

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL

ARQUITETURA GENÉTICA DO CONSUMO
ALIMENTAR RESIDUAL EM BOVINOS NELORE

Beatriz Pressi Molina da Silva

Zootecnista

2018

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

**ARQUITETURA GENÉTICA DO CONSUMO
ALIMENTAR RESIDUAL EM BOVINOS NELORE**

Beatriz Pressi Molina da Silva

Orientador: Prof. Dr. Josineudson Augusto II de Vasconcelos Silva

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

2018

O48f Silva, Beatriz Pressi Molina da
Arquitetura genética do consumo alimentar residual em bovinos
Nelore / Beatriz Pressi Molina da Silva. – – Jaboticabal, 2018
x, 85 p. : il. ; 29 cm

Dissertação (mestrado) - Universidade Estadual Paulista,
Faculdade de Ciências Agrárias e Veterinárias, 2018
Orientador: Josineudson Augusto II de Vasconcelos Silva
Banca examinadora: Rusbel Raul Aspilcueta Borquis, Henrique
Nunes de Oliveira
Bibliografia

1. Estudo de associação. 2. Frequência alélica. 3. Fst.4. QTL. I.
Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 633.34:631.54

Ficha Catalográfica elaborada pela STATI - Biblioteca da UNESP
Campus de Jaboticabal/SP - Karina Gimenes Fernandes - CRB 8/7418



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: ARQUITETURA GENÉTICA DO CONSUMO ALIMENTAR RESIDUAL EM BOVINOS NELORE

AUTORA: BEATRIZ PRESSI MOLINA DA SILVA

ORIENTADOR: JOSINEUDSON AUGUSTO II DE VASCONCELLOS SILVA

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

Prof. Dr. JOSINEUDSON AUGUSTO II DE VASCONCELLOS SILVA
Departamento de Melhoramento e Nutrição Animal-FMVZ/UNESP / Botucatu/SP

Prof. Dr. RUBEL RAUL ASPILCUETA BORQUIS
Departamento de Zootecnia / Universidade Federal da Grande Dourados - Dourados/MS

Prof. Dr. HENRIQUE NUNES DE OLIVEIRA
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal

Jaboticabal, 16 de agosto de 2018

DADOS CURRICULARES DO AUTOR

Beatriz Pressi Molina da Silva, nascida em São Paulo – SP em 02 de Outubro de 1992, iniciou o curso de Zootecnia na Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp), Botucatu – SP em Março de 2011, tornando-se Zootecnista em Dezembro de 2016. Iniciou o curso de Pós-Graduação em Genética e Melhoramento Animal na Universidade Estadual Paulista “Júlio de Mesquita Filho” em Agosto de 2016, sob orientação do Prof. Dr. Josineudson Augusto II de Vasconcelos Silva. Foi bolsista CAPES, obtendo o título de mestre em Genética e Melhoramento Animal em Agosto de 2018.

EPÍGRAFE

“If it's not fun, you're not doing it right.”
Bob Basso

“...Opte por aquilo que faz seu coração vibrar... Apesar de todas as consequências.”
Osho Rajneesh

AGRADECIMENTOS

Agradeço aos meus pais, Neila e Mario, por todo amor, incentivo, dedicação e por serem inspiração de vida. Por eles e com eles sempre.

Ao meu orientador professor Augusto pela oportunidade em que me acolheu, por todo conhecimento e dedicação compartilhados. Foi um privilégio conviver e aprender ao seu lado, principalmente pelo exemplo de profissional que transmitiu a mim. Devo agradecê-lo também por tornar o processo agradável, pela paciência, por acreditar no trabalho e pelos momentos de parceria. O meu muito obrigada!

Ao Nelore Qualitas pela disponibilização dos dados e tornar possível a pesquisa de mestrado.

Aos companheiros de pós-graduação e de departamento: Alejandra, Amanda, Cintia, Felipe, Ricardo, Eduardo, Matheus e Ana Beatriz por toda ajuda e companheirismo, fazendo a caminhada mais leve.

Aos amigos de Botucatu e de Jaboticabal por se fazerem presentes nos melhores e piores momentos, vocês são especiais.

Aos membros da banca examinadora da qualificação, Dr. Diercles Francisco Cardoso e Dr. Henrique Nunes de Oliveira pelas correções e contribuições no trabalho.

Aos membros da banca examinadora da defesa, Dr. Rusbel Raul Aspilcueta Borquis e Dr. Henrique Nunes de Oliveira por colaborarem e fazerem parte na minha pós-graduação.

À Universidade Estadual Paulista, tanto FMVZ – Botucatu, quanto FCAV - Jaboticabal por fazerem parte da minha história.

A agência de fomento Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa e tornar possível a realização do mestrado.

A todos que colaboraram com a minha formação. Obrigada!

