JELTSCH; BROCHE; BASHTRYKOV, 2018; LONG et al., 2016). A metilação de *enhancers* e SLFTs em promotores com pouco conteúdo CpG pode afetar de formas diferentes a expressão gênica, dependentemente do efeito do FT (ativador ou repressor da transcrição) (JELTSCH; BROCHE; BASHTRYKOV, 2018).

A metilação diferencial em regiões reguladoras dos dois alelos provenientes de cada parental pode causar a ASE (TIAN, 2014). A metilação de DNA em ilhas CpG é uma marca epigenética de *imprinting* genômico (PFEIFER, 2000). Normalmente, regiões sob o controle de *imprinting* regulam a expressão de agrupamentos de genes com esse padrão de expressão e são protegidos da demetilação (NILLER; DEMCSÁK; MINAROVITS, 2018). Esse padrão é preservado conforme as gerações, pois as marcas de metilação do DNA depositadas durante a gametogênese nos *loci imprinted* são mantidas mesmo após a fertilização nos locais com *imprinting*, permanecendo apesar da intensa demetilação do DNA que ocorre no embrião pré-implantado (BARTOLOMEI; FERGUSON-SMITH, 2011).

MicroRNAs

Os miRNAs são pequenos RNAs não codificantes capazes de inibir a tradução por se ligarem à sequência alvo do mRNA (JONAS; IZAURRALDE, 2015). Os primeiros trabalhos que identificaram a função dos miRNAs foram a partir de análises dos pequenos RNAs let-7 e lin-4 de *C. elegans* (LEE; FEINBAUM; AMBROS, 1993), os quais se mostraram com capacidade de se ligar a sítios específicos da região 3' do mRNA alvo, agindo como um repressor pós-transcricional (LEE et al., 2002).

Genes que codificam para miRNAs estão proximamente localizados no genoma, formando agrupamentos de genes com o mesmo padrão de expressão, o que indica que a transcrição de miRNAs pode ser controlada por um único promotor (LEE et al., 2002). Lee e

colaboradores (2002) realizaram um dos primeiros trabalhos sobre a biogênese dos miRNAs, separando o processo em duas etapas, sendo que na primeira ocorre o processamento do transcrito em pré-miRNA, e na segunda etapa esse precursor é processado e torna-se o miRNA maduro. A transcrição do miRNA localizado dentro de um gene é realizada pela RNA Polimerase II, simultaneamente à transcrição desse gene hospedeiro. Quando o miRNA está localizado em regiões intergênicas ele é transcrito independentemente e possui um promotor próprio. Na primeira etapa do processamento, os transcritos primários ou pri-miRNAs têm uma estrutura de gancho reconhecida por um complexo microprocessador que possui a função de gerar um novo pré-miRNA de 60 nucleotídeos, cuja conformação permite que o miRNA interaja com proteínas que medeiam o transporte do miRNA ao citoplasma. Na segunda etapa do processamento, o complexo formado pela proteína transportadora e o RNA é reconhecido pela RNase III Dicer, que corta o pré-miRNA em uma forma madura de miRNA (DYKES; EMANUELI, 2017; GULYAEVA; KUSHLINSKIY, 2016).

O miRNA maduro regula a expressão gênica por vários mecanismos. Fatores epigenéticos podem ser afetados por miRNAs, como por exemplo DNA metiltransferases e histonas desacetilases, podendo modificar a conformação da cromatina (BIANCHI et al., 2017). A presença de hipermetilação pode diminuir a expressão de miRNAs (WANG et al., 2015), sendo também descrita uma relação reguladora entre miRNAs e RNAs longos não codificadores (DYKES; EMANUELI, 2017). O modo de ação mais clássico envolvendo miRNAs é a ligação entre as sequências “seed” do miRNA ao mRNA transcrito a partir de genes-alvo, seguido pela degradação desse mRNA (CAI et al., 2009; GULYAEVA; KUSHLINSKIY, 2016; JONAS; IZAURRIALDE, 2015), mas também existem miRNAs que podem realizar a deadenilação dos mRNAs afetados, bem como ativar seus alvos, ao invés de reprimi-los (CHEKULAEVA; FILIPOWICZ, 2009). O principal modo de repressão dos miRNAs sobre o mRNA alvo em mamíferos não exige o emparelhamento perfeito e não necessariamente degrada o mRNA, pois essa repressão depende da interação com outras proteínas. O miRNA pode se associar à proteína Ago, que recruta uma proteína adaptadora que interage com proteínas da cauda Poli A e recruta complexos deadenilases. A deadenilação encurta a cauda Poli A, desestabilizando o mRNA e afetando a tradução (BARTEL, 2018).

Um miRNA individual pode regular mais de um mRNA, e vários miRNAs podem regular o mesmo mRNA (CAI et al., 2009). A interação entre mRNA e miRNAs geralmente se faz através da ligação da região *seed* 5' do miRNA com a 3' do mRNA, podendo também ocorrer interações entre a 3' do miRNA com a 5' do respectivo alvo (CAI et al., 2009). Em menor frequência, miRNAs podem regular regiões intrônicas do mRNA (MENG et al., 2013). Estudos

relacionaram miRNAs e o modelo bovino de diferentes formas. Existem evidências de que miRNAs presentes em leite bovino e ovos de galinha podem interferir na regulação gênica de humanos, podendo ser promissora a aplicabilidade de uma dieta com base em miRNAs relacionados a genes funcionalmente interessantes (ZEMPLINI et al., 2015). Os auroques, animais primitivos que deram origem aos bovinos, provavelmente passaram por processos mutacionais em miRNAs e seus respectivos sítios de ligação, o que pode ter sido um dos fatores responsáveis pelas diferentes características que distinguem as raças taurinas atualmente (BRAUD et al., 2017). Foram descritos miRNAs que podem regular o desenvolvimento do tecido adiposo bovino (JIN et al., 2010), sendo que em Nelore, análises de expressão de miRNAs em músculo e fígado sugerem que os miRNAs podem afetar o consumo alimentar residual (CAR) (DE OLIVEIRA et al., 2018), um fenótipo associado à eficiência alimentar, e marmoreio (OLIVEIRA et al., 2018), demonstrando sua potencial aplicabilidade no melhoramento da raça.

As análises alelo-específicas de sítios de ligação de miRNA fornecem mais informações para a compreensão de variantes com ação em *cis* sobre a expressão gênica, facilitando a busca por genes alvos de características complexas (JACINTA-FERNANDES et al., 2019). A ASE pode ser causada pela presença diferencial de sítios de ligação para miRNA entre os dois alelos (TAN et al., 2007; TCHATCHOU et al., 2009), e pode regular a expressão de genes com funções interessantes para a produção e qualidade animal.

2.6 aseQTLs

Como demonstrado pela ação alelo-específica dos mecanismos de regulação, alterações na expressão de dois alelos podem ocorrer em resposta a variações *cis*-reguladoras entre as sequências. Os SNPs que afetam essas regiões reguladoras podem ser identificados a partir de uma associação entre heterozigotos no contexto da ASE e diferentes genótipos de SNPs candidatos a reguladores (BATTLE et al., 2014; MOHAMMADI et al., 2017; ZOU et al., 2018).

Para descobrir mecanismos reguladores que causam a ASE, é possível executar uma análise de triagem inicial para identificar polimorfismos associados à expressão desigual de alelos: os aseQTLs (BATTLE et al., 2014; ZHANG; CAL; BOREVITZ, 2011). A partir dessa varredura inicial, aseQTLs relacionados a fenótipos interessantes podem ser priorizados para análises adicionais, a fim de entender a regulação da expressão de genes associados a essas características econômicas e como eles podem ser expressos na prole (COCHRAN et al., 2013;

HIGGINS et al., 2018; MANCINI et al., 2013; SANTIAGO et al., 2018; SEABURY et al., 2017b; TIZIOTO et al., 2013).

Aqui, investigamos aseQTLs associados a SNPs ASE descritos anteriormente em uma população experimental de animais Nelore (DE SOUZA et al., em revisão¹). Depois disso, analisamos seu envolvimento com mecanismos reguladores, como a metilação do DNA, e se eles foram inseridos nos sítios de ligação a FTs ou a miRNAs. Este estudo adiciona uma nova camada de informações genômicas em Nelore, descrevendo possíveis mecanismos envolvidos na expressão alélica de regiões associadas a fenótipos econômicos relevantes.

¹DE SOUZA, M. M. et al. Allele-specific expression is widespread in *Bos indicus* muscle and affects meat quality candidate genes. **Scientific Reports**, [s.d.].

3. Justificativa

A ocorrência de expressão alelo-específica em músculo de Nelore foi demonstrada, porém ainda são necessários estudos que investiguem a participação de variantes genéticas e mecanismos reguladores desse padrão. A identificação de tais variantes deve contribuir para o conhecimento de regiões e mecanismos reguladores do genoma de *Bos indicus* e para a associação dessas variantes com fenótipos de interesse agropecuário.

4. Hipótese

Variações de sequência de DNA estão associadas a padrões de expressão alelo-específica em músculo de Nelore.

5. Objetivos

5.1 Objetivo geral

- Predizer os mecanismos pelos quais SNPs possam regular a ASE de SNPs relacionados a fenótipos de interesse agropecuário em bovinos Nelore.

5.2 Objetivos específicos

- Identificar SNPs candidatos a ter efeito regulador sobre a ASE, por abordagem de aseQTL;
- Identificar possíveis mecanismos reguladores pelos quais os aseQTLs possam afetar a ASE, como metilação de DNA, SLFTs e miRNAs;
- Integrar dados obtidos na mesma população para relacionar essas variantes a fenótipos de interesse agropecuário.

6. Material e métodos

6.1 Métodos realizados anteriormente pelo grupo de pesquisa

Nesta subseção estão os métodos utilizados anteriormente pelo grupo de pesquisa, de onde foram obtidos os dados para a execução deste trabalho.

6.1.1 Criação dos animais e coleta das amostras

A produção dos novilhos Nelore foi realizada mediante aprovação do comitê de uso e bem-estar animal da EMBRAPA, e em compartilhamento com outros trabalhos envolvidos no projeto em rede “Bife de Qualidade”, a partir do qual muitos estudos complementares foram realizados.

Para a seleção dos animais que produziram a nossa população, foi utilizado o sêmen de 34 touros pouco aparentados representantes das principais linhagens da raça Nelore, monetariamente acessíveis a quaisquer produtores do Brasil, mediante consulta a catálogos. Foram realizadas as inseminações e os novilhos nascidos cresceram até os 18 meses em três fazendas, sendo depois confinados recebendo a mesma nutrição e manejo em três estações de monta de 90 a 130 dias na Embrapa Pecuária Sudeste (São Carlos, São Paulo, Brazil). O abate foi realizado com os animais por volta dos 25 meses de idade, pesando em média 383,2 Kg. Nesse momento foram coletadas as amostras de sangue e do músculo *L. thoracis*.

6.1.2 Extração de DNA e genotipagem

O DNA dos touros foi extraído do sêmen congelado, pelo método fenol-clorofórmio padrão (SAMBROOK J., FRITSCH E. F., 1989) e o DNA dos novilhos foi extraído de amostras de sangue utilizando-se o método *salting out*, na mesma metodologia descrita por Tizioto e colaboradores (2012). A concentração de DNA foi medida no espectrofotômetro NanoDrop®, e a qualidade analisada por meio da relação da absorbância óptica 260/280, seguido de inspeção da integridade por meio de eletroforese em gel de agarose.

A genotipagem dos touros e novilhos foi realizada em *Illumina BovineHD BeadChip* (Illumina Inc, San Diego, CA, EUA) no laboratório de genômica funcional bovina do USDA ARS em Beltsville, MD, EUA e no Centro de Genômica da ESALQ, em Piracicaba, SP, Brasil (MUDADU et al., 2016). O controle de qualidade foi realizado com o programa PLINK

(CHANG et al., 2015) a partir dos SNPs autossômicos com as coordenadas genômicas correspondendo ao genoma de referência bovino mais recente (versão: ARS-UCD1.2). Foram descartados SNPs com MAF (frequência alélica mínima) menor que 5%, valores de Hardy-Weinberg menores que 0,0001 e também amostras e SNPs com um *call rate* menor que 95%. Os filtros de qualidade resultaram em 429.513 SNPs de 224 animais (34 touros e 190 novilhos), cujos genótipos foram faseados utilizando o *software* BEAGLE (BROWNING; BROWNING, 2008), sendo organizados em arquivo no formato VCF.

6.1.3 Extração de RNA e sequenciamento

Para a extração de RNA, cerca de 100 mg do músculo *L. thoracis* de 190 animais foi coletado imediatamente após o abate. Esses tecidos foram conservados em nitrogênio líquido (Ambion, Austin, TX, United States) e congelados a -80°C até a extração. A extração de RNA foi realizada utilizando o reagente TRIzol® (Life Technologies, Carlsbad, CA, USA), de acordo com as instruções do fabricante. O equipamento Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA) foi utilizado na verificação da qualidade do RNA e o TruSeq RNA Sample Preparation kit (Illumina, San Diego, CA) foi usado na preparação das bibliotecas de RNA a partir de 2µg de RNA. O tamanho médio das bibliotecas também foi determinado pelo Bioanalyzer 2100 e foram realizadas qPCRs com o kit KAPA Library Quantification (KAPA Biosystems, Foster City, CA, USA) para quantificar essas bibliotecas. O agrupamento e o sequenciamento foram realizados em plataforma Illumina HiSeq 2500® (Illumina, San Diego, CA, USA), e o programa SeqyClean (<https://github.com/ibest/seqyclean>) foi utilizado para remover os adaptadores de sequenciamento e leituras de baixa complexidade. Pelo *software* FASTQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) foi realizado o controle de qualidade. O sequenciamento foi realizado no Centro de Genômica da ESALQ (Piracicaba, SP, Brasil). A metodologia do sequenciamento foi descrita mais detalhadamente em trabalho anterior (CESAR et al., 2018), mas foi publicada com a versão anterior do genoma (UMD_3.1). Neste trabalho o genoma de referência utilizado foi o ARS-UCD1.2.

6.1.4 Dados de ASE

SNPs heterozigotos foram identificados pela genotipagem em *Illumina BovineHD BeadChip*, e o algoritmo ALEA (YOUNESY et al., 2014) foi usado para criar dois genomas haplóides para cada animal a partir dos genótipos faseados com o genoma de referência bovino (versão: ARS-UCD1.2). Depois disso, os dois genomas foram concatenados em um genoma diploide para cada animal e foi feito o alinhamento das leituras de RNA-seq a cada alelo. Foi aplicado o teste binomial para determinar se as contagens das leituras dos alelos foram estatisticamente divergentes para cada SNP de cada animal. Este trabalho obteve 820 SNPs ASE (DE SOUZA et al. em revisão¹).

Para o início deste projeto foram cedidos pela Dra. Marcela Maria de Souza os arquivos contendo essas informações, desde os dados de ASE aos de genotipagem de todos os animais. A Tabela 1 demonstra quais são esses arquivos, e apresenta um breve resumo sobre quais informações cada arquivo fornece.

Tabela 1 - Arquivos com os dados de entrada para utilização no projeto.

Arquivo	Informações	Utilização
Resultados.txt	Apresenta todas as informações referentes à ASE. Contêm: SNP, posição, ID da amostra, <i>counts</i> por haplótipo, p-valor após correção FDR, significância e classificação da ASE, número de testes por SNP, razão ASE significativa/SNP, razão de amostras significativas/SNP, genótipo no SNP e genótipos parentais.	Seleção de animais que possuem os SNPs ASE; Seleção do mínimo de 10 heterozigotos significativos no SNP ASE; Cálculo dos desequilíbrios alélicos para a associação com aseQTL.
Chip.vcf	Arquivo VCF com os genótipos de todos os animais, em fase. Nesse arquivo existem as seguintes informações: identificação da amostra, do SNP, cromossomos, posições, alelos (referência ou alternativos), genótipos em formato DS e GT, e outros.	Arquivo utilizado para a seleção dos SNPs candidatos e animais da análise; Arquivo que origina demais VCFs filtrados com todos os SNPs candidatos, dentro das janelas especificadas.

¹ DE SOUZA, M. M. et al. Allele-specific expression is widespread in *Bos indicus* muscle and affects meat quality candidate genes. **Scientific Reports**, [s.d.].

6.2 Análises realizadas nesse trabalho

Nessa seção estão descritas todas as atividades realizadas no presente trabalho.

6.2.1 Análise de associação de variantes à expressão alelo-específica (aseQTL)

A metodologia utilizada foi semelhante à descrita por Battle e colaboradores (2014) devido à alta confiabilidade estatística e bibliográfica, pois foi citada na maior parte das publicações relacionadas ao aseQTL. O estudo buscou identificar SNPs que possuíssem efeito regulador significativo na expressão alélica, com base em sua associação ao desequilíbrio da razão alélica (DRA). Para facilitar a compreensão das siglas utilizadas nesse estudo, os SNPs em que se investiga a possibilidade de serem aseQTLs foram denominados como SNPs candidatos. Os SNPs que apresentaram ASE em músculo de Nelore previamente (DE SOUZA et al., em revisão¹) foram denominados como SNPs ASE, os genes que os contém foram denominados como genes ASE.