SUMÁRIO

1. INTRODUÇÃO.....	1
2. REVISÃO DE LITERATURA.....	4
2.1 Considerações gerais acerca da eficiência alimentar em bovinos de corte ..	4
2.2 Fatores metabólicos e fisiológicos que afetam o CAR	7
2.3 Abordagem e estudos moleculares na eficiência alimentar	10
2.4 Associação genômica ampla.....	11
2.5 Metodologia ssGBLUP ponderado (WssGBLUP).....	16
3. OBJETIVO GERAL	17
3.1 Objetivos específicos.....	17
4. MATERIAL E MÉTODOS.....	17
4.1 Dados fenotípicos e pedigree.....	17
4.2 Estimativa do consumo alimentar residual	19
4.3 Dados genotípicos – Controle de qualidade e imputação de genótipos.....	19
4.4 Estimativa dos componentes de variância	21
4.5 Estudo de associação ampla do genoma.....	22
4.6 Anotação de genes e análise funcional	25
4.7 Diferenciação genética entre grupos divergentes para CAR.....	25
4.8 Estimativa do efeito da substituição alélica	28
5. RESULTADOS E DISCUSSÃO	29
5.1 Estrutura populacional e desequilíbrio de ligação	29
5.2 Estudo de associação genômica e enriquecimento funcional	33
5.3 Distribuição da frequência alélica, efeito da substituição alélica e índice de fixação de Wright (Fst).	45
6. CONCLUSÃO	57
7. REFERÊNCIAS	58
8. APÊNDICE.....	73

ARQUITETURA GENÉTICA DO CONSUMO ALIMENTAR RESIDUAL EM BOVINOS NELORE

RESUMO – A alimentação é o componente mais dispendioso e relevante na produção de bovinos de corte. Com a necessidade de tornar a bovinocultura de corte mais rentável e sustentável, características de eficiência alimentar ganharam importância, entretanto, apresentam dificuldades pela onerosa obtenção da informação. A característica relacionada a eficiência alimentar e com mais estudos na última década, consumo alimentar residual (CAR), possui independência fenotípica de características de crescimento e de produção. O avanço na utilização da tecnologia de marcadores moleculares, como os SNPs, possibilitou o uso de novas ferramentas genômicas. Neste sentido, é de interesse econômico e científico investigar mecanismos genéticos e biológicos que influenciam o CAR. O objetivo do estudo foi identificar regiões genômicas associadas ao CAR, avaliar diferenças genômicas em grupos contrastantes de animais, além de estimar os efeitos da substituição alélica de SNP que apresentem diferentes frequências alélicas em bovinos Nelore (*Bos indicus*). Foram utilizadas informações fenotípicas de 946 animais participantes do teste de eficiência alimentar nos anos de 2010 a 2017, com informação de genótipos de 956 animais. O modelo abordado para predição dos valores genéticos foi o WssGBLUP/S2, e a cada iteração foram reestimados efeitos de SNPs a partir de valores genético genômico (GEBVs). Foram encontradas oito regiões explicando variância genética total acima de 1%, totalizando 12,25% da variância genética total do CAR, localizadas nos cromossomos 5, 6, 9, 11, 14, 17 e 27, com 71 genes identificados, indicando existência de QTL associados ao CAR nestas regiões do genoma. A classificação de genes quanto a função biológica foi realizada pelo pacote topGO e as principais vias metabólicas associadas ao CAR foram relacionadas com função mitocondrial e metabolismo energético. O CAR apresentou estimativa de herdabilidade moderada (0,24) utilizando matriz genômica associada à matriz de parentesco (H). A partir do fenótipo ajustado, 20% dos animais presentes nos extremos da característica foram divididos em eficientes e não eficientes. Foi realizado teste exato de Fisher ($p < 0,05$) e corrigido para taxa de falso positivo (FDR) ($q < 0,05$), resultando em 104 SNPs significativos com diferenças alélicas entre os grupos. Foi calculado, para cada SNP diferenciado entre os grupos divergentes, o efeito da substituição aditiva alélica, resultando a soma para cada efeito o valor de 0,51% da variância genética aditiva para o CAR. Para identificar regiões do genoma selecionadas de forma divergente em relação aos grupos contrastantes para o CAR foi utilizado o estimador índice de fixação (Fst). O valor máximo de Fst para janelas de 100kb com 50kb de sobreposição foi de 0,025, com média de 0,022. O baixo valor de Fst demonstra baixa diferenciação genética entre os grupos, e sustenta a estrutura da população visualizada na análise de PCA. As 10 janelas com maiores valores de Fst estavam presentes nos cromossomos 4, 6, 9, 14, 23 e 25, apresentando 8 genes associados a estas regiões. Estas regiões demonstram quais genes possivelmente sofrem modificações para adaptação a seleção do CAR. A contínua seleção para o CAR provavelmente irá alterar as frequências alélicas e evidenciar as forças de seleção em diferentes regiões do genoma, diferenciando os grupos contrastantes. Os genes identificados podem fornecer informações adicionais sobre regiões QTL, devem

ser considerados em estudos de associação, genômica funcional e utilizados em programas de melhoramento genético, como ferramenta para aumentar a eficiência da seleção do CAR.

Palavras-chave: estudo de associação, frequência alélica, Fst, QTL

GENETIC ARCHITECTURE OF RESIDUAL FEED INTAKE IN NELORE CATTLE

ABSTRACT – Feeding is the most expensive and relevant component in the production of beef cattle. With the need to make beef cattle production more profitable and sustainable, feed efficiency traits have gained importance, however, they present difficulties due to the onerous obtaining of the information. The trait related to feed efficiency and with more studies in the last decade, residual feed intake (RFI), has phenotypic independence of growth and production traits. Advances in the use of molecular marker technology, such as SNPs, provide new genomic tools. In this sense, it is of economic and scientific interest to investigate genetic and biological mechanisms that influence RFI. The objective of the study was to identify genomic regions associated with RFI, and in contrasting groups of animals, to evaluate genomic differences and to estimate the effects of allelic substitution of SNPs that present different allelic frequencies in Nelore (*Bos indicus*) cattle. Phenotypic information from 946 animals participating in the feed efficiency test was used from 2010 to 2017, with information on 956 animals genotypes. The model addressed for prediction of genetic values was the WssGBLUP/S2, and at each iteration were re-estimated effects of SNPs from genomic genetic values (GEBVs). Eight regions were found explaining over 1% of the total genetic variance, and accounting for 12.25% of the total genetic variance of the RFI, located on chromosomes 5, 6, 9, 11, 14, 17 and 27, with 71 genes identified, indicating existence of QTLs associated with RFI in these regions of the genome. The classification of genes for biological function was performed by the topGO package and the main metabolic pathways associated with CAR were related to mitochondrial function and energy metabolism. The RFI presented an estimate of moderate heritability (0.24) using pedigree-genomic relationship matrix (H). From the adjusted phenotype, 20% of the animals present at the extremes of the trait were divided into efficient and non-efficient. Fisher's exact test ($p < 0.05$) and corrected for false positive rate (FDR) ($q < 0.05$) were performed, resulting in 104 significant SNPs with allelic differences between groups. The effect of the allelic additive substitution was calculated for each differentiated SNP between the divergent groups, resulting in the sum of 0.51% of the additive genetic variance for the RFI for each effect. In order to identify regions of the genome that were differently selected in relation to the groups contrasting for CAR, the fixation index (Fst) was used. The maximum value of Fst for 100kb windows with 50kb of overlap was 0.025, with an average of 0.022. The low Fst value demonstrates low genetic differentiation between the groups, and supports the population structure visualized in the PCA analysis. The 10 windows with higher Fst values were present in chromosomes 4, 6, 9, 14, 23 and 25, presenting 8 genes associated to these regions. These regions demonstrate which genes may undergo modifications to fit RFI selection. The continuous selection for the RFI probably will alter the allelic frequencies and evidence the selection forces in different regions of the genome, differentiating the contrasting groups. The identified genes can provide additional information on QTL regions, should be considered in association studies, functional genomics and used in breeding programs as a tool to increase the efficiency of RFI selection.

Keywords: association study, allele frequency, Fst, QTL

1. INTRODUÇÃO

O componente mais dispendioso e relevante na produção de bovinos de corte é a alimentação, representando até 80% do custo total, dependendo do sistema de produção (ARTHUR; ARCHER; HERD, 2004; LOWE et al., 2009). Considerando o aumento da demanda da produção, aliada a questões ambientais e custos com alimentação, o bovino eficiente torna-se cada vez mais desejável.

A eficiência alimentar é importante característica, capaz de identificar animais com menor consumo alimentar produzindo mesma quantidade de carne. A seleção de animais eficientes otimiza a produção de bovinos de corte, aumenta a rentabilidade e sustentabilidade ao reduzir o impacto ambiental, diminuindo uso de recursos naturais, gases de efeito estufa e de dejetos sólidos, sem comprometer o desempenho animal (ARTHUR; HERD, 2008; HEGARTY et al., 2007; NKRUMAH et al., 2006).

Vários índices estimam a eficiência alimentar, o CAR foi um índice proposto por Koch et al. (1963), definido como a diferença entre o consumo observado e o consumo predito, baseado no ganho médio diário (GMD) e peso metabólico ($PV^{0,75}$) dos animais. O resíduo da regressão da predição determina o valor do CAR e sua forma de calcular atribui independência fenotípica de características de crescimento, como taxa de crescimento e tamanho corporal, representadas respectivamente pelo ganho médio diário (GMD) e peso vivo metabólico ($PV^{0,75}$) (HERD; ARTHUR, 2009). Assim, animais negativos para o CAR são mais eficientes, apresentando o consumo observado menor do que o predito para o ganho observado, e os animais menos eficientes tem CAR positivo, caracterizando consumo observado maior do que o predito.