As manipulações do arquivo em formato VCF, o qual contém os genótipos, foram realizadas em terminal Linux, com os programas Vcftools 0.1.15 (DANECEK et al., 2011) e Bedtools 2.26.0 (<https://github.com/arq5x/bedtools>). As informações sobre os SNPs ASE e contagens dos alelos foram visualizados e filtrados em R 3.5.1, através da utilização de funções básicas e do pacote `data.table` 1.12.2 (<https://github.com/Rdatatable/data.table/>) para manipulação de tabelas. A execução dos milhares de testes foi automatizada através da elaboração de linhas de comando distribuídas em comandos de controles de fluxo (*loops*) e condições lógicas em ambas as plataformas.

O *script* foi desenvolvido para imprimir na tela “console” do RStudio todas as informações de maior importância para a execução dos testes, como o tamanho dos grupos amostrais, se os animais possuem o SNP ASE ao qual se está avaliando a associação, o número de heterozigotos para o SNP ASE e SNPs candidatos, qual teste foi realizado e a significância do teste, de acordo com o valor p.

Todas as etapas da análise passaram por condições lógicas que garantiam sua aceitação e prosseguimento, conforme o requerido pela estatística do teste. Dentre essas condições, estava o número mínimo de heterozigotos para cada teste, pois foram utilizados apenas os SNPs

¹ DE SOUZA, M. M. et al. Allele-specific expression is widespread in *Bos indicus* muscle and affects meat quality candidate genes. **Scientific Reports**, [s.d.].

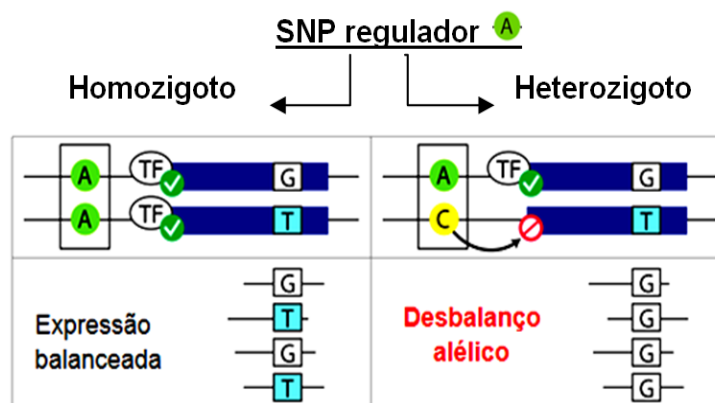
candidatos e SNPs ASE em que o grupo possuía no mínimo 10 animais heterozigotos. Outra condição lógica adotada determinava que os testes deveriam ser realizados com os grupos homozigotos mais frequentes para cada SNP candidato. Além dessas, foi estabelecida a condição de retornar uma resposta quando houvesse algo que levasse ao bloqueio do *loop*, como por exemplo, quando o *loop* não gerasse os arquivos intermediários que possibilitassem sua continuidade.

Para identificar os aseQTLs, foram realizados testes para todos os SNPs contidos em 1Mb *upstream* e *downstream* às posições de cada SNP ASE. Como mencionado anteriormente, o DRA é a variável dependente da associação, que pode ser calculado a partir da equação:

$$\left| \frac{r}{r+a} - 0,5 \right|,$$

onde r e a são o número de leituras de RNA-seq mapeadas para os alelos referência e alternativos, respectivamente. Segundo essa equação, existe ASE quando o resultado é qualquer número diferente de zero. Após o cálculo dessas variáveis, foi realizado o teste da Soma dos Ranques de Wilcoxon para comparar o DRA no SNP ASE entre grupos de animais com genótipos diferentes (homozigotos e heterozigotos) para o SNP candidato a ser um aseQTL. Se animais do grupo heterozigoto para o SNP candidato apresentaram o DRA maior que os do grupo homozigoto para o mesmo SNP, a um nível de significância de 5%, esse SNP candidato foi considerado potencial regulador da ASE e classificado como um aseQTL. Para facilitar a compreensão do teste, um esquema foi adaptado do artigo com a metodologia original na Figura 2.

Figura 2 – Esquema de um *locus* quantitativo associado à expressão alelo-específica (aseQTL).



O aseQTL é um SNP *cis*-regulador que somente afeta a ASE quando heterozigoto, quando os diferentes nucleotídeos contribuem diferencialmente com a expressão dos alelos. No exemplo, a adenina do SNP regulador (verde) permite a ligação de um fator de transcrição que promove a expressão no alelo no mesmo haplótipo que ele. Esse SNP resulta em desbalanço alélico quando heterozigoto, expressando apenas o alelo que contém o nucleotídeo A e silenciando o alelo que contém o nucleotídeo C.

A frequência alélica dos aseQTLs foi calculada com o programa Vcftools 0.1.15 (DANECEK et al., 2011). A anotação dos genes ASE foi realizada a partir dos SNPs ASE com o *software Variant Effect Predictor (VEP)* (MCLAREN et al., 2016). Para comparar a posição dos aseQTLs com os sítios de início de transcrição (TSSs, do inglês *transcription start sites*), os TSS para todos os genes ASE foram extraídos do banco de dados Ensembl pela ferramenta BiomaRt (DURINCK et al., 2005).

6.2.2 Análise de desequilíbrio de ligação

As análises de DL foram realizadas de três maneiras diferentes com aseQTLs e SNPs ASE, usando o *software Haploview 4.2* (BARRETT et al., 2005): I) Análise de DL entre aseQTLs e seus respectivos SNPs ASE associados, para entender se um grande DL pode ser relevante para a significância da associação. II) Análise de DL entre todos os aseQTLs, para investigar possíveis resultados ambíguos e indicar elementos *cis*-reguladores conservados dentro de um bloco de DL e III) Análise de DL entre todos os SNPs ASE para os quais foram identificados aseQTLs, para entender se os mesmos aseQTLs podem afetar mais de um SNP ASE em decorrência do alto DL entre os SNPs ASE.

6.2.3 Predição de mecanismos de regulação

Os aseQTLs foram anotados utilizando o programa VEP (MCLAREN et al., 2016).

A identificação de aseQTLs que mudam SLFTs foi realizada pelo uso do programa *Transcription Factor Affinity Prediction (TRAP)* para múltiplas sequências (THOMAS-CHOLLIER et al., 2011). As análises foram feitas com as sequências flanqueadoras dos aseQTLs, que abrangem regiões de 25 pb antes e depois dos SNPs. Nesta etapa foi utilizada a matriz Transfac 12.1 e como modelo de fundo o que contém os elementos conservados em cordados. Os FTs resultantes foram comparados com uma lista dos fatores expressos em músculo bovino curada manualmente (DE SOUZA et al., 2018). Nós consideramos significativas as interações com os valores p corrigidos menores que 0,05 e somente se a presença de um alelo do aseQTL (referência ou alternativo) mudasse a ligação de algum FT.

Os dados de aseQTL foram sobrepostos aos produzidos pelo metiloma de músculo (dados ainda não publicados), que compreendeu o perfil de metilação de uma amostra de 12 animais da nossa população por RRBS (Do inglês *reduced representation bisulfite sequencing*) (MEISSNER, 2005). A taxa de metilação foi calculada pela porcentagem de metilação em um

determinado sítio, dividindo o número de citosinas metiladas num sítio pela soma das citosinas metiladas e não metiladas, multiplicado por 100.

Para filtrar os dados de aseQTL para a predição de sítios de ligação de miRNA, utilizamos apenas aseQTLs que se localizam fora de regiões intergênicas, e que integraram com dados de GWAS produzidos pelo grupo de pesquisa. A predição de sítios de ligação de miRNA nas sequências flanqueadoras dos aseQTLs foi realizada com o *software* RNA *hybrid* 2.2.1 (KRUGER; REHMSMEIER, 2006). O programa RNA *hybrid* calcula a hibridização energeticamente mais favorável entre as sequências do mRNA e miRNA. Nesse trabalho nós avaliamos a interação entre as sequências flanqueadoras dos aseQTLs e as sequências dos miRNAs maduros bovinos extraídos do banco de dados MiRBase (Versão 22.1) (KOZOMARA; GRIFFITHS-JONES, 2014). Foram mantidos para as análises os pares de miRNA/aseQTLs com energia mínima livre (mfe) menor ou igual a -18,0. Esses resultados foram comparados à lista de miRNAs expressos em músculo de 185 animais da nossa população (OLIVEIRA et al., 2018) e também foram selecionados apenas os aseQTLs que criam um sítio de ligação de miRNA com apenas um dos alelos.

Depois disso, foi utilizado o *software* MirWalk 3.0 (STICHT et al., 2018) para que fosse possível identificar pares de miRNA e genes alvo conhecidos no modelo bovino e depois comparar com os genes ASE regulados por aseQTLs que também afetassem os mesmos miRNAs. Assim, o programa mostra quais genes foram previamente descritos na literatura como alvos para estes miRNAs.

6.2.4 Integração de dados

Foi realizada a integração dos aseQTLs e os genes ASE associados com regiões de QTL de fenótipos interessantes para a pecuária bovina. Nós realizamos a sobreposição das posições dos aseQTLs com as regiões de QTL do genoma bovino obtidas pela lista disponível no *Cattle QTL database* (versão do genoma de referência: UMD_3.1) (<https://www.animalgenome.org/QTLdb/cattle/> baixado em junho de 2019). Os aseQTLs também foram integrados com QTLs descritos em trabalhos realizados na mesma população pelo nosso grupo de pesquisa, identificados por GWAS para fenótipos de qualidade de carne (TIZIOTO et al., 2013), conteúdo de minerais (TIZIOTO et al., 2015a), eficiência alimentar (DE OLIVEIRA et al., 2014), conteúdo e deposição de marmoreio (CESAR et al., 2014) e também comparamos os aseQTLs com eQTLs descritos na mesma população (CESAR et al., 2018).

Na análise de integração com QTLs em janelas genômicas, foi realizado um teste de permutação com 1.000 simulações utilizando o pacote para R *RegioneR* 1.14.0 (GEL et al., 2015) para investigar a aleatoriedade das posições dos aseQTLs nessas regiões potencialmente reguladoras.

Todos os resultados de GWAS utilizados na integração de dados que foram obtidos a partir da versão UMD_3.1.1 do genoma de referência foram convertidos para a versão ARS-UCD1.2 no programa da UCSC *Lift Genome Annotations* (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>).

7. Resultados

7.1 aseQTLs

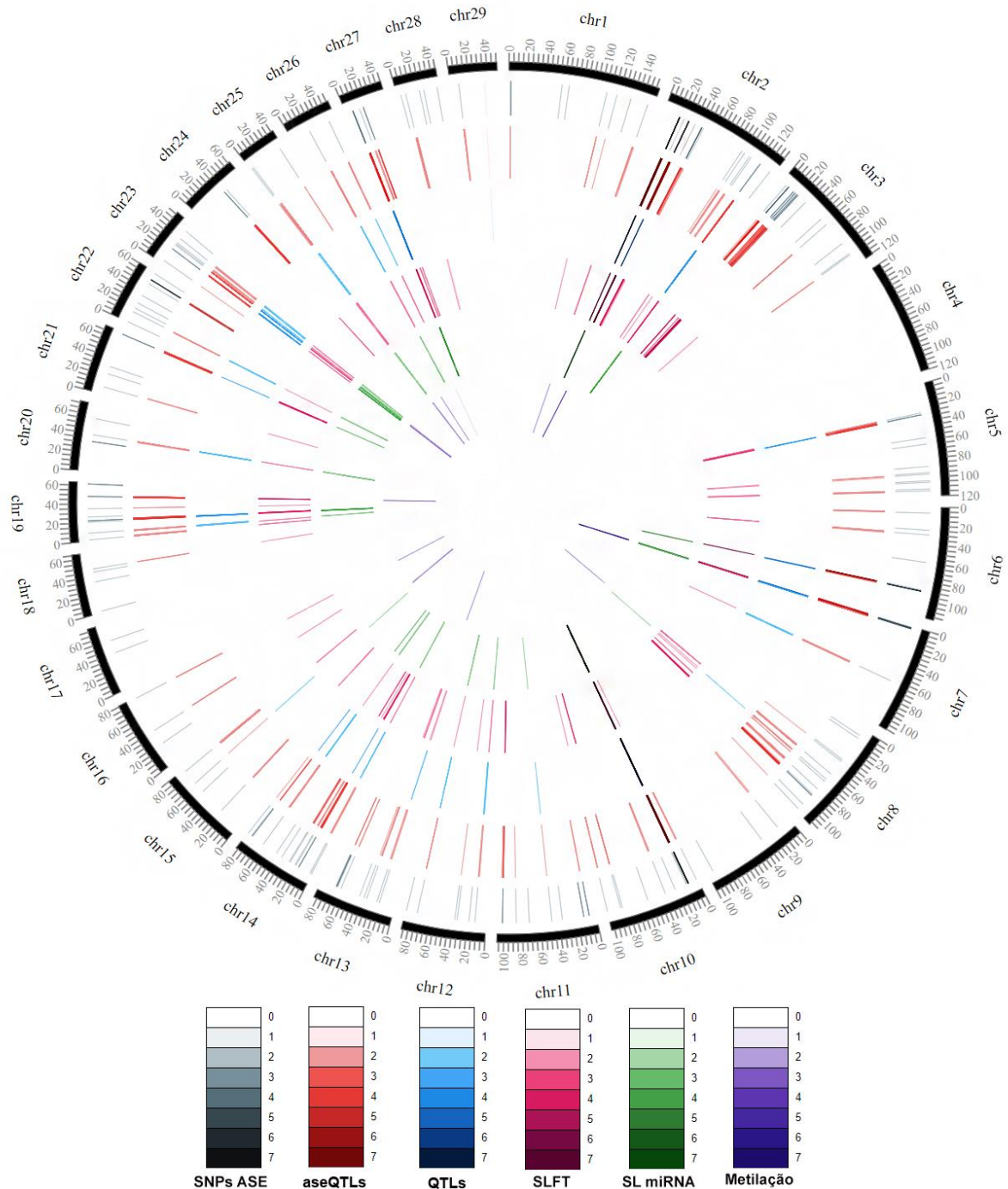
Neste trabalho foram identificados 1.134 aseQTLs (de 22.470 testes), associados com os padrões de ASE de 126 SNPs ASE, distribuídos em 85 genes. Os SNPs ASE estão demonstrados no anel cinza da Figura 3. Nós observamos que três SNPs ASE também foram identificados como aseQTLs: o SNP ASE rs136209194, que atua como um aseQTL associado ao SNP ASE rs132817153; o SNP ASE rs109842586 (dentro do gene *XIRP2*), é aseQTL do SNP ASE rs109372848; e o SNP ASE rs136097891 (que pertence ao gene *MTUS1*), é um aseQTL associado ao SNP ASE rs109005284 (*MTUS1*).

O número médio de aseQTLs por SNP ASE foi nove, variando de um a 56. O SNP ASE rs134422650 obteve associação com o maior número de aseQTLs. O segundo foi o SNP ASE rs110850310, localizado no gene *HSPA1A*, com 38 aseQTLs. A Figura 3 mostra a distribuição dos aseQTLs por cromossomo (anel vermelho) e a Figura 4 contém os resultados gerais do teste de associação. Alguns pontos tem muitos aseQTLs aparentemente porque estão associados a múltiplos SNPs ASE vizinhos, como pode-se observar por comparação ao anel cinza da Figura 3, na posição entre 15 e 20 kb no cromossomo 2 e entre as posições de 10 e 15 kb no cromossomo 10.

A frequência alélica dos aseQTLs varia homogeneamente de 0,05507 para 0,94493, com um valor médio de 0,5. Os resultados da associação entre aseQTLs e SNPs ASE estão presentes na Tabela Suplementar 1.

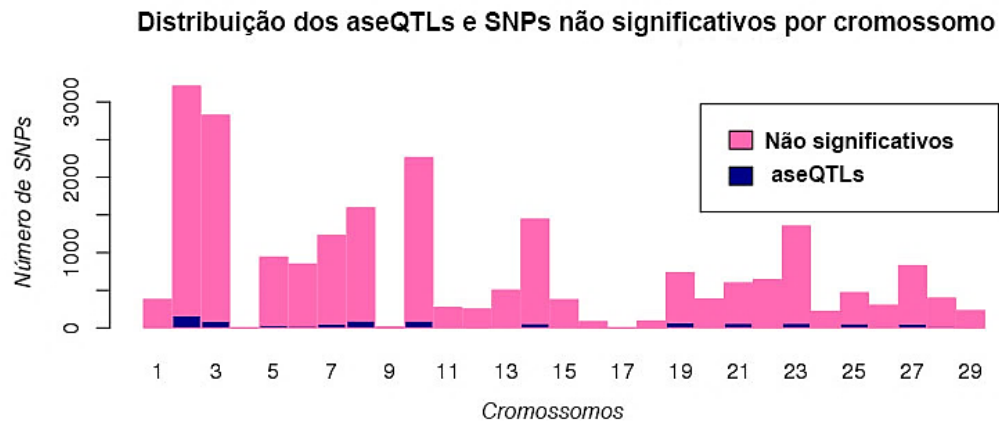
Como foi demonstrado, um SNP ASE pode ser afetado por vários aseQTLs, mas o contrário, ou seja, um aseQTLs afetando mais de um SNP ASE, foi pouco observado. Os aseQTLs rs132798564 e rs110663707 foram associados a quatro e três SNPs ASE do gene *CMYA5*, respectivamente. Outros 54 aseQTLs foram associados simultaneamente a diferentes pares de SNPs ASE.