O potencial econômico da característica nos EUA foi estimado e constatado que a redução de 10% do CAR, equivalente a redução de 0,91kg de MS/dia do consumo diário por animal, reduz mais de um bilhão de dólares por ano (WEABER et al., 2013). A utilização das características de eficiência alimentar possui dificuldades de obtenção da informação para serem utilizadas no processo de seleção (ROLF et al., 2011). Em bovinos Nelore (*Bos taurus indicus*), há escasso conhecimento das estimativas de componentes de variância, parâmetros genéticos e correlações genéticas com outras características de importância

econômica e poucos estudos genômicos relatados na raça (SANTANA et al., 2014).

O sequenciamento do genoma bovino possibilitou o desenvolvimento de painéis comerciais com grande número de marcadores moleculares do tipo SNP, permitindo capturar a variação genética e detecção de *loci* de características quantitativas (QTL – do inglês *Quantitative trait locus*) de características complexas por meio de estudos de associação genômica ampla (GWAS – do inglês *Genome Wide Association Studies*) (ELSIK; TELLAM; WORLEY, 2009; HAYES et al., 2007). Os estudos de GWAS tem como base o desequilíbrio de ligação (LD) entre marcador e a variante causal com frequência suficiente para detecção, pressupondo que pelo menos um SNP esteja em LD com a variante associada ao fenótipo (CANTOR; LANGE; SINSHEIMER, 2010)

As diferenças nas frequências alélicas e extensão do desequilíbrio de ligação dos marcadores entre taurinos (*Bos taurus*) e zebuínos refletem, consequentemente, diferentes efeitos de marcadores identificados entre as raças (TIZIOTO et al., 2013). A montagem do genoma e grande maioria dos GWAS foram realizados em animais taurinos, portanto, identificar regiões genômicas e genes do CAR em zebuínos, principalmente na raça Nelore, contribui com informações genômicas especificamente da raça de maior representatividade nacional. A integração de informação molecular a métodos tradicionais de seleção diminui o intervalo de gerações, aumenta acurácia e incrementa a resposta de seleção, surgindo como auxílio aos programas de melhoramento genético animal (DEKKERS, 2007).

A comparação das frequências alélicas entre populações selecionadas para diferentes finalidades fornece indícios de regiões no genoma sujeitas a diferenciação genética. A seleção artificial através da domesticação resultou em aumento da diversidade entre recentes populações bovinas. Esta diversidade inclui variação na morfologia, fisiologia, produção e fertilidade (CHEN et al., 2016). Avaliar os diferentes padrões de diferenciação genética tem aplicabilidade em estudos de domesticação, formação de raças, estrutura populacional e consequências de seleção (KIJAS et al., 2012). Porém, além da seleção, a divergência genética nas frequências alélicas pode ter origem devido à deriva

genética e endogamia. Entretanto, a endogamia deve afetar todos *loci* igualmente, e a deriva genética deve afetar os *loci* de forma aleatória, não originando LD entre *loci* adjacentes, que é o esperado em seleção artificial (MACEACHERN et al., 2009).

Para detectar alterações ocasionadas por seleção em diferentes regiões do genoma, diferentes métodos foram propostos, e um deles é o índice de fixação (F_{st}), que é medida proposta capaz de identificar diferenciação genética (WEIR; CAROLINA; HILL, 2002; WRIGHT, 1950)

Estudos de GWAS são utilizados para identificação de QTL para características de importância produtiva (HAWKEN et al., 2012; MACEACHERN et al., 2009). Dessa forma, o GWAS informa relevantes descobertas sobre a arquitetura genética de características complexas e oferece oportunidade para análises posteriores. Análises como a distribuição da frequência alélica, diferenciação genética entre grupos divergentes e o efeito da substituição dos SNPs fornecem subsídios adicionais para entender a base biológica da característica e a estrutura populacional (ALLEN et al., 2010).

A motivação do trabalho foi identificar genes de acordo com suas funções biológicas que influenciam o consumo alimentar residual, avaliar diferenças genômicas que ocorrem entre grupos contrastantes e a contribuição para variância genética em grupos divergentes do CAR, fornecendo informações sobre a arquitetura genética da característica em animais contrastantes. Diferenças encontradas na estrutura populacional podem auxiliar estudos de sequenciamento, mapeamento fino em busca de variantes causais e estimativas acuradas dos efeitos de cada marcador.

2. REVISÃO DE LITERATURA

2.1 Considerações gerais acerca da eficiência alimentar em bovinos de corte

O Brasil, segundo maior produtor e maior exportador de carne bovina, destaca-se em âmbito internacional na cadeia produtiva de carne, com mais de 226 milhões de cabeças (USDA, 2017). A origem do rebanho de corte nacional possui 87,5% de composição zebuína e destes, 78% possui a raça Nelore na composição genética (ROSA; MENEZES; SULEIMAN, 2016). O componente mais dispendioso e relevante no processo de produção de bovinos de corte é a alimentação, representando até 80% do custo total, dependendo do sistema de produção (ARTHUR; ARCHER; HERD, 2004; LOWE et al., 2009). Assim, o estudo da associação da ingestão de alimentos e produção geraram diversas medidas de eficiência alimentar.