Figura 3 - Distribuição de SNPs com expressão alelo-específica (SNPs ASE) e *loci* quantitativos associados à expressão alelo-específica (aseQTLs) de acordo com a sua posição genômica, mecanismos de regulação preditos e sobreposições com QTLs obtidos na mesma população.



SNPs ASE: SNPs ASE para os quais foram identificados aseQTLs. AseQTLs: aseQTLs identificados. QTLs: aseQTLs que sobrepõem com QTLs identificados previamente na nossa população de Nelore. SLFTs: aseQTLs preditos a modificar a ligação do fator de transcrição com apenas um alelo (ferramenta *TRAP*). SL miRNA: aseQTLs que potencialmente mudam a ligação de um miRNA a seu sítio com apenas um alelo (ferramenta *RNA hybrid*). Metilação: aseQTLs metilados na nossa população. A intensidade da cor aumenta conforme a densidade dos SNPs na escala do gráfico (janelas de 1pb). No gráfico, "chr" indica o número do cromossomo em que os aseQTLs estão localizados.

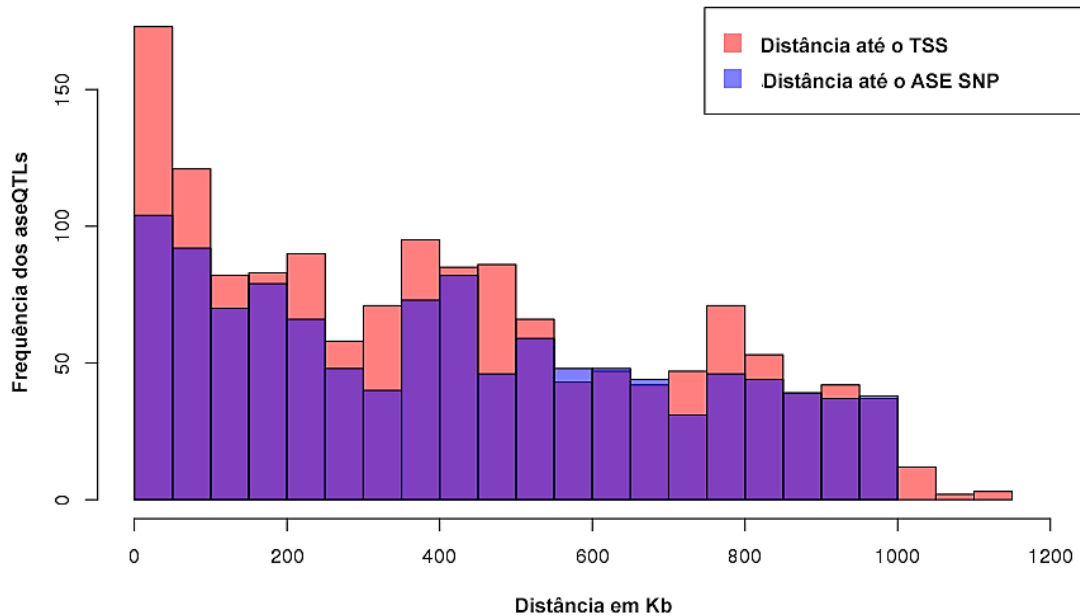
Figura 4 - Resultados gerais da análise de associação de SNPs à expressão alelo-específica (ASE).



Em rosa estão as barras para os SNPs que não foram estatisticamente significativos e em azul estão os aseQTLs.

Uma distância média de 404.983 pb foi observada entre todos os aseQTLs e os respectivos TSSs dos genes ASE afetados. As distâncias variam de 281 pb a 1,1 Mb. Os aseQTLs estão distribuídos mais frequentemente entre 0 e 100 kb de distância do TSS, dispersando conforme o aumento da distância, sendo menos comuns em intervalos superiores a 1Mb. Além disso, a distância média entre um aseQTL e o seu SNP ASE associado foi de 417.698 pb, com uma distância mínima de 2.238 pb e máxima de 999.650 pb. Similarmente à distância entre os aseQTLs e os TSSs, as maiores frequências de aseQTLs estão presentes até 100kb de distância para o SNP ASE associado. A distribuição dos aseQTLs em comparação aos SNPs ASE e TSSs dos genes ASE associados estão dispostos na Figura 5.

Figura 5 - Distribuição dos *loci* quantitativos associados à expressão alelo-específica (aseQTLs) em relação à sua distância aos sítios de início de transcrição (TSSs) dos genes com expressão alelo-específica (genes ASE, em rosa) e em relação à sua distância aos SNPs com expressão alelo-específica (SNPs ASE) associados (em azul), em Kb.

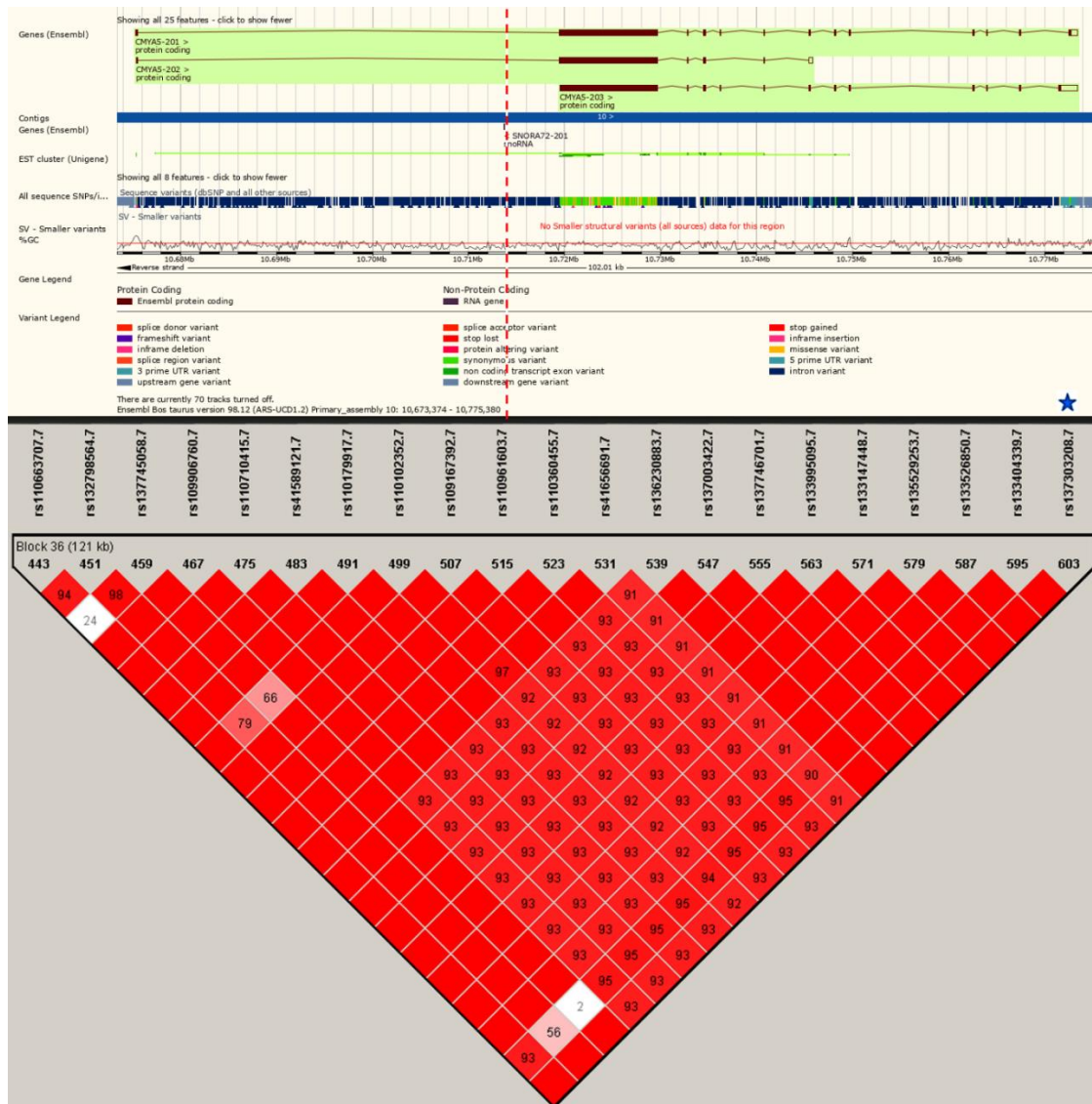


7.2 Desequilíbrio de Ligação

Foi analisado o DL entre todos os aseQTLs para investigar blocos com elementos reguladores associados. No total, 3.030 pares de aseQTLs mostraram D' maior que 0,8, e 1.813 deles estavam em total desequilíbrio ($D'=1$). Foram identificados 117 blocos de DL (Tabela Suplementar 2), incluindo de dois a 21 aseQTLs. O maior bloco de DL encontrado possui 21 aseQTLs (Figura 6) que estão afetando 5 SNPs ASE presentes no gene ASE *CMYA5*.

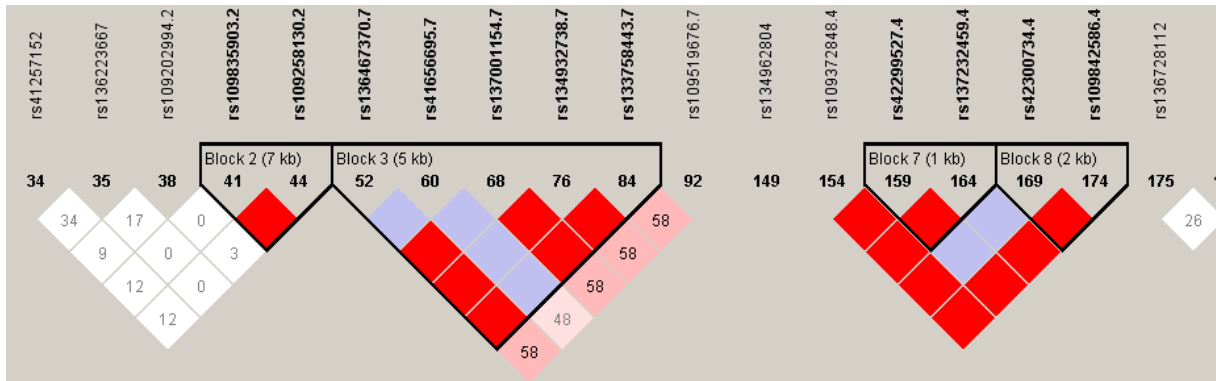
Depois, nós realizamos a mesma análise com todos os SNPs ASE, para observar o impacto do DL em efeitos redundantes de aseQTLs sobre estes SNPs ASE. Um total de 32 pares de SNPs ASE apresentaram $D' > 0,8$, sendo 31 com $D'=1$, distribuídos em 11 blocos de DL (Figura 7, Tabela Suplementar 3).

Figura 6 - Maior bloco de desequilíbrio de ligação (DL) entre os *loci* quantitativos associados à expressão alelo-específica (aseQTLs).



Os quadrados com números indicam a intensidade do DL de 1 a 99, e quando sem números indicam que os pares de SNPs estão em total DL. Os tons de vermelho se intensificam conforme o aumento do DL. Todos os SNPs desse bloco mudam sítios de ligação para diferentes miRNAs, sendo que o SNP marcado com uma estrela azul, rs137303208, também foi predito estar presente em sítios de ligação para os FTs DR1 e SP3 com o alelo referência. A linha tracejada mostra a posição genômica do SNP marcado. Corte da imagem e bloco de DL gerados pelo Ensembl Viewer e Haploview, respectivamente.

Figura 7 - Principais blocos de desequilíbrio de ligação (DL) entre os SNPs com expressão alelo-específica (SNPs ASE).



Os quadrados com números indicam a intensidade do DL de 1 a 99, e quando sem números indicam que os pares de SNPs estão em total DL. Os tons de vermelho se intensificam conforme o aumento do DL. Quadrados em azul também representam $D'=1$, mas com a estimativa LOD (logaritmo de probabilidades) < 2 . Cortes da imagem gerada pelo Haploview.

Por fim, nós avaliamos o DL entre SNPs ASE e seus aseQTLs associados. De 1.134 pares de SNP para a análise, 238 mostraram D' maior que 0.8, com 71 pares aseQTL/SNP ASE encontrados em total desequilíbrio ($D'=1$) (Tabela Suplementar 4). Esses pares de SNPs correspondem a 35 ASE genes, sendo o gene *CMYA5* aquele que apresentou maior número de interações entre aseQTLs e SNPs ASE (38), seguido do gene *EIF5*, com 33, e *VCP*, com 27.

7.3 Integração de dados

Após a conversão dos QTLs do *Cattle QTL Database* para a versão mais atualizada do genoma de referência, nós identificamos uma sobreposição de 173 aseQTLs dentro de regiões de QTL de 21 características, distribuídas em 215 ocorrências (Tabela Suplementar 5). O teste de permutação foi significativo (valor $p = 9.999e-05$). As três características com mais aseQTLs sobrepostos foram peso corporal (com 81 aseQTLs), teor de ácido palmítico do leite (com 42 aseQTLs) e marmoreio (com 28 aseQTLs). O aseQTL rs136717535 está localizado em mais regiões de QTL: porcentagem de kappa-caseína do leite, porcentagem de proteína do leite, força de cisalhamento, tamanho da panturrilha e índice de nascimento.

Nós também integramos os dados de ASE com outros estudos realizados em uma amostra mais ampla oriunda da mesma população amostral. Todos os testes de permutação realizados para QTLs identificados em janelas genômicas foram significativos (valor $p \leq 0.05$). Nós identificamos 234 aseQTLs distribuídos em 894 sobreposições com 25 características (Tabela

suplementar 6). Dentre esses, 847 sobreposições ocorreram com características de qualidade de carne (TIZIOTO et al., 2013) (Tabela 2).

Tabela 2 - *Loci* quantitativos associados à expressão alelo-específica (aseQTLs) integrando com QTLs de características de qualidade de carne identificados em músculo de Nelore.

Característica	Número de aseQTLs	Genes ASE afetados
Área de olho de lombo (AOL)	192	<i>ACOT13, ASB5, CA2, CAB39, CMYA5, CUEDC1, DNAJC21, ITGB1, KLF10, MBNL2, NXN, PTP4A3, RGCC, SPARC, STBD1, XIRP1, XIRP2</i>
Luminosidade da Gordura (L*gordura)	127	<i>ACOT13, ASB5, CAB39, CMYA5, DNAJC21, ITGB1, KLF10, NXN, SPARC, STBD1,</i>
Força de cisalhamento 7 dias após o abate (FC7)	122	<i>ACOT13, ASB5, CAB39, CMYA5, DNAJC21, ITGB1, KLF10, SPARC, STBD1</i>
Capacidade de retenção de água (CRA)	113	<i>ACOT13, ASB5, CAB39, CMYA5, DNAJC21, ITGB1, KLF10, SPARC, STBD1</i>
Perdas por cozimento (PC)	105	<i>ACOT13, ASB5, CMYA5, DNAJC21, ITGB1, SPARC, STBD1</i>
Intensidade da coloração amarelada da gordura (b*)	100	<i>ACOT13, CAB39, CMYA5, DNAJC21, ITGB1, KLF10, SPARC, STBD1</i>
Força de cisalhamento 24 horas após o abate (FC24)	46	<i>CMYA5, PFKM</i>
Luminosidade do músculo (L*músculo)	23	<i>ACOT13, SPARC,</i>
Espessura de gordura subcutânea (EGS)	17	<i>ACOT13, STBD1</i>
Força de cisalhamento 14 dias após o abate (FC14)	2	<i>SPARC</i>

O SNP ASE rs109550233, uma variante exônica transcrita não-codificante, teve mais aseQTLs associados que se sobrepuseram a fenótipos de qualidade de carne, com nove aseQTLs dentro de regiões de nove características: b*, EGS, PC, L*gordura, L*músculo, AOL, FC, FC7 e CRA. O segundo SNP ASE com mais aseQTLs em QTLs foi o rs135906938 (do gene *ACOT13*), com oito aseQTLs em oito regiões de QTL para fenótipos de qualidade de carne (todas as características integradas aos aseQTLs do SNP ASE rs109550233, exceto FC). Além disso, 35 aseQTLs associados a seis SNPs ASE do gene *CMYA5* se sobrepuseram a QTLs de sete características.