A eficiência alimentar é importante característica, capaz de identificar animais com menor consumo alimentar produzindo mesma quantidade de carne. Deste modo, esta característica otimiza a produção na pecuária de corte, aumentando a rentabilidade e sustentabilidade ao reduzir o impacto ambiental diminuindo o uso de recursos naturais, de gases de efeito estufa e dos dejetos sólidos (ARTHUR; HERD, 2008; HEGARTY et al., 2007; NKRUMAH et al., 2006).

O CAR foi um índice proposto por Koch et al. (1963), definido como a diferença entre o consumo observado e o consumo predito, baseado no ganho médio diário (GMD) e peso metabólico ($PV^{0,75}$) dos animais. O resíduo da regressão da predição determina o valor do CAR e sua forma de calcular atribui independência fenotípica de características de crescimento, como taxa de crescimento e tamanho corporal, representadas respectivamente pelo GMD e $PV^{0,75}$ (HERD; ARTHUR, 2009).

Animais mais eficientes são classificados como baixo CAR (negativo) em que consomem menos alimento do que o esperado (BASARAB et al., 2003). A utilização de animais negativos reduz os custos de produção, bem como diminuem o impacto ambiental, reduzindo de 15 a 30% da produção de metano (NKRUMAH et al., 2006) e de 15 a 20% da produção de estrume (HEGARTY et al., 2007). Logo, ao selecionar animais com menores valores genéticos ou

fenotípicos de CAR, melhora simultaneamente a eficiência alimentar e não aumenta o valor genético de características de crescimento, atribuindo outra vantagem ao critério de seleção (DEL CLARO; MERCADANTE; SILVA, 2012).

Dentre as características que mensuram a eficiência alimentar, muitas expressam razão de duas características, por outro lado, o CAR é combinação linear composto de outras características. Características representadas por uma razão possuem pressão de seleção desigual devido as diferenças da variação genética dos componentes, causando resposta a seleção imprevisível. Segundo Gunsett (1984), a abordagem do índice linear obteve melhor resposta a seleção comparada a seleção direta para uma razão composta pelos componentes do índice.

Para determinar o CAR o *Beef Improvement Federation Guidelines* (BIF, 2010) sugere 21 dias de adaptação seguidos por 70 dias de teste, utilizando cochos automatizados. A mensuração dos testes de eficiência alimentar para o CAR divergem em diferentes períodos, com autores sugerindo de 56 dias (CULBERTSON et al., 2015) 63 dias (WANG et al., 2006) e 70 dias (ARCHER et al., 1997). Segundo Manafiazar et al. (2017), o CAR requer testes de eficiência alimentar de maior duração em relação ao consumo alimentar (42 dias) devido a necessidade de mensurar com acurácia o crescimento dos animais. As diferenças podem ser explicadas pelo sistema dos testes, pelo tipo de coleta do consumo alimentar, se é realizado individual ou com equipamentos automatizados e pela frequência das pesagens dos animais, o qual se aplicado, tende a obter maior acurácia nos resultados dos testes mesmo com menor período de duração.

Testes individuais são onerosos, trabalhosos e com maior período de duração, enquanto os realizados com equipamentos automatizados diminui o tempo do teste, enquanto as instalações para mensurar o consumo alimentar são limitadas, além de suportar pequeno número de animais por teste (CULBERTSON et al., 2015; NIELSEN et al., 2013). Logo, reduzir o tempo dos testes permite testar mais animais, além de diminuir os custos da produção por teste, permitindo maior intensidade de seleção. As características de eficiência alimentar envolvem alto custo devido a necessidade de infraestrutura especializada para coleta de dados fenotípicos, dificultando que sejam adotadas em massa pela indústria da carne e

incluídas nos índices de seleção (MANAFIAZAR et al., 2017).

Em bovinos Nelore (*Bos indicus*) há escasso conhecimento das estimativas de componentes de variância, parâmetros genéticos e correlações genéticas com outras características (SANTANA et al., 2014). Apesar do desvio-padrão para o CAR variar entre raças e apresentar valores entre 0,66 kg/dia a 0,74 kg/dia– (Basarab et al., 2003; Arthur, Renand e Krauss, 2001), a variabilidade fenotípica para o Nelore demonstra valores similares aos do *Bos taurus*, com valores de 0,28, 0,31 e 0,69 kg/dia (Castilhos et al., 2010; Sobrinho et al., 2011 e Gomes et al., 2012). Segundo Arthur e Herd (2008), progênies de touros com CAR 26% menor e consumo de matéria seca de -1.00kg/dia obtiveram ganho de peso 19% maior sem aumento na ingestão de alimentos, e conversão alimentar 41% menor.