Vinte aseQTLs integraram com regiões de *cis*-eQTL e quatro com regiões de *trans*-eQTL (CESAR et al., 2018); 10 aseQTLs estão em regiões de QTL para conteúdo dos minerais Co, Mn, Zn, Ca, Cr, Ar, K, Mg e S (TIZIOTO et al., 2015a); 7 se sobrepuseram com QTLs para

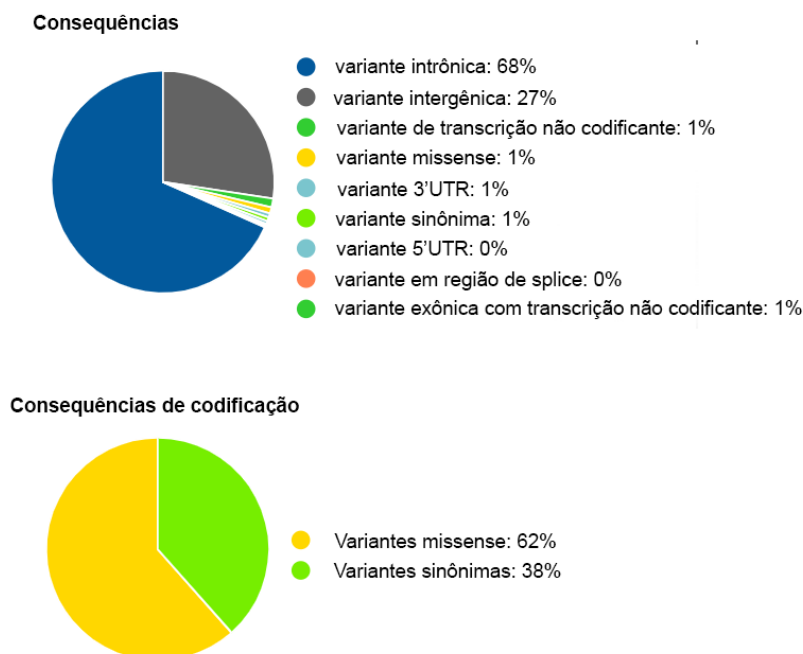
composição de marmoreio, todos associados ao ácido octadecenóico (CESAR et al., 2014); seis aseQTLs estão em regiões de QTLs para fenótipos de eficiência alimentar: eficiência de ganho (EG), eficiência de manutenção (ME) e eficiência parcial do crescimento (PEG) (DE OLIVEIRA et al., 2014). O anel azul da Figura 3 apresenta a distribuição genômica dos aseQTLs que integraram com esses trabalhos.

7.4 Mecanismos de regulação

7.4.1 Anotação dos SNPs

De acordo com a localização genômica dos aseQTLs, nós realizamos a anotação funcional com o VEP (MCLAREN et al., 2016), para prever as possíveis consequências destes SNPs de acordo com sua posição (Figura 8, Tabela Suplementar 7). Os aseQTLs estão predominantemente distribuídos em regiões intrônicas (68%), 27% estão dentro de regiões intergênicas e 1% em regiões de transcritos não codificantes, 1% em regiões 3'UTR, 1% são variantes *missense* e 1% são variantes sinônimas. Com o programa VEP também foi possível prever que 98,47% dos aseQTLs têm um impacto modificante sobre os genes afetados por eles, 0,72% têm impacto moderado e 0,81% tem pouco impacto.

Figura 8 - Anotação funcional dos *loci* quantitativos associados à expressão alelo-específica (aseQTLs).



Adaptado a partir da saída do VEP.

7.4.2 aseQTLs em SLFTs

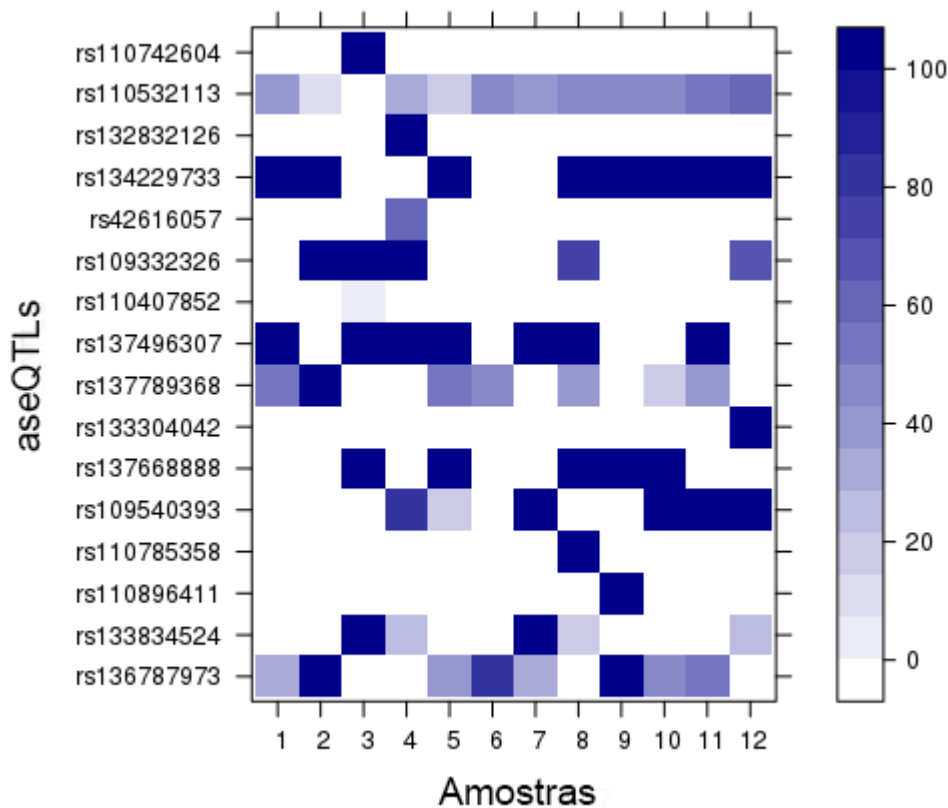
Usando as sequências flanqueadoras de todos os aseQTLs como entrada para o programa TRAP (THOMAS-CHOLLIER et al., 2011), nós encontramos 83 SLFTs presentes em nossas sequências (FDR < 0.05) (anel rosa da Figura 3, Tabela Suplementar 9). Destes, apenas 12 fatores de transcrição foram expressos em músculo bovino (DE SOUZA et al., 2018). Considerando apenas os aseQTLs que mudam estes sítios, foram encontrados 215 aseQTLs, sendo que 150 potencialmente afetam SLFTs com o alelo referência e 117 com o alelo alternativo. O fator de transcrição SMAD4 pode ter sua ligação afetada por 37 aseQTLs; o sítio de ligação para o fator de transcrição AHR pode ser afetado por 31 aseQTLs; em sítios de ligação para os fatores de transcrição VDR e SP3 foram encontrados 26 aseQTLs; o sítio de ligação do *STAT3* pode ser afetado por 25 aseQTLs; sítios de ligação para o fator de transcrição PPARA foram encontrados em 24 aseQTLs; DR1, TAL1, ZNF219 potencialmente possuem sítios com 20 aseQTLs; os fatores de transcrição STAT6, SP1 e MZF1 foram preditos a se ligar a 16, 12 e 10 aseQTLs, respectivamente. O aseQTL rs137692633 é o que pode modificar o maior número de SLFTs, afetando a ligação de 6 fatores de transcrição: SP1, SP3, STAT3, STAT6, VDR e ZNF219.

Os dois principais genes associados aos aseQTLs que potencialmente modificam SLFTs, com o maior número de aseQTLs dentro de SLFTs foram o gene *EIF5*, com 18 aseQTLs modificando sítios de ligação para 9 fatores de transcrição e o *CMYA5*, ao qual estão associados 14 aseQTLs que potencialmente modificam a ligação de sete fatores de transcrição.

7.4.3 aseQTLs com metilação

Para analisar o padrão de metilação nos nossos aseQTLs, foram comparadas as posições destes aseQTLs com dados de metilação por RRBS produzidos na mesma população (dados não publicados). Nós identificamos 69 citosinas metiladas em 16 aseQTLs distribuídos em 12 animais (Figura 9, Tabela Suplementar 8). O aseQTL rs110532113, associado com o SNP ASE rs133877387 do gene *MSRB1*, foi metilado em 11 dos 12 animais para os quais a metilação foi testada. Trinta e sete citosinas tiveram 100% de metilação considerando todas as amostras, compreendendo 13 aseQTLs. Oito animais tiveram 100% de metilação no aseQTL rs134229733 e sete no aseQTL rs137496307, associado com um SNP ASE do gene *VIM*.

Figura 9 - Porcentagem de metilação dos *loci* quantitativos associados à expressão alelo-específica (aseQTLs) em 12 animais Nelore.



A intensidade da cor aumenta de acordo com a porcentagem de metilação de cada SNP por amostra.

Seis aseQTLs metilados estão associados a SNPs ASE dentro de sete genes ASE: *HSPA1A*, *MSRB1*, *VCP*, *VIM*, *SCN4A*, *PECAMI* e *ZDHHC4*, sendo o aseQTL rs137668888 associado com SNPs ASE dos genes *SCN4A* e *PECAMI*. O aseQTL rs110742604 é localizado a 761.297 pb do TSS do gene *HSPA1A* e rs110407852 foi o aseQTL metilado mais próximo do TSS do respectivo gene associado, *VCP*, distante 51.026 pb dele. A distância média dos aseQTLs metilados aos TSSs dos genes ASE associados foi de 216,169 pb. Os outros dez aseQTLs metilados foram associados a 9 SNPs ASE. Os aseQTLs rs42616057 e rs109332326 foram ambos metilados e associados com o mesmo SNP ASE: rs110095343. A distribuição dos aseQTLs metilados está presente no anel roxo da Figura 3.

7.4.4 aseQTLs em sítios de ligação de miRNAs

Foram selecionados 163 aseQTLs para essa análise com o critério de que eles estivessem fora de regiões intergênicas e dentro de regiões de QTL identificadas na nossa população Nelore. Foram identificados 153 sítios de ligação de miRNA em regiões de 163 aseQTLs,

distribuídos em 1.448 pares de miRNA e sequências alvo (aseQTLs), sendo 693 com o alelo alternativo dos aseQTLs e 755 com o alelo referência do aseQTL intersectando com o sítio de ligação de miRNA (anel verde da Figura 3, Tabela Suplementar 10).

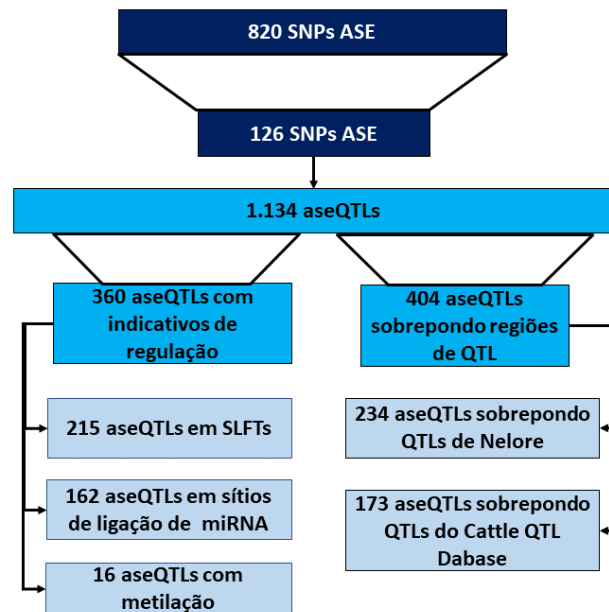
Vinte ASE genes com aseQTLs que potencialmente mudam a afinidade da ligação de miRNAs aos respectivos sítios foram utilizados como entrada para a análise realizada com o programa MiRWalk (STICHT et al., 2018). Os resultados mostraram que esse genes são alvos para 583 miRNAs, sendo que 31 miRNAs tiveram seus sítios afetados por aseQTLs preditos como reguladores de 14 genes ASE que, por sua vez, também são alvos conhecidos na literatura para esses mesmos miRNAs.

Nos próximos exemplos, os aseQTLs modificam sítios de ligação de miRNAs conhecidos por afetarem o ASE gene associado: O aseQTL rs42437277, associado com um SNP ASE no gene *ASB5*, foi predito como pertencente a sítios de cinco miRNAs. O aseQTL rs133661649, associado com um SNP ASE do gene *LPL* e o aseQTL rs109408013 associado a um SNP ASE no gene *NXN*, foram ambos alvos de três miRNAs. Sete aseQTLs afetam dois sítios de ligação de miRNA e potencialemtne regulam os respectivos genes: rs41255633 (*NXN*), rs135834852 (*NXN*), rs136463427 (*CUEDC1*), rs110417686 (*CUEDC1*), rs132705216 (*MBNL2*), rs137489440 (*MBNL2*), rs133710221 (*CAB39*). Os resultados do programa MiRWalk estão na Tabela Suplementar 11.

7.4.5 Integração dos resultados

Em resumo, nós exploramos três mecanismos de regulação atuando sobre os nossos aseQTLs: metilação, SLFTs e sítios de ligação de miRNAs. Apenas um aseQTL, rs109540393, associado ao SNP ASE rs109550233, foi metilado em seis animais da nossa população e o alelo referência possibilita a ligação do fator de transcrição AHR e de seis miRNAs expressos em músculo, sendo 5 afetados pelo aseQTL com o alelo referência e um com o alelo alternativo. Trinta e três aseQTLs estão potencialmente sendo afetados por dois mecanismos de regulação, 354 aseQTLs apresentaram apenas um mecanismo de regulação e 746 não apresentaram nenhum. A Tabela Suplementar 12 contém um resumo dos aseQTLs que foram preditos a possuírem esses mecanismos de regulação ou que sobrepueram-se a QTLs descritos anteriormente. Um resumo dos resultados, desde a obtenção dos aseQTLs está na Figura 10.

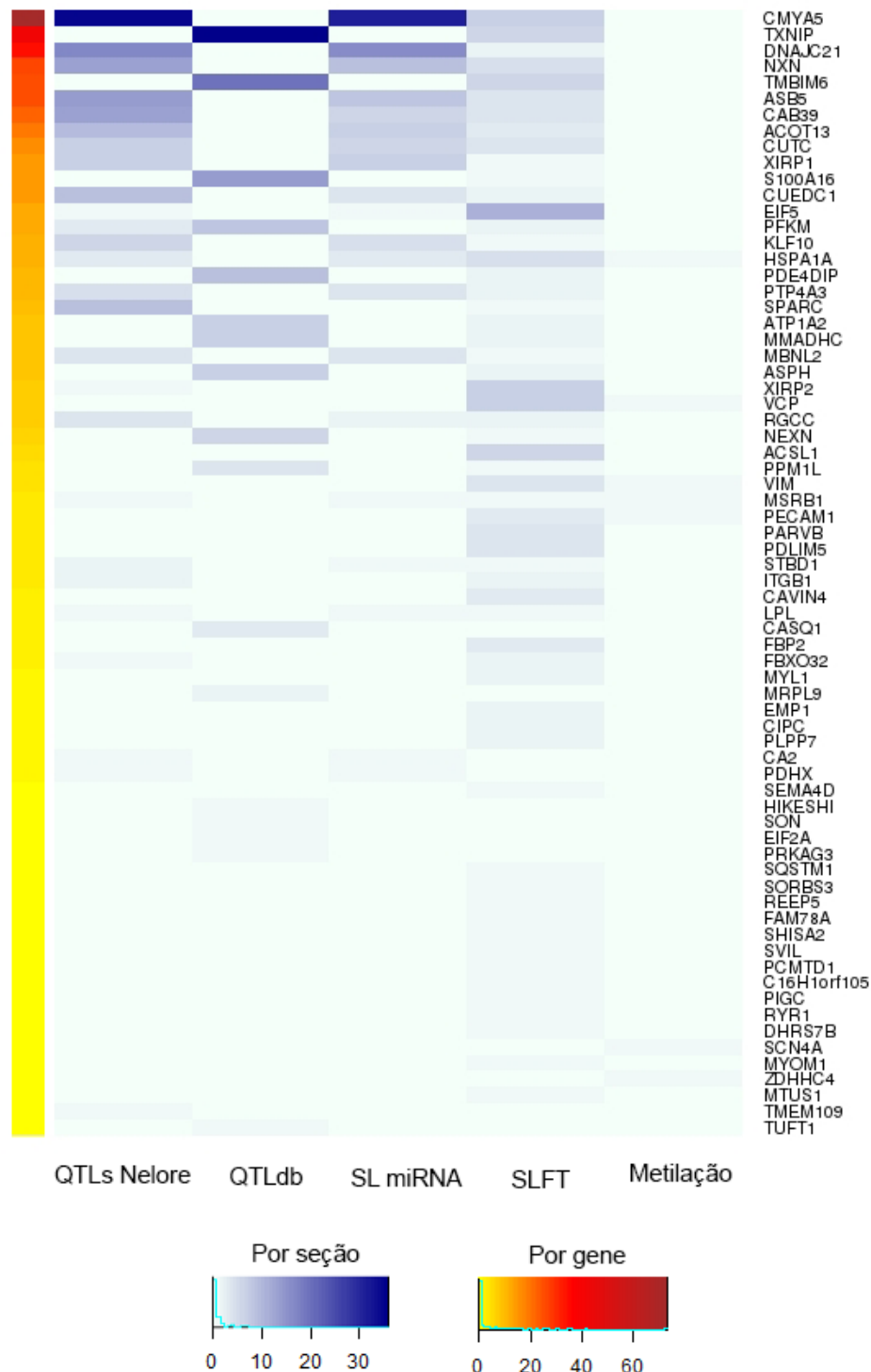
Figura 10 - Síntese esquemática dos resultados desse trabalho



A partir de 820 SNPs ASE, foram identificados 1.134 aseQTLs associados a 126 SNPs ASE; Desses aseQTLs, 360 foram preditos a possuir algum mecanismo de regulação, sendo 215 aseQTLs potencialmente capazes de alterar a afinidade de FTs expressos em músculo bovino pelos seus sítios, 162 aseQTLs podem afetar sítios de ligação de miRNAs expressos em músculo dos animais da mesma população e 16 aseQTLs apresentaram metilação. Além disso, 404 aseQTLs se sobrepuseram a regiões de QTL, sendo 234 aseQTLs identificados em dados GWAS da mesma população Nelore e 173 aseQTLs foram sobrepostos aos QTLs publicados no *Cattle QTL Database*.