Em revisão considerando 22 trabalhos científicos com uso de meta-análise, Del Claro; Mercadante; Silva, (2012) encontraram valor de herdabilidade para CAR e características componentes, sendo consumo alimentar residual (CMS), ganho médio diário (GMD) e peso vivo metabólico ($PV^{0,75}$) de $0,298\pm 0,127$, $0,411\pm 0,131$, $0,353\pm 0,127$ e $0,430\pm 0,114$, respectivamente, e descreveram que cerca de 67% da variação da estimativa de herdabilidade foi explicada pelos fatores sexo, país e raça. Logo, parte das variações observadas são devido a ação de genes aditivos, com possível manipulação da característica. Arthur; Renand; Krauss. (2001) observaram variabilidade fenotípica do CAR, indicando sua utilização como critério de seleção, sendo que as estimativas de herdabilidade apresentados na literatura para o CAR variaram entre 0,28 e 0,45 (ARTHUR; RENAND; KRAUSS, 2001; CROWLEY et al., 2010; KOCH et al., 1963).

A seleção do CAR em programas de melhoramento se apresenta como um desafio, por ser de difícil mensuração e de custo elevado, necessitando também de equipamentos específicos para a coleta rotineira dos dados. Mas, ao considerar o valor econômico, a melhoraria de 5% na eficiência alimentar tem cerca de quatro vezes o impacto econômico como melhoria de 5% do GMD, significando que os animais ganham peso mais rápido, conseqüentemente são vendidos mais cedo, aumentando a rotatividade do sistema de produção e justificando o investimento da característica (Gibb e McAllister, 1999).

A aplicação do CAR nos programas de melhoramento com emprego da

informação genômica auxiliam a melhor compreensão e identificação de regiões e os mecanismos biológicos, principalmente na raça Nelore, de maior representatividade no Brasil (SANTANA et al., 2014; OLIVIERI et al., 2016).

2.2 Fatores metabólicos e fisiológicos que afetam o CAR

Os fatores biológicos que influenciam o CAR são: diferenças na digestão dos alimentos, composição corporal e atividade física dos animais, sendo mais da metade desta diferença devida a outros processos fisiológicos (HERD; ODDY; RICHARDSON, 2004). Os mecanismos fisiológicos que contribuem para a variação consumo alimentar residual encontradas por Herd e Arthur (2009) foram: *turnover* protéico e metabolismo do tecido e estresse (37%); digestibilidade (10%); incremento de calor e de fermentação (9%); atividade física (9%); composição corporal (5%); padrões de alimentação (2%) e outros mecanismos, como transporte iônico (27%). Em GWAS para CAR, Barendse *et al.* (2007) encontraram 161 SNPs evidenciando associação de genes em diferentes processos metabólicos, incluindo todos processos descritos acima.

A variação nos processos fisiológicos que contribuem para a síntese e degradação de proteínas e metabolismo de gordura podem resultar em diferentes gastos de energia (RICHARDSON; HERD, 2004). Os custos energéticos do *turnover* protéico contribui em cerca de 20 a 25% do gasto energético de manutenção (REEDS et al., 1998). Desse modo, os principais processos fisiológicos influenciando o CAR estão relacionados com exigência de manutenção e transporte de íons, além dos diferentes fatores metabólicos que interferem na variação (ARTHUR; HERD, 2008; RICHARDSON et al., 2004). A regulação fisiológica é controlada por diferentes genes, assim a compreensão dos mecanismos fisiológicos relacionados diretamente ao animal é fundamental para compreender as causas das variações da eficiência alimentar (BARENDSE et al., 2007; MOORE; MUJIBI; SHERMAN, 2009).

Richardson e Herd (2004) constataram que animais menos eficientes permanecem 5% mais tempo se alimentando e em pé, exigindo maior consumo dos animais. Em relação ao CMS, os mesmos autores encontraram correlação

negativa entre digestibilidade da matéria seca e CAR, podendo associar com a menor taxa de passagem dos alimentos, originando maior disponibilidade de nutrientes, ocasionando redução nas exigências de manutenção. Com resultado similar, Nkrumah et al. (2006) encontraram menor tempo de consumo em animais baixo CAR e também menor número de visitas quando comparados com novilhos de alto CAR.

Richardson et al. (2002) estudaram a associação dos parâmetros sanguíneos com o CAR em animais taurinos, e obtiveram diferenças consistentes nos indivíduos menos eficientes, que demandaram maior transporte de oxigênio e são menos adaptados ao confinamento. Os autores também avaliaram o estresse no transporte e não houve diferenças nos parâmetros das células vermelhas, mas as células brancas apresentaram valores diferentes antes e depois do transporte, com relação negativa com o CAR, indicando maior contagem de células brancas em animais mais eficientes.

Nos resultados quanto a estimativa dos metabólitos sanguíneos, Richardson et al. (2004), encontraram maiores níveis de insulina, cortisol e ureia em animais menos eficientes, em contrapartida, triglicerídeos em menores níveis, explicado como resposta relacionada a mudança na composição corporal, reciclagem tecidual e eficiência na utilização de nutrientes, além de diferenças entre os animais quanto a resposta ao estresse causado durante o transporte.

A leptina regula o consumo alimentar e a composição corporal, possuindo relação com o CAR, dado que menores concentrações plasmáticas da leptina aumenta a eficiência alimentar (FARJALLA, 2009; SUZUKI et al., 2009). A leptina atua em todo metabolismo, atuando no hipotálamo e ocasionando redução da ingestão alimentar e aumento do gasto energético (BARTHA; SAYED-AHMED; RUDAS, 2005). O efeito da leptina na regulação do consumo alimentar se deve a este hormônio regular o balanço energético e estar correlacionado ao apetite, que interfere no consumo e influencia a formação do tecido adiposo (ZIEBA; AMSTALDEN; WILLIAMS, 2005). Outros hormônios e fatores de crescimento regulam a utilização de nutrientes no metabolismo, como insulina, glucagon, hormônios do eixo somatotrópico e hormônios da tireoide são considerados decisivos na taxa de metabolismo basal (BELLMANN et al., 2004).

A eficiência dos animais possui relação com a exigência de manutenção, que varia de 10 a 12% em bovinos de corte e é afetada pela produção de calor (HOTOVY et al., 1991; SWANSON; MILLER, 2008). Desta forma, a produção de calor tem efeito sobre a eficiência alimentar de bovinos. Segundo Nkrumah et al. (2006), em estudo de calorimetria indireta observou que animais baixo CAR produziram 21% menos calor, tanto quanto menor produção de metano (HEGARTY et al., 2007). Tanto a produção de calor quanto taxas metabólicas já foram associadas a diferenças na síntese protéica (WEBSTER, 1980). Segundo Herd e Arthur (2009), o tecido adiposo tem maior eficiência comparado com o tecido muscular na deposição e uso de nutrientes, esta diferença ocorre devido o *turnover* protéico apresentar maior variabilidade entre os órgãos, o que não ocorre no *turnover* de lipídeos.

A maioria dos tecidos viscerais como trato gastrointestinal e fígado passam por processos de síntese e degradação protéica energeticamente dispendiosos, com alta taxa de *turnover* protéico, alta demanda de aminoácidos, alto fluxo sanguíneo e alto consumo de oxigênio (SEAL; REYNOLDS, 1993). Esta síntese e degradação protéica possui alto custo energético no metabolismo dos animais de produção e representa até 30% do gasto energético com manutenção (BALDWIN; SAINZ, 1995). Richardson et al. (2001) evidenciaram correlações negativas entre CAR e percentagem de proteína corporal, descrevendo maior eficiência dos animais que conseguem depositar proteína com maior eficácia ou com menor taxa de degradação.

Compreender processos metabólicos como a utilização de energia e nutrientes, digestibilidade, metabolismo basal e atividade favorece a identificação de possíveis biomarcadores preditores de eficiência alimentar. Segundo Meale *et al.* (2017), a eficiência alimentar está relacionada a componentes metabólicos, porém identificar um marcador biológico continua não conclusivo devido sua complexidade, apesar de encontrar diferenças biológicas em grupos divergentes para o CAR.

O CAR é caracterizado por complexas interações entre constituição celular e proteínas, sendo influenciada por diversos processos biológicos (BARABÁSI; OLTVAI, 2004). Desse modo, é necessário abordar os mecanismos biológicos

que originam diferenças fenotípicas para entender os processos biológicos associados ao CAR (KARISA; MOORE; PLASTOW, 2013). A eficiência alimentar é afetada por muitas redes moleculares e características fisiológicas. Desse modo, obter o conhecimento dessas funções e interações genéticas proporcionam melhor conhecimento da relação dos genes e a anotação funcional destes fornecem melhor compreensão de mecanismos genômicos e componentes regulatórios que afetam características de difícil mensuração, como o CAR (SNELLING et al., 2013).

2.3 Abordagem e estudos moleculares na eficiência alimentar

Os fatores genéticos, fisiológicos e ambientais interferem na produção de bovinos. Estudos em genética e biologia molecular possibilitaram a identificação de marcadores moleculares que podem estar relacionados a diferenças fenotípicas entre os animais. A cada geração variações na molécula de DNA podem ocorrer, sendo que estas podem ser do tipo mutação, como as inserções e deleções (InDels) e polimorfismos de nucleotídeo único (SNP – do inglês *Single Nucleotide Polymorphism*) que são o tipo mais abundante de variação no genoma, ou ainda, variações em regiões repetitivas em tandem (minissatélites e microssatélites) (BROOKES, 1999). A frequência das mutações é afetada por seleção, recombinações homólogas, deriva gênica e fatores exógenos. Caso atinjam frequência mínima de 1% ($MAF \geq 1\%$) na população são denominados polimorfismos, e caso tenha alterado apenas um nucleotídeo, são chamados de SNPs (KWOK; GU, 1999). Os marcadores do tipo SNP são alterações elementares e podem ser associados a *loci* de características quantitativas (QTL) de características de produção e diferenciam dois ou mais indivíduos.