Os 10 ASE genes com o maior número de aseQTLs dentro de regiões relevantes associadas a características ou mecanismos de regulação foram *CMYA5*, *TXNIP*, *DNAJC21*, *NXN*, *TMBIM6*, *ASB5*, *CAB39*, *ACOT13*, *CUTC* e *CUEDC1*, variando de 73 a 15 aseQTLs cada, respectivamente. Considerando os principais genes, de acordo com cada seção de resultados representados na Figura 11 como colunas, temos o gene *CMYA5*, com 35 aseQTLs sobrepondo a sete fenótipos de qualidade de carne de QTLs identificados na nossa população Nelore; *TXNIP*, que apresentou 36 aseQTLs localizados em regiões de QTL obtidas a partir do banco de dados bovino; *CMYA5* também possui 31 aseQTLs que potencialmente modificam a ligação de miRNAs e 11 aseQTLs do gene *EIF5* podem afetar SLFTs. Adicionalmente, sete genes tiveram um aseQTL metilado cada.

Figura 11 – *Heatmap* da distribuição de *loci* quantitativos associados à expressão alelo-específica (aseQTLs) de acordo com os genes com expressão alelo-específica (genes ASE) associados e as regiões reguladoras previstas nesse trabalho.



Os genes ASE para os quais foi encontrado ao menos um aseQTL estão em cada linha. As colunas correspondem aos aseQTLs que foram previamente identificados como QTLs na nossa população experimental, aseQTLs que foram QTLs no *Cattle QTL Database*, aseQTLs que mudam sítios de ligação para miRNAs, aseQTLs que mudam SLFTs e aseQTLs metilados, respectivamente da esquerda para a direita. A intensidade dos tons de azul é proporcional à presença de aseQTLs associados a cada ASE gene dentro de cada seção de regiões reguladoras e as cores quentes do *heatmap* à esquerda escurecem proporcionalmente ao total de aseQTLs dentro de todas as seções de resultados, para cada gene.

8 Discussão

8.1 aseQTLs

Neste estudo, foram identificados 1.134 aseQTLs associados com 126 SNPs ASE no músculo *L. thoracis* de novilhos Nelore. A exigência de 10 genótipos heterozigotos entre os pares de SNPs a serem associados possibilitou que apenas fossem testados 192 SNPs ASE, correspondentes a 23% de todos os 820 SNPs ASE previamente descritos na mesma população (DE SOUZA et al., em revisão¹). Após a análise, foram identificados aseQTLs para 126 SNPs ASE, que correspondem a 15% do número total. Embora essa exigência faça com que o teste seja mais restritivo, com a adoção de um número maior de animais utilizados nessa análise, aumentaria também a probabilidade de existirem mais genótipos heterozigotos para os SNPs de teste, possibilitando assim a realização de testes para um número maior de SNPs ASE e, possivelmente, a identificação dos respectivos aseQTLs. Trinta e sete SNPs ASE para os quais encontramos aseQTLs também foram descritos por Guillocheau e colaboradores (2019) em músculo bovino e dois SNPs ASE provenientes deste mesmo estudo também foram identificados como aseQTLs: rs109170993, associado com o SNP ASE rs110850310 e rs109842586, associado com o SNP ASE rs109372848.

8.2 Integração com dados de GWAS

Regiões de QTL obtidas do banco de dados bovino foram integradas com 16% dos nossos dados de aseQTLs, mostrando um enriquecimento para características de leite e produção, apesar de que é necessário considerar que possa existir um viés na frequência de publicações para cada característica nesse banco de dados. O aseQTL rs136717535 está presente em mais regiões de QTL do *Cattle QTL Database*, sendo associado ao SNP ASE rs41257152. Esse SNP ASE está localizado no gene *HIKESHI*, que realiza a mediação do transporte nuclear das proteínas Hsp70s, protegendo as células do dano por estresse térmico (IMAMOTO, 2018; KOSE; FURUTA; IMAMOTO, 2012). Um alelo do gene *Hsp70* foi associado com termotolerância em gado (BHAT et al., 2016) e foi classificado como um marcador importante para a quantificação de estresse térmico (ARCHANA et al., 2017). Como a raça Nelore é amplamente visada devido à sua tolerância ao calor (PASSOS et al., 2014; RIBEIRO et al., 2009), a ASE pode afetar a adaptação característica desses animais. A integração entre os

¹ DE SOUZA, M. M. et al. Allele-specific expression is widespread in *Bos indicus* muscle and affects meat quality candidate genes. **Scientific Reports**, [s.d.].

aseQTLs e QTLs descritos previamente indica que os aseQTLs podem estar em regiões que afetam características de interesse agropecuário em outras populações e raças bovinas.

A integração com a nossa população Nelore revelou um enriquecimento para características de qualidade de carne, com 94,7% dos aseQTLs integrados se sobrepondo a dez fenótipos dessa categoria (TIZIOTO et al., 2013). Apesar de considerarmos possível a existência de um viés no tecido de estudo (músculo), que pode privilegiar a identificação de QTLs associados a fenótipos de qualidade de carne, o enriquecimento funcional dos genes ASE em relação a fenótipos de qualidade de carne também foi encontrado previamente em músculo bovino de populações Limousin e Nelore (GUILLOCHEAU et al., 2019, DE SOUZA, em revisão¹).

O gene *ACOT13* foi diferencialmente expresso entre amostras contrastantes para o conteúdo de ácido linoleico (CESAR et al., 2016) e foi associado com metabolismo lipídico hepático e muscular (BAKSHI et al., 2018; KANG et al., 2012). *ACOT13* foi associado a aseQTLs que integraram com oito fenótipos de qualidade de carne, como EGS, FC7, L*gordura e b*.

O gene *SPARC*, com aseQTLs associados sobrepondo com L*gordura e b*, FC7 e FC14, foi previamente identificado na nossa população como diferencialmente expresso entre grupos extremos para conteúdo de marmoreio (CESAR et al., 2015), classificado como um alvo potencial para marmoreio e maciez em suínos (DAMON et al., 2012, 2013). Sete regiões de QTLs sobrepuseram com aseQTLs associados com um SNP ASE do gene *STBD1*, que também foi diferencialmente expresso entre amostras contrastantes para marmoreio (CESAR et al., 2015). O gene *ITGB1* foi apontado como alvo para o miRNA bta-miR-423, diferencialmente expresso entre amostras extremas pra marmoreio em nossa população (OLIVEIRA et al., 2018); aseQTLs desse gene foram também integrados com QTLs de PC e de FC. O gene ASE *CMYA5* teve 245 sobreposições entre alguns dos seus aseQTLs e as características b*, PC, L*gordura, AOL, FC, FC7 e CRA.

8.3 Influência da distância genômica e do desequilíbrio de ligação

A maioria dos aseQTLs que se associaram a múltiplos SNPs ASE apresentaram total DL entre si ($D'=1$), indicando que aseQTLs regulam o mesmo haplótipo transcrito. 15 aseQTLs foram associados com dois SNPs ASE subsequentes dentro do gene *EIF5*, com total

¹ DE SOUZA, M. M. et al. Allele-specific expression is widespread in *Bos indicus* muscle and affects meat quality candidate genes. **Scientific Reports**, [s.d.].

desequilíbrio. Esse gene foi previamente associado com marmoreio em gado Nelore (CARVALHO et al., 2019). Os aseQTLs rs110966557 e rs110621075, foram associados com dois SNPs ASE próximos dentro do gene *XIRP2*, com $D'=1$; esse gene foi relacionado com a textura e capacidade de retenção de água da carne suína (PIÓRKOWSKA et al., 2017), assim como CAR no modelo suíno (DO et al., 2014) e bovino (SEABURY et al., 2017). O aseQTL rs132798564 foi associado com quatro SNPs ASE do gene *CMYA5*, e também apresentou total desequilíbrio de ligação com eles. O aseQTL rs110663707 foi associado simultaneamente com três SNPs ASE. Quinze aseQTLs foram associados com dois diferentes pares de SNPs ASE do gene *CMYA5*, todos os pares de SNPs ASE apresentaram $D'=1$.

O gene *CMYA5* foi associado a marmoreio no modelo suíno (XU et al., 2011), teve variantes associadas com AOL em gado Nelore (CARVALHO et al., 2019) e seis SNPs associados com características de qualidade de carne na mesma população (TIZIOTO et al., 2013). Esse gene codifica para uma proteína associada à desmina (KOULOUMENTA; MAVROIDIS; CAPETANAKI, 2007), cuja interação pode regular o tráfego vesicular e lisossomal (TSOUPRI; CAPETANAKI, 2013). A via da autofagia por lisossomos é um sistema de degradação envolvido na perda de massa muscular e atrofia (SANDRI, 2013). Além disso, a proteína *CMYA5* é um substrato proteolítico para a *CAPN3*, que pode proteger a *CAPN3* da autólise (SARPARANTA et al., 2010). Em gado zebuíno, foi encontrada uma associação entre variantes do gene *CAPN3* e a maciez da carne (BARENDSE et al., 2008). A maioria dos aseQTLs no cromossomo 10 está associada ao gene ASE *CMYA5*, pertencendo a um forte bloco de DL que pode estar afetando a expressão alélica desse gene. Além disso, a presença de aseQTLs sobrepondo-se a QTLs para fenótipos de qualidade de carne também colabora com os indicativos de que a regulação do gene *CMYA5* pode fazer parte de uma rede de proteínas interagentes relacionadas ao metabolismo e estrutura muscular, podendo afetar características de interesse relacionadas à saúde do animal e maciez da carne.

Considerando todos os aseQTLs, os maiores valores de DL foram encontrados entre SNPs próximos, como foi descrito anteriormente em populações Nelore (ESPIGOLAN et al., 2013; MOKRY et al., 2014). O DL é um fator essencial quando afirma-se que um SNP está efetivamente associado com o SNP ASE, e não por uma interação indireta ou redundante com outro SNP em DL com ele (BUSH; MOORE, 2012; HIRSCHHORN; DALY, 2005). Assim, esses resultados indicam que a análise de DL é indispensável para a interpretação de testes de associação à ASE na busca por mutações causais.

8.4 Atividade cis-reguladora dos aseQTLs

Os resultados de DL, bem como a presença de aseQTLs próximos aos SNPs ASE associados e seus respectivos TSSs, são indicativos da ação em *cis* dos aseQTLs. Alguns estudos mostraram que eQTLs cujos mecanismos atuam nas proximidades das regiões promotoras dos genes podem estar envolvidos com a ASE (PICKRELL et al., 2010), e *cis*-eQTLs e aseQTLs foram identificados concentrados próximos ao TSS dos genes afetados (BATTLE et al., 2014; JOSEPHS et al., 2015; LARSON et al., 2015). AseQTLs estão predominantemente distribuídos em regiões intergênicas e intrônicas, como nós esperávamos, pois SNPs nessas regiões tendem a possuir ação reguladora (BARRETT; FLETCHER; WILTON, 2012; NELSON; HERSH; CARROLL, 2004). Essa informação está de acordo com outro estudo realizado na nossa população com *cis*-eQTLs (CESAR et al., 2018), que se sobrepuseram a 20 aseQTLs identificados aqui. Battle e colaboradores (2014) também identificaram 641 *cis*-eQTLs associados com a ASE em humanos; numa única população da planta *Capsella grandiflora*, 83% dos aseQTLs mostraram resultados consistentes com a regulação em *cis* (JOSEPHS et al., 2015). Como nós comparamos grupos de diferentes genótipos e a significância do teste depende de grandes valores de desequilíbrio alélico no grupo heterozigoto, nossos resultados corroboram com a literatura, e nós assumimos que os aseQTLs são majoritariamente afetados por elementos reguladores com ação em *cis*.

8.5 SLFT podem ser afetados por aseQTLs

Variantes *cis*-reguladoras podem modificar SLFTs (ALMLÖF et al., 2012), sendo a presença alelo-específica de SLFTs a causa para a perda de ligação de FTs (MCDANIELL et al., 2010), o que pode levar à ASE em animais heterozigotos. Nós encontramos 215 aseQTLs dentro de 12 sítios de ligação para fatores de transcrição expressos em músculo bovino (DE SOUZA et al., 2018). Cinco SLFTs preditos podem afetar mais de um aseQTL: SMAD4, AHR, VDR, STAT3 e PPARA. A deficiência de SMAD4 induzida por *knockout* murino causa atrofia muscular (SARTORI et al., 2013). SMAD4 regulado por TGF- β mostrou uma interação com o FT PPARA ativado por ligante (KINTSCHER et al., 2002), sendo este FT um receptor alfa de PPAR, relacionado com o metabolismo de lipídios (BURRI; THORESEN; BERGE, 2010). A isoforma gama do PPAR mostrou uma grande expressão e baixa metilação em regiões de marmoreio do músculo *L.thoracis* do que na porção muscular (BAIK et al., 2014). O aseQTL mais metilado deste estudo, rs134229733, foi predito a beneficiar a ligação do FT PPARA no

seu respectivo sítio com o alelo referência. PPARs e *VDR* foram também relacionados ao câncer (MATSUDA; KITAGISHI, 2013). *VDR* foi o principal FT numa rede gênica em resultados do nosso grupo (MUDADU et al., 2016), relacionado com o crescimento do gado e também apontado como regulador do desenvolvimento muscular (ENDO et al., 2003). Esse FT é um receptor de vitamina D, que foi sugerido ter participação na regulação do marmoreio junto à homeostase de cálcio (SMITH et al., 2009).

A sinalização TGF- β também foi relacionada com o FT AHR (GUO et al., 2004), associado ao envelhecimento humano (ECKERS et al., 2016). *STAT3* foi relacionado com a perda de massa muscular (BONETTO et al., 2011) e a ativação da via de sinalização do *STAT3* induz regeneração muscular (KAMI; SENBA, 2002; ZHANG et al., 2013). O alelo alternativo do aseQTL rs135348406 foi identificado dentro de um sítio de ligação para *STAT3* e foi associado com o desequilíbrio alélico de um SNP ASE do gene *HSPA1A*, que codifica para a proteína do choque térmico Hsp70. STATs podem modular proteínas do choque térmico, incluindo a Hsp70 (STEPHANOU; LATCHMAN, 1999).

Segundo os nossos resultados, o alelo referência do aseQTL rs137692633 pode beneficiar a ligação dos FTs STATs encontrados, *STAT3* e *STAT6*, e de outros quatro FTs (*SP1*, *SP3*, *VDR* e *ZNF219*). A sobreposição deste aseQTL com regiões de QTL para AOL e FC7 (TIZIOTO et al., 2013) contribuem com os indicativos apresentados sobre a relação do FT *STAT3* com o tecido muscular. *SP1* e *SP3* regulam o promotor da utrofina, sendo candidatos para os tratamentos da distrofia muscular de Duchenne (GALVAGNI; CAPO; OLIVIERO, 2001).

Quanto à posição relativa do aseQTL ao TSS do gene ASE associado, dentre todos os aseQTLs que alteram SLFTs, o rs137303208 está localizado mais próximo do TSS do gene ASE associado (*CMYA5*), distante 5.358 pb. Ele foi associado a dois SNPs ASE do gene *CMYA5* (rs137001154 e rs136467370), e seu alelo referência cria um sítio com afinidade para os FTs *DR1* e *SP3*. Esse aseQTL é o tagSNP do maior bloco de DL entre aseQTLs.

8.6 A metilação de DNA pode afetar aseQTLs na sua relação com a ASE

Mecanismos epigenéticos também podem ser associados com a ASE (GUILLOCHEAU et al., 2019; SERRE et al., 2008), como a metilação de DNA (FERGUSON-SMITH, 2001; TIAN, 2014). Para identificar a importância do efeito da metilação desses aseQTLs sobre a ASE, nós observamos o nível de DRA dos SNPs ASE associados com esses aseQTLs que foram metilados nos mesmos animais. Poucos animais com metilação tiveram ASE significativa para

o SNP ASE associado ao aseQTL metilado. Para o aseQTL rs134229733, que teve 100% de metilação em oito animais, nós observamos um pequeno desequilíbrio alélico para o SNP ASE associado, variando de 0,003 a 0,137 em seis animais. Três desses animais apresentaram o genótipo do SNP ASE 1|0 (sendo a barra um indicador da separação entre dois haplótipos de um animal e os números 1 e 0 representativos de um alelo alternativo ou referência em cada haplótipo, respectivamente) com o alelo alternativo mais expresso, sendo o genótipo do aseQTL 0|1. Dois animais tiveram os mesmos genótipos para o aseQTL e o SNP ASE, mas expressaram apenas o alelo referência, com os menores valores de razão do desbalanço alélico. Curiosamente, como o alelo referência do aseQTL é um G e o alternativo é um T, o alelo do SNP ASE mais expresso nos três casos com maiores valores de DRA está supostamente sobre o controle de um aseQTL cujo nucleotídeo está metilado. A partir desses resultados, nós acreditamos que esse aseQTL está em DL com algum SNP responsável pela ASE neste caso.

O aseQTL rs110532113 apresentou mais animais metilados, mas somente um deles com ASE significativa no SNP ASE associado. O aseQTL apresentou 14% de metilação nesse animal e o SNP ASE teve um valor de razão do desequilíbrio alélico de 0,104. O aseQTL metilado mais próximo do TSS do respectivo SNP ASE associado, foi o rs110407852, com metilação no alelo alternativo. O SNP ASE correspondente, dentro do gene *VCP*, foi monoalélico em nove animais, expressando apenas o alelo em fase com o alelo referência do aseQTL, assim, indicando o silenciamento mediado pela metilação do alelo alternativo do aseQTL sob o alelo do SNP ASE associado a ele.

Três animais mostraram uma razão de desequilíbrio alélico de 0,5 para o SNP ASE rs42719199, expressando somente o alelo alternativo, com taxas de metilação 52,631%, 50% e 36,84% para o aseQTL associado rs137789368 e 32,30%, 80%, 55,75% para outro aseQTL associado, rs136787973. Esses aseQTLs não estão em DL entre si e o aseQTL rs137789368 teve os genótipos 0|1, 1|1 e 1|1. Os genótipos para o aseQTL rs136787973 foram 1|0, 0|1 e 0|1. Ambos os aseQTLs tem o alelo referência T e o alelo alternativo C. Considerando somente os aseQTLs com genótipos heterozigotos, para o aseQTL rs137789368, a citosina metilada pode impossibilitar a expressão do alelo referência do SNP ASE em um animal, no qual o aseQTL heterozigoto mostrou a maior porcentagem de metilação entre os demais (52,63%). Para o aseQTL rs136787973, com três genótipos heterozigotos, nós supomos que para animais com os genótipos dos aseQTLs 0|1, a citosina metilada no alelo alternativo pode silenciar a expressão do alelo referência do SNP ASE. Esses dois animais mostraram elevadas taxas de metilação neste aseQTL (80% e 55,75%). A metilação destes dois aseQTLs citados pode ser um dos fatores que podem regular a ASE do SNP rs42719199, favorecendo a expressão do alelo

alternativo. A Tabela 3 apresenta uma tabela exemplificando os dois aseQTLs que apresentaram metilação e foram associados a esse SNP ASE.

Tabela 3 – Genótipos e taxa de metilação dos *loci* quantitativos associados à expressão alelo-específica (aseQTLs) rs136787973 e rs137789368, associados ao SNP com expressão alelo-específica (SNP ASE) rs42719199, com os respectivos genótipos e valores de desequilíbrio da razão alélica (DRA) em três animais da população.

Animal	DRA SNP ASE: rs42719199	Genótipos SNP ASE: rs42719199	Genótipos aseQTL: rs137789368	Genótipos aseQTL: rs136787973	% Metilação aseQTL: rs137789368	% Metilação aseQTL: rs136787973
1	0,5	1(exp) 0	0 1 (met)	1(met) 0	52,631	32,30
2	0,5	1(exp) 0	1(met) 1(met)	0 1(met)	50,0	80,0
3	0,5	1(exp) 0	1(met) 1(met)	0 1(met)	36,84	55,75

Genótipos: “1” indica a presença do alelo alternativo e “0” o alelo referência. A barra “|” separa dois haplótipos de cada animal. O termo (exp) indica o alelo expresso e (met) indica o alelo metilado.

Nós observamos que é inconclusivo uma análise de metilação considerando apenas a exata posição da metilação em relação à posição dos aseQTLs, devido à dificuldade de se determinar a variante reguladora causativa. Em análises futuras, o perfil de metilação será estatisticamente associado à outros SNPs, que podem ser comparados aos SNPs ASE e aseQTLs. Nós acreditamos que essa análise indicará múltiplos SNPs metilados afetando a expressão alélica dos SNPs ASE.

8.7 Sítios de ligação pra miRNAs são potencialmente modificados por alelos de aseQTLs

Identificamos 162 aseQTLs nos locais de ligação ao miRNA, alterando a afinidade de ligação de 149 miRNAs, distribuídos em 1.448 pares de miRNA e alvo. Alguns estudos relacionaram o desequilíbrio alélico com a ação *cis*-reguladora dos sítios de ligação ao miRNA (KIM; BARTEL, 2009; RAMACHANDRAN et al., 2016), incluindo em músculo bovino (GUILLOCHEAU et al., 2019). Os miRNAs bta-miR-423-5p e bta-miR-486 foram diferencialmente expressos entre grupos extremos para marmoreio (OLIVEIRA et al., 2018) e CAR (DE OLIVEIRA et al., 2018), respectivamente. Seis genes foram potencialmente afetados por aseQTLs contendo sítios para o miRNA bta-miR-423-5p: *CMYA5*, *CUEDC1*, *DNAJC21*,

EIF5, *NXN* e *XIRP1*. O miRNA bta-miR-486 pode se ligar a aseQTLs que afetam a ASE dos genes *ACOT13*, *CAB39*, *CMYA5*, *CUTC*, *DNAJC21*, *MBNL2*, *NXN*, *STBD1* e *XIRP1*.

Catorze ASE genes foram afetados por aseQTLs em sítios de ligação para miRNAs que também foram preditos como reguladores desses ASE genes, como o aseQTL rs42437277, associado com um SNP ASE no gene *ASB5*, que foi predito a afetar a ligação de cinco miRNAs. A integração dos nossos dados com interações conhecidas de miRNAs e seus genes alvo obtidos a partir do banco de dados do MiRWalk oferece maior confiabilidade a esses resultados.

A maioria dos aseQTLs preditos a afetar sítios de ligação de miRNAs estão localizados em regiões de introns. Pequenos RNAs de interferência (siRNAs) podem se ligar à regiões de introns e regular o splicing, sendo que uma proteína Ago é necessária para esse efeito (ALLÓ et al., 2009). Como o modo de ação mais comum nos mamíferos é de que os miRNAs associados a proteínas Ago podem levar à deadenilação e conseqüentemente a desestabilização do mRNA (BARTEL, 2018), podemos relacionar a presença de aseQTLs intrônicos em sítios de ligação de miRNAs com a queda no número de transcritos do SNP ASE associado, o que pode também regular o splicing alternativo.

8.8 Integração dos resultados e conclusões finais

Interessantemente, o aseQTL rs109540393, que sobrepôs a nove regiões de QTL associadas a fenótipos de qualidade de carne obtidas na nossa população (TIZIOTO et al., 2013), foi predito a possuir os três mecanismos reguladores analisados nesse estudo. Esse aseQTL foi metilado, potencialmente se encontra dentro de um sítio de ligação para o FT expresso em músculo bovino AHR, e também afeta a ligação de seis miRNAs expressos em músculo da nossa população. O FT AHR foi negativamente associado com envelhecimento (ECKERS et al., 2016) e angiogênese (ICHIHARA et al., 2007). MiRNAs podem regular genes envolvidos com doenças musculares (ALEXANDER; KUNKEL, 2015), bem como proliferação de células do músculo, hipertrofia e atrofia (GÜLLER; RUSSELL, 2010).

Os aseQTLs associados ao gene ASE *CMYA5* mostraram mais sobreposições com QTLs de Nelore (35) e o maior número de aseQTLs que afetam sítios de ligação de miRNA (7) também foram associados a esse gene, incluindo dois miRNAs diferencialmente expressos entre animais divergentes para marmoreio e CAR (DE OLIVEIRA et al., 2018; OLIVEIRA et al., 2018). AseQTLs e seus respectivos SNPs ASE do gene *CMYA5* apresentaram grandes blocos de DL, sendo o tagSNP do maior bloco de aseQTLs, rs137303208, localizado em um intron específico presente em duas das três isoformas do *CMYA5*. Esse aseQTL intrônico

apresentou potencial em criar sítios pra dois FTs com o alelo referência, estando bem próximo do TSS do gene *CMYA5*, com cerca de 5kb de distância. Além disso, também foi predito o potencial deste aseQTL em criar sítios energeticamente favoráveis para a ligação dos miRNAs expressos em músculo da nossa população bta-miR-11992, bta-miR-2483-3p, bta-miR-330, bta-miR-10175-5p, bta-miR-219, bta-miR-382, bta-miR-362-5p, bta-miR-301b e bta-miR-2432. Esse mesmo aseQTL está presente em uma região de QTL associada às características de qualidade de carne BFAT, CL, LFAT, REA, WBSF0, WBSF7 e WHC. Tendo em vista a relação dos siRNAs com o splicing alternativo (ALLÓ et al., 2009) e diante desses resultados, nós acreditamos na possibilidade de que esse bloco possa participar da regulação do *splicing* do gene *CMYA5*, e todos esses resultados nos fazem acreditar no potencial desse gene como um marcador da qualidade de carne em músculo de Nelore.

Esses resultados demonstraram que os aseQTLs estão concentrados em regiões potencialmente associadas a fenótipos de qualidade de carne, incluindo AOL, marmoreio e maciez. A maciez é um fator essencial para a aceitabilidade do consumidor (HUFFMAN et al., 1996; MILLER et al., 2001). A gordura intramuscular é associada à maciez, palatabilidade e suculência da carne (WOOD et al., 2004), sendo outra característica sensorial que afeta a escolha do consumidor e o preço da carne (CHENG et al., 2015b). Assim, o impacto regulador dos aseQTLs pode afetar características de marmoreio e maciez, e os mecanismos envolvidos devem ser profundamente analisados em programas que visam o melhoramento sensorial da carne.

No entanto, apesar dessas múltiplas indicações do potencial regulador dos aseQTLs, a validação experimental é necessária para confirmar os mecanismos reguladores que influenciam a ASE, bem como verificar o impacto dessas variantes em outras populações. Com isso é possível o estudo da aplicabilidade destes resultados no melhoramento da qualidade de carne de Nelore.

9 Conclusão

Esse trabalho identificou 1.134 aseQTLs que podem afetar a expressão alélica de 85 genes. A metodologia utilizada para a identificação dos aseQTLs foi restritiva em relação à cobertura dos testes que, aliada ao baixo poder estatístico do teste não-paramétrico e à distribuição de heterozigotos na amostra, muitas vezes insuficiente para a execução dos testes, resultou na identificação de aseQTLs para uma parcela pequena dos SNPs ASE de Nelore. Apesar disso, acreditamos que esse critério agrega confiabilidade aos resultados obtidos, por proporcionar uma menor taxa de falsos positivos.

Grande parte dos aseQTLs demonstraram potencial capacidade de alterar sítios de ligação de miRNAs e de FTs, e também apresentaram metilação na nossa população. Além disso, as variantes desse estudo (SNPs ASE e aseQTLs) foram caracterizadas por possuírem alto DL entre si, principalmente em relação ao genes com múltiplos SNPs ASE, para os quais foram observadas associações simultâneas entre aseQTLs e mais de um SNP ASE. A dificuldade de se encontrar um aseQTL causal com base na predição dos mecanismos de regulação e do DL indicam que diferentes variantes possam estar agindo em conjunto na regulação da ASE dos SNPs desse estudo. Porém, por se tratar de predição *in silico* em uma espécie que possui poucos dados de anotação experimental do genoma funcional, experimentos adicionais são necessários para confirmar o potencial regulatório dos aseQTLs aqui descritos.

Os aseQTLs apresentaram potencial influência sobre fenótipos de qualidade de carne, por meio de numerosas sobreposições a regiões de QTLs identificadas previamente na mesma população experimental. Assim, esse tipo de variação pode afetar não somente a ASE, como também a expressão fenotípica de características de qualidade de carne em Nelore.

REFERÊNCIAS BIBLIOGRÁFICAS

- AFONSO, J. et al. Muscle transcriptome analysis reveals genes and metabolic pathways related to mineral concentration in *Bos indicus*. **Scientific Reports**, n. January, p. 1–11, 2019.
- AGRICULTURAL INTERAGENCY COMMITTEE, P. USDA Agricultural Projections to 2028. **USDA**, v. 1, p. 188, 2019.
- AJMONE-MARSAN, P. et al. On the origin of cattle: how aurochs became domestic and colonized the world. **Evolutionary Anthropology**, v. 19, n. June, p. 148–157, 2010.
- ALEXANDER, M. S.; KUNKEL, L. M. “Skeletal Muscle MicroRNAs: Their Diagnostic and Therapeutic Potential in Human Muscle Diseases”. **Journal of Neuromuscular Diseases**, v. 2, n. 1, p. 1–11, 2015.
- ALLIS, C. D.; JENUWEIN, T. The molecular hallmarks of epigenetic control. **Nature Reviews Genetics**, v. 17, n. 8, p. 487–500, 27 ago. 2016.
- ALLÓ, M. et al. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. **Nature Structural & Molecular Biology**, v. 16, n. 7, p. 717–724, 21 jul. 2009.
- ALMLÖF, J. C. et al. Powerful Identification of Cis-regulatory SNPs in Human Primary Monocytes Using Allele-Specific Gene Expression. **PLoS ONE**, v. 7, n. 12, p. e52260, 26 dez. 2012.
- ARCHANA, P. R. et al. Role of heat shock proteins in livestock adaptation to heat stress. **Journal of Dairy, Veterinary & Animal Research**. v. 5, n. 1, p. 13–19, 2017.
- ARIMBASSERI, A. G.; RIJAL, K.; MARAIA, R. J. Transcription termination by the eukaryotic RNA polymerase III. **Biochimica et Biophysica Acta - Gene Regulatory Mechanisms**, v. 1829, n. 3–4, p. 318–330, 2013.
- BAIK, M. et al. Association of DNA Methylation Levels with Tissue-specific Expression of Adipogenic and Lipogenic Genes in Longissimus dorsi Muscle of Korean Cattle. **Asian-Australasian Journal of Animal Sciences**, v. 27, n. 10, p. 1493–1498, 20 ago. 2014.
- BAKSHI, I. et al. Increasing Acyl CoA thioesterase activity alters phospholipid profile without effect on insulin action in skeletal muscle of rats. **Scientific Reports**, v. 8, n. 1, p. 13967, 18 dez. 2018.
- BARBOSA, B. R. P. et al. Tolerância ao calor em bovinos das raças Nelore branco, Nelore vermelho e Pantaneira. **Revista Brasileira de Saúde e Produção Animal**, v. 15, n. 4, p. 854–865, dez. 2014.
- BARENDSE, W. et al. Variation at the Calpain 3 gene is associated with meat tenderness in zebu and composite breeds of cattle. **BMC Genetics**, v. 9, n. 1, p. 41, 1 dez. 2008.
- BARRETT, L. W.; FLETCHER, S.; WILTON, S. D. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. **Cellular and Molecular Life Sciences**, v. 69, n. 21, p. 3613–3634, 27 nov. 2012.
- BARTEL, D. P. Metazoan MicroRNAs. **Cell**, v. 173, n. 1, p. 20–51, mar. 2018.

- BARTOLOMEI, M. S.; FERGUSON-SMITH, A. C. Mammalian Genomic Imprinting. **Cold Spring Harbor Perspectives in Biology**, v. 3, n. 7, p. a002592–a002592, 1 jul. 2011.
- BATTLE, A. et al. Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. **Genome Research**, v. 24, n. 1, p. 14–24, 2014.
- BECKER, P. B.; WORKMAN, J. L. Nucleosome Remodeling and Epigenetics. **Cold Spring Harbor Perspectives in Biology**, v. 5, n. 9, p. a017905–a017905, 1 set. 2013.
- BENJAMINI, Y.; HOCHBERG, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 57, n. 1, p. 289–300, jan. 1995.
- BHAT, S. et al. Effect of heat shock protein 70 polymorphism on thermotolerance in Tharparkar cattle. **Veterinary World**. v. 9, p. 113–117, 2016.
- BHAT, Z. F. et al. Role of calpain system in meat tenderness: A review. **Food Science and Human Wellness**, v. 7, n. 3, p. 196–204, 2018.
- BIANCHI, M. et al. Coordinated Actions of MicroRNAs with other Epigenetic Factors Regulate Skeletal Muscle Development and Adaptation. **International Journal of Molecular Sciences**, v. 18, n. 4, p. 840, 15 abr. 2017.
- BONETTO, A. et al. STAT3 Activation in Skeletal Muscle Links Muscle Wasting and the Acute Phase Response in Cancer Cachexia. **PLoS ONE**, v. 6, n. 7, p. e22538, 20 jul. 2011.
- BRAUD, M. et al. Genome-Wide microRNA Binding Site Variation between Extinct Wild Aurochs and Modern Cattle Identifies Candidate microRNA-Regulated Domestication Genes. **Frontiers in Genetics**, v. 8, 31 jan. 2017.
- BROOKS, J. C. et al. National Beef Tenderness Survey-1998. **Journal of Animal Science**, v. 78, n. 7, p. 1852, 2000.
- BROWNING, B. L.; BROWNING, S. R. Haplotypic analysis of Wellcome Trust Case Control Consortium data. **Human Genetics**, v. 123, n. 3, p. 273–280, 26 abr. 2008.
- BURRI, L.; THORESEN, G. H.; BERGE, R. K. The Role of PPAR α Activation in Liver and Muscle. **PPAR Research**, v. 2010, p. 1–11, 2010.
- BUSH, W. S.; MOORE, J. H. Chapter 11: Genome-Wide Association Studies. **PLoS Computational Biology**, v. 8, n. 12, p. e1002822, 27 dez. 2012.
- BUSHEY, A. M.; DORMAN, E. R.; CORCES, V. G. Chromatin insulators: regulatory mechanisms and epigenetic inheritance. **Molecular Cell**, v. 32, n. 4, p. 1–9, 2009.
- CAI, Y. et al. A Brief Review on the Mechanisms of miRNA Regulation. **Genomics, Proteomics & Bioinformatics**, v. 7, n. 4, p. 147–154, dez. 2009.
- CARVALHO, M. E. et al. Genomic regions and genes associated with carcass quality in Nelore cattle. **Genetics and Molecular Research**, v. 18, n. 1, 2019.
- CESAR, A. S. et al. Genome-wide association study for intramuscular fat deposition and composition in Nelore cattle. **BMC Genetics**, v. 15, n. 1, p. 39, 2014.

- CESAR, A. S. M. et al. Putative Regulatory Factors Associated with Intramuscular Fat Content. **PLOS ONE**, v. 10, n. 6, p. e0128350, 4 jun. 2015.
- CESAR, A. S. M. et al. Differences in the skeletal muscle transcriptome profile associated with extreme values of fatty acids content. **BMC Genomics**, v. 17, n. 1, p. 961, 22 dez. 2016.
- CESAR, A. S. M. et al. Identification of putative regulatory regions and transcription factors associated with intramuscular fat content traits. **BMC Genomics**, v. 19, n. 1, p. 499, 27 dez. 2018.
- CHAMBERLAIN, A. J. et al. Extensive variation between tissues in allele specific expression in an outbred mammal. **BMC Genomics**, v. 16, n. 1, p. 1–20, 2015.
- CHANG, C. C. et al. Second-generation PLINK: Rising to the challenge of larger and richer datasets. **GigaScience**, v. 4, n. 1, p. 1–16, 2015.
- CHEKULAEVA, M.; FILIPOWICZ, W. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. **Current Opinion in Cell Biology**, v. 21, n. 3, p. 452–460, jun. 2009.
- CHEN, X.-F. et al. An Osteoporosis Risk SNP at 1p36.12 Acts as an Allele-Specific Enhancer to Modulate LINC00339 Expression via Long-Range Loop Formation. **The American Journal of Human Genetics**, v. 102, n. 5, p. 776–793, maio 2018.
- CHENG, H. H. et al. Fine mapping of QTL and genomic prediction using allele-specific expression SNPs demonstrates that the complex trait of genetic resistance to Marek's disease is predominantly determined by transcriptional regulation. **BMC Genomics**, v. 16, n. 1, p. 1–9, 2015a.
- CHENG, W. et al. Marbling Analysis for Evaluating Meat Quality: Methods and Techniques. **Comprehensive Reviews in Food Science and Food Safety**, v. 14, n. 5, p. 523–535, set. 2015b.
- CHUANG, T.-J. et al. Assessment of imprinting- and genetic variation-dependent monoallelic expression using reciprocal allele descendants between human family trios. **Scientific Reports**, v. 7, n. 1, p. 7038, 1 dez. 2017.
- CHURCHILL, G. A.; DOERGE, R. W. Empirical threshold values for quantitative trait mapping. **Genetics**, v. 138, n. 3, p. 963–71, nov. 1994.
- CLAPIER, C. R.; CAIRNS, B. R. Chromatin remodeling complexes. **Fundamentals of Chromatin**, p. 69–146, 2014.
- CLÉMENT-ZIZA, M. et al. Natural genetic variation impacts expression levels of coding, non-coding, and antisense transcripts in fission yeast. **Molecular Systems Biology**, v. 10, n. 11, p. 764, 28 nov. 2014.
- COCHRAN, S. D. et al. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. **BMC Genetics**, v. 14, 2013.
- DAMON, M. et al. Comparison of Muscle Transcriptome between Pigs with Divergent Meat Quality Phenotypes Identifies Genes Related to Muscle Metabolism and Structure. **PLoS ONE**, v. 7, n. 3, p. e33763, 21 mar. 2012.

- DAMON, M. et al. Associations between muscle gene expression pattern and technological and sensory meat traits highlight new biomarkers for pork quality assessment. **Meat Science**, v. 95, n. 3, p. 744–754, nov. 2013.
- DANECEK, P. et al. The variant call format and VCFtools. **Bioinformatics**, v. 27, n. 15, p. 2156–2158, 1 ago. 2011.
- DE MELO COSTA, C. C. et al. Thermal equilibrium of Nelore cattle in tropical conditions: an investigation of circadian pattern. **Journal of Thermal Biology**, v. 74, p. 317–324, maio 2018.
- DE OLIVEIRA, P. S. et al. Identification of genomic regions associated with feed efficiency in Nelore cattle. **BMC Genetics**, v. 15, n. 1, p. 100, 26 dez. 2014.
- DE OLIVEIRA, P. S. N. et al. An integrative transcriptome analysis indicates regulatory mRNA-miRNA networks for residual feed intake in Nelore cattle. **Scientific Reports**, v. 8, n. 1, p. 17072, 20 dez. 2018.
- DE SOUZA, M. M. et al. Allele-specific expression is widespread in *Bos indicus* muscle and affects meat quality candidate genes. **Scientific Reports**, [s.d.].
- DE SOUZA, M. M. et al. A comprehensive manually-curated compendium of bovine transcription factors. **Scientific Reports**, v. 8, n. 1, p. 13747, 13 dez. 2018.
- DELBARE, S. Y. N.; CLARK, A. G. Allele-specific expression elucidates cis-regulatory logic. **PLOS Genetics**, v. 14, n. 11, p. e1007690, 1 nov. 2018.
- DESTEFANIS, G. et al. Relationship between beef consumer tenderness perception and Warner–Bratzler shear force. **Meat Science**, v. 78, n. 3, p. 153–156, mar. 2008.
- DO, D. N. et al. Genome-wide association and pathway analysis of feed efficiency in pigs reveal candidate genes and pathways for residual feed intake. **Frontiers in Genetics**, v. 5, 9 set. 2014.
- DOS SANTOS, R. **Nelore: A vitória Brasileira**. [s.l.] Editora Agropecuária Tropical, 2000.
- DUARTE, M. et al. **Calpastatin activity and beef tenderness of Nelore and Angus cattle fed two feeding strategies**. Proceedings of the 59th International Congress of Meat Science and Technology. **Anais...Izmir, Turkey.**: 2013
- DUNNE, P. G. et al. Colour of bovine subcutaneous adipose tissue: A review of contributory factors, associations with carcass and meat quality and its potential utility in authentication of dietary history. **Meat Science**, v. 81, n. 1, p. 28–45, jan. 2009.
- DURINCK, S. et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. **Bioinformatics**, v. 21, n. 16, p. 3439–3440, 15 ago. 2005.
- DYKES, I. M.; EMANUELI, C. Transcriptional and Post-transcriptional Gene Regulation by Long Non-coding RNA. **Genomics, Proteomics and Bioinformatics**, v. 15, n. 3, p. 177–186, 2017.
- ECKERS, A. et al. The aryl hydrocarbon receptor promotes aging phenotypes across species. **Scientific Reports**, v. 6, n. 1, p. 19618, 21 abr. 2016.

- EDWARDS, J. R. et al. Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns. **Genome Research**, v. 20, n. 7, p. 972–980, 1 jul. 2010.
- EDWARDS, J. R. et al. DNA methylation and DNA methyltransferases. **Epigenetics & Chromatin**, v. 10, n. 1, p. 23, 8 dez. 2017.
- ENDO, I. et al. Deletion of Vitamin D Receptor Gene in Mice Results in Abnormal Skeletal Muscle Development with Deregulated Expression of Myoregulatory Transcription Factors. **Endocrinology**, v. 144, n. 12, p. 5138–5144, dez. 2003.
- ESPIGOLAN, R. et al. Study of whole genome linkage disequilibrium in Nellore cattle. **BMC Genomics**, v. 14, n. 1, p. 305, 2013.
- FERGUSON-SMITH, A. C. Imprinting and the Epigenetic Asymmetry Between Parental Genomes. **Science**, v. 293, n. 5532, p. 1086–1089, 2001.
- GALVAGNI, F.; CAPO, S.; OLIVIERO, S. Sp1 and Sp3 physically interact and co-operate with GABP for the activation of the utrophin promoter. **Journal of Molecular Biology**, v. 306, n. 5, p. 985–996, mar. 2001.
- GASZNER, M.; FELSENFELD, G. Insulators: Exploiting transcriptional and epigenetic mechanisms. **Nature Reviews Genetics**, v. 7, n. 9, p. 703–713, 2006.
- GEL, B. et al. regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. **Bioinformatics**, p. btv562, 30 set. 2015.
- GUILLOCHEAU, G. M. et al. Survey of allele specific expression in bovine muscle. **Scientific Reports**, v. 9, n. 1, p. 4297, 12 dez. 2019.
- GÜLLER, I.; RUSSELL, A. P. MicroRNAs in skeletal muscle: their role and regulation in development, disease and function. **The Journal of Physiology**, v. 588, n. 21, p. 4075–4087, 1 nov. 2010.
- GULYAEVA, L. F.; KUSHLINSKIY, N. E. Regulatory mechanisms of microRNA expression. **Journal of Translational Medicine**, v. 14, n. 1, p. 143, 20 dez. 2016.
- GUO, J. et al. Expression of genes in the TGF- β signaling pathway is significantly deregulated in smooth muscle cells from aorta of aryl hydrocarbon receptor knockout mice. **Toxicology and Applied Pharmacology**, v. 194, n. 1, p. 79–89, jan. 2004.
- HIGGINS, M. G. et al. GWAS and eQTL analysis identifies a SNP associated with both residual feed intake and GFRA2 expression in beef cattle. **Scientific Reports**, v. 8, n. 1, p. 1–12, 2018.
- HIRSCHHORN, J. N.; DALY, M. J. Genome-wide association studies for common diseases and complex traits. **Nature Reviews Genetics**, v. 6, n. 2, p. 95–108, fev. 2005.
- HUFFMAN, K. L. et al. Effect of beef tenderness on consumer satisfaction with steaks consumed in the home and restaurant. **Journal of Animal Science**, v. 74, n. 1, p. 91, 1996.
- ICHIHARA, S. et al. A Role for the Aryl Hydrocarbon Receptor in Regulation of Ischemia-Induced Angiogenesis. **Arteriosclerosis, Thrombosis, and Vascular Biology**, v. 27, n. 6, p. 1297–1304, jun. 2007.

IMAMOTO, N. Heat stress-induced nuclear transport mediated by Hkeshi confers nuclear function of Hsp70s. **Current Opinion in Cell Biology**, v. 52, p. 82–87, 2018.

INTERNATIONAL HAPMAP CONSORTIUM. A haplotype map of the human genome. **Nature**, v. 437, n. 7063, p. 1299–1320, out. 2005.

JACINTA-FERNANDES, A. et al. Allele-specific miRNA-binding analysis identifies candidate target genes for breast cancer risk. **BioRxiv**, 2019.

JELTSCH, A.; BROCHE, J.; BASHTRYKOV, P. Molecular Processes Connecting DNA Methylation Patterns with DNA Methyltransferases and Histone Modifications in Mammalian Genomes. **Genes**, v. 9, n. 11, p. 566, 21 nov. 2018.

JEREMIAH, L. E.; CARPENTER, Z. L.; SMITH, G. C. BEEF COLOR AS RELATED TO CONSUMER ACCEPTANCE AND PALATABILITY. **Journal of Food Science**, v. 37, n. 3, p. 476–479, maio 1972.

JIANG, Z. et al. DNA methylomes of bovine gametes and in vivo produced preimplantation embryos. **Biology of Reproduction**, jun. 2018.

JIN, W. et al. Characterization of microRNA expression in bovine adipose tissues: a potential regulatory mechanism of subcutaneous adipose tissue development. **BMC Molecular Biology**, v. 11, n. 1, p. 29, 2010.

JOHNSON, R. C. et al. Accounting for multiple comparisons in a genome-wide association study (GWAS). **BMC Genomics**, v. 11, n. 1, p. 2–7, 2010.

JONAS, S.; IZAURRALDE, E. Towards a molecular understanding of microRNA-mediated gene silencing. **Nature Reviews Genetics**, v. 16, n. 7, p. 421–433, 16 jul. 2015.

JOSEPHS, E. B. et al. Association mapping reveals the role of purifying selection in the maintenance of genomic variation in gene expression. **Proceedings of the National Academy of Sciences**, v. 112, n. 50, p. 15390–15395, 15 dez. 2015.

KAMI, K.; SENBA, E. In Vivo Activation of STAT3 Signaling in Satellite Cells and Myofibers in Regenerating Rat Skeletal Muscles. **Journal of Histochemistry & Cytochemistry**, v. 50, n. 12, p. 1579–1589, 26 dez. 2002.

KANG, H. W. et al. Thioesterase superfamily member 2/acyl-CoA thioesterase 13 (Them2/Acot13) regulates hepatic lipid and glucose metabolism. **The FASEB Journal**, v. 26, n. 5, p. 2209–2221, maio 2012.

KHALIL, A. M. et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. **Proceedings of the National Academy of Sciences**, v. 106, n. 28, p. 11667–11672, 14 jul. 2009.

KHANSEFID, M. et al. Comparing allele specific expression and local expression quantitative trait loci and the influence of gene expression on complex trait variation in cattle. **BMC Genomics**, v. 19, n. 1, p. 793, 3 dez. 2018.

KIM, J.; BARTEL, D. P. Allelic imbalance sequencing reveals that single-nucleotide polymorphisms frequently alter microRNA-directed repression. **Nature Biotechnology**, v. 27, n. 5, p. 472–477, 26 maio 2009.

- KINTSCHER, U. et al. PPAR α Inhibits TGF- β -Induced β 5 Integrin Transcription in Vascular Smooth Muscle Cells by Interacting With Smad4. **Circulation Research**, v. 91, n. 11, 29 nov. 2002.
- KNOWLES, D. A. et al. Allele-specific expression reveals interactions between genetic variation and environment. **Nature Methods**, v. 14, n. 7, p. 699–702, 22 jul. 2017.
- KONHEIM, Y. L.; WOLFORD, J. K. Association of a promoter variant in the inducible cyclooxygenase-2 gene (PTGS2) with type-2 diabetes mellitus in Pima Indians. **Human Genetics**, v. 113, n. 5, p. 377–381, 1 out. 2003.
- KOSE, S.; FURUTA, M.; IMAMOTO, N. Hikeshi , a Nuclear Import Carrier for Hsp70s , Protects Cells from Heat Shock-Induced Nuclear Damage. **Cell**, v. 149, n. 3, p. 578–589, 2012.
- KOULOUMENTA, A.; MAVROIDIS, M.; CAPETANAKI, Y. Proper Perinuclear Localization of the TRIM-like Protein Myospryn Requires Its Binding Partner Desmin. **Journal of Biological Chemistry**, v. 282, n. 48, p. 35211–35221, 30 nov. 2007.
- KOZOMARA, A.; GRIFFITHS-JONES, S. miRBase: annotating high confidence microRNAs using deep sequencing data. **Nucleic Acids Research**, v. 42, n. D1, p. D68–D73, jan. 2014.
- KRAVITZ, S. N.; GREGG, C. New subtypes of allele-specific epigenetic effects: implications for brain development, function and disease. **Current Opinion in Neurobiology**, v. 59, p. 69–78, dez. 2019.
- KRUGER, J.; REHMSMEIER, M. RNAhybrid: microRNA target prediction easy, fast and flexible. **Nucleic Acids Research**, v. 34, n. Web Server, p. W451–W454, 1 jul. 2006.
- KWAK, H.; LIS, J. T. Control of Transcriptional Elongation. **Annual Review of Genetics**, v.47, p.483-508, 2013.
- LAGE, J. F. et al. Influence of genetic type and level of concentrate in the finishing diet on carcass and meat quality traits in beef heifers. **Meat Science**, v. 90, n. 3, p. 770–774, mar. 2012.
- LARSON, N. B. et al. Comprehensively Evaluating cis -Regulatory Variation in the Human Prostate Transcriptome by Using Gene-Level Allele-Specific Expression. **The American Journal of Human Genetics**, v. 96, n. 6, p. 869–882, jun. 2015.
- LEE, R. C.; FEINBAUM, R. L.; AMBROS, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. **Cell**, v. 75, n. 5, p. 843–854, dez. 1993.
- LEE, S.; LEE, D. K. What is the proper way to apply the multiple comparison test? **Korean Journal of Anesthesiology**, v. 71, n. 5, p. 353–360, 2018.
- LEE, Y. et al. MicroRNA maturation: Stepwise processing and subcellular localization. **EMBO Journal**, v. 21, n. 17, p. 4663–4670, 2002.
- LEE, Y.; RIO, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. **Annual Review of Biochemistry**, v. 84, n. 1, p. 291–323, 2 jun. 2015.

- LOFTUS, R. T. et al. Evidence for two independent domestications of cattle. **Proceedings of the National Academy of Sciences**, v. 91, n. 7, p. 2757–2761, 29 mar. 1994.
- LONG, H. K. et al. Protection of CpG islands from DNA methylation is DNA-encoded and evolutionarily conserved. **Nucleic Acids Research**, v. 44, n. 14, p. 6693–6706, 19 ago. 2016.
- LOPEZ, A. J. ALTERNATIVE SPLICING OF PRE-mRNA: Developmental Consequences and Mechanisms of Regulation. **Annual Review of Genetics**, v. 32, n. 1, p. 279–305, dez. 1998.
- LOURENCO, B. D. et al. Introduction to Genomic Selection. **Notes for a short course taught in 2017 in Maringa, Brazil**. p. 1–23, 2017.
- LUSE, D. S. The RNA polymerase II preinitiation complex: Through what pathway is the complex assembled? **Transcription**, v. 5, n. NOV, 2013.
- MANCINI, G. et al. Association between single nucleotide polymorphisms (SNPs) and milk production traits in Italian Brown cattle. **Livestock Science**, v. 157, n. 1, p. 93–99, 2013.
- MARINO, R. et al. Proteolytic pattern of myofibrillar protein and meat tenderness as affected by breed and aging time. **Meat Science**, v. 95, n. 2, p. 281–287, out. 2013.
- MASON, V. C. et al. eQTL discovery and their association with severe equine asthma in European Warmblood horses. **BMC Genomics**, v. 19, n. 1, p. 1–15, 2018.
- MATSUDA, S.; KITAGISHI, Y. Peroxisome Proliferator-Activated Receptor and Vitamin D Receptor Signaling Pathways in Cancer Cells. **Cancers**, v. 5, n. 4, p. 1261–1270, 21 out. 2013.
- MCDANIELL, R. et al. Heritable Individual-Specific and Allele-Specific Chromatin Signatures in Humans. **Science**, v. 328, n. 5975, p. 235–239, 9 abr. 2010.
- MCLAREN, W. et al. The Ensembl Variant Effect Predictor. **Genome Biology**, v. 17, n. 1, p. 122, 6 dez. 2016.
- MENG, Y. et al. Introns targeted by plant microRNAs: a possible novel mechanism of gene regulation. **Rice**, v. 6, n. 1, p. 8, 15 dez. 2013.
- MERCER, T. R.; MATTICK, J. S. Structure and function of long noncoding RNAs in epigenetic regulation. **Nature Structural and Molecular Biology**, v. 20, n. 3, p. 300–307, 2013.
- MEYDAN, H.; YILDIZ, M. A. Monitoring of the genes that confer genetic resistance to Marek's disease in chicken using allele-specific expression analyses. **Journal of Biotechnology**, v. 280, p. S36–S37, ago. 2018.
- MILLER, M. F. et al. Consumer thresholds for establishing the value of beef tenderness. **Journal of Animal Science**, v. 79, n. 12, p. 3062, 2001.
- MOHAMMADI, P. et al. Quantifying the regulatory effect size of cis -acting genetic variation using allelic fold change. **Genome Research**, v. 27, n. 11, p. 1872–1884, nov. 2017.
- MOKRY, F. et al. Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. **BMC Genomics**, v. 15, n. Suppl 7, p. S6, 2014.

- MUDADU, M. A. et al. Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. **BMC Genomics**, v. 17, n. 1, p. 235, 15 dez. 2016.
- NELSON, C. E.; HERSH, B. M.; CARROLL, S. B. The regulatory content of intergenic DNA shapes genome architecture. **Genome biology**, v. 5, n. 4, p. R25, 2004.
- NICOLAE, D. L. et al. Trait-Associated SNPs Are More Likely to Be eQTLs : Annotation to Enhance Discovery from GWAS. **PLoS Genetics**. v. 6, n. 4, 2010.
- NILLER, H. H.; DEMCSÁK, A.; MINAROVITS, J. DNA Methylation in Eukaryotes: Regulation and Function. In: **Cellular Ecophysiology of Microbe: Hydrocarbon and Lipid Interactions**. Cham: Springer International Publishing, 2018. p. 509–570.
- NOGALES, E.; LOUDER, R. K.; HE, Y. Structural Insights into the Eukaryotic Transcription Initiation Machinery. **Annual Review of Biophysics**. v. 46, p.59-83, 2017.
- NYHOLT, D. R. A Simple Correction for Multiple Testing for Single-Nucleotide Polymorphisms in Linkage Disequilibrium with Each Other. **American Journal of Human Genetics**, v. 74, n. 4, p. 765–769, 2004.
- OLIVEIRA, G. B. et al. Integrative analysis of microRNAs and mRNAs revealed regulation of composition and metabolism in Nelore cattle. **BMC Genomics**, v. 19, n. 1, p. 126, 7 dez. 2018.
- PACIS, A. et al. Gene activation precedes DNA demethylation in response to infection in human dendritic cells. **Proceedings of the National Academy of Sciences of the United States of America**, v. 116, n. 14, p. 6938–6943, 2019.
- PASSOS, B. R. et al. Tolerância ao calor em bovinos das raças Nelore branco , Nelore vermelho e Pantaneira. **Revista Brasileira de Saúde e Produção Animal**. p. 854–865, 2014.
- PASTINEN, T. Genome-wide allele-specific analysis: Insights into regulatory variation. **Nature Reviews Genetics**, v. 11, n. 8, p. 533–538, 2010.
- PATNALA, R.; CLEMENTS, J.; BATRA, J. Candidate gene association studies: A comprehensive guide to useful in silico tools. **BMC Genetics**, v. 14, 2013.
- PENNY, I. F.; DRANSFIELD, E. Relationship between toughness and troponin T in conditioned beef. **Meat Science**, v. 3, n. 2, p. 135–141, abr. 1979.
- PEREIRA, A. S. C. et al. Growth performance, and carcass and meat quality traits in progeny of Poll Nellore, Angus and Brahman sires under tropical conditions. **Animal Production Science**, v. 55, n. 10, p. 1295, 2015.
- PFEIFER, K. Mechanisms of Genomic Imprinting. **The American Journal of Human Genetics**, v. 67, n. 4, p. 777–787, out. 2000.
- PICKRELL, J. K. et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. **Nature**, v. 464, n. 7289, p. 768–772, 10 abr. 2010.
- PIÓRKOWSKA, K. et al. Transcript variants of a region on SSC15 rich in QTLs associated with meat quality in pigs. **Annals of Animal Science**, v. 17, n. 3, p. 703–715, 26 jul. 2017.

- QIN, W.; LEONHARDT, H.; PICHLER, G. Regulation of DNA methyltransferase 1 by interactions and modifications. **Nucleus**, v. 2, n. 5, p. 392–402, 27 set. 2011.
- QUACH, H. et al. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. **Cell**, v. 167, n. 3, p. 643– 656.e17, out. 2016.
- RAMACHANDRAN, S. et al. Cis-acting single nucleotide polymorphisms alter MicroRNA-mediated regulation of human brain-expressed transcripts. **Human Molecular Genetics**, p. ddw317, 16 set. 2016.
- REESE, J. C. Basal transcription factors. **Current Opinion in Genetics & Development**, v. 13, n. 2, p. 114–118, abr. 2003.
- RIBEIRO, A. R. B. et al. Heat tolerance of Nelore , Senepol x Nelore and Angus x Nelore heifers in the southeast region of Brazil. **South African Journal of Animal Science**. v. 39, n. Supplement 1, p. 2008–2010, 2009.
- RODRIGUES, R. T. DE S. et al. Differences in Beef Quality between Angus (*Bos taurus taurus*) and Nelore (*Bos taurus indicus*) Cattle through a Proteomic and Phosphoproteomic Approach. **PLOS ONE**, v. 12, n. 1, p. e0170294, 19 jan. 2017.
- ROSENTHAL, R.; RUBIN, D. B. Multiple contrasts and ordered Bonferroni procedures. **Journal of Educational Psychology**, v. 76, n. 6, p. 1028–1034, 1984.
- SAINSBURY, S.; BERNECKY, C.; CRAMER, P. Structural basis of transcription initiation by RNA polymerase II. **Nature Reviews Molecular Cell Biology**, v. 16, n. 3, p. 129–143, 18 mar. 2015.
- SAMBROOK J., FRITSCH E. F., M. T. **Molecular Cloning: A Laboratory Manual**. [s.l.] Cold Spring Harbor Laboratory Press, 1989.
- SANDRI, M. Protein breakdown in muscle wasting: Role of autophagy-lysosome and ubiquitin-proteasome. **The International Journal of Biochemistry & Cell Biology**, v. 45, n. 10, p. 2121–2129, out. 2013.
- SANTANA, M. L. et al. Structure and genetic diversity of Brazilian Zebu cattle breeds assessed by pedigree analysis. **Livestock Science**, v. 187, p. 6–15, maio 2016.
- SANTIAGO, A. A. Gado nelore: 100 anos de seleção. **Ver Curiosidades**. São Paulo: Dos criadores, 1987.
- SANTIAGO, G. G. et al. Genomewide association study for production and meat quality traits in Canchim beef cattle. **Journal of Animal Science**. v.95, n.8. p. 3381-3390, 2018.
- SARPARANTA, J. et al. Interactions with M-band Titin and Calpain 3 Link Myospryn (CMYA5) to Tibial and Limb-girdle Muscular Dystrophies. **Journal of Biological Chemistry**, v. 285, n. 39, p. 30304–30315, 24 set. 2010.
- SARTORI, R. et al. BMP signaling controls muscle mass. **Nature Genetics**, v. 45, n. 11, p. 1309–1318, 29 nov. 2013.
- SCHEFERS, J. M.; WEIGEL, K. A. Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. **Animal Frontiers**, v. 2, n. 1, p. 4–9, 2012.

SEABURY, C. M. et al. Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. **BMC Genomics**, v. 18, n. 1, p. 386, 18 dez. 2017.

SERRE, D. et al. Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression. **PLoS genetics**, v. 4, n. 2, p. e1000006, 29 fev. 2008.

SILVA-VIGNATO, B. et al. Comparative muscle transcriptome associated with carcass traits of Nellore cattle. **BMC Genomics**, v. 18, n. 1, p. 506, 3 dez. 2017.

SMITH, S. B. et al. Cellular regulation of bovine intramuscular adipose tissue development and composition¹². **Journal of Animal Science**, v. 87, n. suppl_14, p. E72–E82, 1 abr. 2009.

SOUZA, M. M. DE. **Avaliação da expressão alélica em genes associados à maciez de carne em animais Nelore por meio de análise transcriptômica**. [s.l.] Universidade Federal de Alfenas, 2016.

SPITZ, F. Gene regulation at a distance: From remote enhancers to 3D regulatory ensembles. **Seminars in Cell & Developmental Biology**, v. 57, p. 57–67, set. 2016.

SPITZ, F.; FURLONG, E. E. M. Transcription factors: from enhancer binding to developmental control. **Nature Reviews Genetics**, v. 13, n. 9, p. 613–626, 7 set. 2012.

STACHOWIAK, M.; FLISIKOWSKI, K. Analysis of allele-specific expression of seven candidate genes involved in lipid metabolism in pig skeletal muscle and fat tissues reveals allelic imbalance of ACACA, LEP, SCD, and TNF. **Journal of Applied Genetics**, v. 60, n. 1, p. 97–101, 26 fev. 2019.

STACHOWIAK, M.; SZCZERBAL, I.; FLISIKOWSKI, K. Investigation of allele-specific expression of genes involved in adipogenesis and lipid metabolism suggests complex regulatory mechanisms of PPARGC1A expression in porcine fat tissues. **BMC Genetics**, v. 19, n. 1, p. 107, 29 dez. 2018.

STEPHANOU, A.; LATCHMAN, D. S. Transcriptional regulation of the heat shock protein genes by STAT family transcription factors. **Gene expression**, v. 7, n. 4–6, p. 311–9, 1999.

STICHT, C. et al. miRWalk: An online resource for prediction of microRNA binding sites. **PLOS ONE**, v. 13, n. 10, p. e0206239, 18 out. 2018.

TAN, Z. et al. Allele-Specific Targeting of microRNAs to HLA-G and Risk of Asthma. **The American Journal of Human Genetics**, v. 81, n. 4, p. 829–834, out. 2007.

TAYLOR, R. G. et al. Is Z-disk degradation responsible for postmortem tenderization? **Journal of Animal Science**, v. 73, n. 5, p. 1351–1367, 1 maio 1995.

TCHATCHOU, S. et al. A variant affecting a putative miRNA target site in estrogen receptor (ESR) 1 is associated with breast cancer risk in premenopausal women. **Carcinogenesis**, v. 30, n. 1, p. 59–64, jan. 2009.

TEMPLETON, A. R. **Genética de populações e teoria microevolutiva**. [s.l.] SBG, 2011.

THOMAS-CHOLLIER, M. et al. Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. **Nature Protocols**, v. 6, n. 12, p. 1860–1869, 3 dez. 2011.

- TIAN, X. (CINDY). Genomic Imprinting in Farm Animals. **Annual Review of Animal Biosciences**, v. 2, n. 1, p. 23–40, 2014.
- TIZIOTO, P. C. et al. A SNP in ASAP1 gene is associated with meat quality and production traits in Nelore breed. **Meat Science**, v. 92, n. 4, p. 855–857, dez. 2012.
- TIZIOTO, P. C. et al. Genome scan for meat quality traits in Nelore beef cattle. **Physiological Genomics**, v. 45, n. 21, p. 1012–1020, 2013.
- TIZIOTO, P. C. et al. Detection of quantitative trait loci for mineral content of Nelore longissimus dorsi muscle. **Genetics Selection Evolution**, v. 47, n. 1, p. 15, 2015a.
- TIZIOTO, P. C. et al. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC Genomics**, v. 16, n. 1, p. 242, 25 dez. 2015b.
- TONG, F. NIH Public Access. **Vision Research**, v. 4, n. 7, p. 539–551, 2006.
- TRIANAPHYLLOPOULOS, K. A.; IKONOMOPOULOS, I.; BANNISTER, A. J. Epigenetics and inheritance of phenotype variation in livestock. **Epigenetics and Chromatin**, v. 9, n. 1, p. 1–18, 2016.
- TSOUPRI, E.; CAPETANAKI, Y. Myospryn: a multifunctional desmin-associated protein. **Histochemistry and Cell Biology**, v. 140, n. 1, p. 55–63, 9 jul. 2013.
- VERDIN, E.; OTT, M. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. **Nature Reviews Molecular Cell Biology**, v. 16, n. 4, p. 258–264, 30 abr. 2015.
- WANG, M. et al. Putative bovine topological association domains and CTCF binding motifs can reduce the search space for causative regulatory variants of complex traits. **BMC Genomics**, v. 19, n. 1, p. 1–17, 2018.
- WANG, S. et al. Tagging SNP-set selection with maximum information based on linkage disequilibrium structure in genome-wide association studies. **Bioinformatics**, v. 33, n. 14, p. 2078–2081, 2017.
- WANG, Z. et al. Downregulation of microRNA-132 by DNA hypermethylation is associated with cell invasion in colorectal cancer. **OncoTargets and Therapy**, p. 3639, dez. 2015.
- WATANABE, K. et al. Functional mapping and annotation of genetic associations with FUMA. **Nature Communications**, v. 8, n. 1, p. 1–10, 2017.
- WIGGANS, G. R. et al. Genomic Selection in Dairy Cattle: The USDA Experience. **Annual Review of Animal Biosciences**, v. 5, n. 1, p. 309–327, 2017.
- WILLIAMSON, J. F. et al. Genomics Genomic evolution in domestic cattle : Ancestral haplotypes and healthy beef. **Genomics**, v. 97, n. 5, p. 304–312, 2011.
- WOOD, J. . et al. Effects of fatty acids on meat quality: a review. **Meat Science**, v. 66, n. 1, p. 21–32, jan. 2004.
- WOOD, J. D. et al. Fat deposition, fatty acid composition and meat quality: A review. **Meat Science**, v. 78, n. 4, p. 343–358, abr. 2008.

- XIANG, R. et al. Quantifying the contribution of sequence variants with regulatory and evolutionary significance to 34 bovine complex traits. **Proceedings of the National Academy of Sciences of the United States of America**, v. 116, n. 39, p. 19398–19408, 2019.
- XU, X. et al. The molecular characterization and associations of porcine cardiomyopathy associated 5 (CMYA5) gene with carcass trait and meat quality. **Molecular Biology Reports**, v. 38, n. 3, p. 2085–2090, 22 mar. 2011.
- YAN, H. et al. Allelic variation in human gene expression. **Science**, v. 297, n. 5584, p. 1143, 2002.
- YAO, C. et al. Dynamic Role of trans Regulation of Gene Expression in Relation to Complex Traits. **American Journal of Human Genetics**, v. 100, n. 4, p. 571–580, 2017.
- YAO, J. Imaging Transcriptional Regulation of Eukaryotic mRNA Genes: Advances and Outlook. **Journal of Molecular Biology**, v. 429, n. 1, p. 14–31, jan. 2017.
- ZEMPLINI, J. et al. Gene regulation by dietary microRNAs. **Canadian Journal of Physiology and Pharmacology**, v. 93, n. 12, p. 1097–1102, dez. 2015.
- ZHANG, C. et al. Interleukin-6/Signal Transducer and Activator of Transcription 3 (STAT3) Pathway Is Essential for Macrophage Infiltration and Myoblast Proliferation during Muscle Regeneration. **Journal of Biological Chemistry**, v. 288, n. 3, p. 1489–1499, 18 jan. 2013.
- ZHANG, X.; CAL, A. J.; BOREVITZ, J. O. Genetic architecture of regulatory variation in *Arabidopsis thaliana*. **Genome Research**, v. 21, n. 5, p. 725–733, 1 maio 2011.
- ZOU, J. et al. Leveraging allele-specific expression to refine fine-mapping for eQTL studies. **bioRxiv**, p. 257279, 2018.