Os marcadores são utilizados devido à dificuldade de demarcar com precisão o QTL, permitindo identificar a presença ou posição de um gene de interesse. Se os efeitos dos marcadores são estimados com acurácia, o uso de informação genômica torna-se estratégia para melhorar a seleção de características como CMS e CAR (SANTANA et al., 2014). Ao estabelecer a relação entre polimorfismos e a característica se estima o valor genético

individual, fundamentado no valor genômico, independente de avaliação fenotípica (MEUWISSEN; HAYES; GODDARD, 2001).

Esse avanço contribuiu para o melhor conhecimento do genoma e das regiões que influenciam determinadas características, principalmente de genes candidatos. A integração de informação molecular por meio de uso de painéis, em vez de métodos tradicionais de seleção, diminui o intervalo de gerações, aumenta a acurácia e incrementa a resposta de seleção, surgindo como alternativa aos programas de melhoramento genético animal (DEKKERS, 2007).

2.4 Associação genômica ampla

O genoma bovino (*Bos taurus*) UMD 3.1.1 disponível em “National Center for Biotechnology Information” (NCBI) abrange 2,67 bilhões de pares de bases anotados, distribuídos entre cromossomos autossômicos e cromossomo X (ZIMIN et al., 2009). O sequenciamento do genoma permitiu identificar milhares de SNPs, além de desenvolver arranjos de SNPs que caracterizam o genoma e desse modo, realizar estudos de associação em grande escala, pressupondo-se que o SNP associado ao fenótipo esteja em LD com o QTL (DAETWYLER et al., 2014).

Os SNPs se tornaram eficazes para estudos de GWAS pela abundância no genoma e facilidade de genotipagem que apresentam (GODDARD; HAYES, 2009). Com estudos de associação genômica é possível localizar genes ou regiões genômicas que estejam influenciando o fenótipo estudado (ELSIK; TELLAM; WORLEY, 2009; HAYES et al., 2007). Os estudos de GWAS permitem a identificação de associações estatísticas entre milhares de *loci* genômicos e características poligênicas, aumentando o entendimento de vias moleculares e biológicas da característica.

O desenvolvimento de painéis densos de marcadores SNP permitiu análise de milhares de SNPs simultaneamente, possibilitando análises genômicas de estrutura genética, da relação entre perda de alelos e aumento de endogamia, estimação do grau de diversidade e divergência genética dentro e entre populações e identificação de regiões sujeitas a seleção (BARENDSE et al., 2009; ZENGER et al., 2006). Em geral, estes processos são consequências da associação ou ligação entre dois *loci* em função da proximidade física (ArLDie,

Kruglyak e Seielstad, 2002; Bolormaa *et al.*, 2011). O LD é medida estatística da associação não aleatória entre alelos de diferentes *loci*, que ocorre pela proximidade entre diferentes *loci*, sendo propriedade de marcadores em determinada região adjacente que explica a proporção em que um alelo de um SNP é herdado ou correlacionado com um alelo de outro SNP na população. Assim, o LD descreve cenário onde algumas combinações de alelos ou marcadores acontecem com maior ou menor frequência, portanto, a origem do LD reside nas frequências alélicas e fatores que o alteram, como seleção, recombinação e mutação, endogamia e a distância entre os *loci*, em que quanto mais próximo o SNP estiver do QTL, maior a ligação.

Pode-se estimar o LD a partir da frequência alélica pela medida r^2 , a qual representa o quadrado da correlação entre dois *loci*, sendo preferido em estudos de associação genômica, considerando a existência de relação inversa entre r^2 e o tamanho da amostra necessária para identificação da associação entre marcador e QTL (PRITCHARD; PRZEWORSKI, 2001). Os valores de r^2 variam de zero a um, indicando respectivamente, ausência ou completo LD, o qual valores acima de 0,2 são considerados eficazes para estudos genômicos (HAYES *et al.*, 2007).

As predições genômicas dependem do conhecimento do LD entre marcadores moleculares e variáveis causais desconhecidas na população estudada, visto que que indicam associação entre marcadores e QTL (SNELLING *et al.*, 2013). O elevado LD em determinada região indica que os genes desta região são herdados em blocos, ou seja, menor chance de recombinação e possuem frequências alélicas coincidentes.

O estudo de associação genômica ampla é delineamento experimental validado para identificar associações entre *loci*, responsáveis pela variação genética de determinado fenótipo em uma população (VISSCHER *et al.*, 2017). O GWAS explora o LD entre marcador e variante causal a nível da população e considera todos marcadores nas análises estatísticas, pressupondo que pelo menos um SNP esteja em LD com genes ou regiões genômicas associadas com o fenótipo e que esteja em frequência suficiente para identificação (CANTOR; LANGE; SINSHEIMER, 2010). A eficácia do GWAS depende de quantos *loci*

afetam a característica, a distribuição, tamanho e frequência alélica em determinado *loci*, tamanho da população, densidade do painel utilizado para genotipagem (quanto maior densidade do painel, maior probabilidade do QTL estar em LD com algum marcador) e o nível de heterogeneidade da característica analisada (VISSCHER et al., 2017).

A presença do LD em estudos de associação genômica permite inferir sobre associação direta e a associação indireta. Os SNPs são associados diretamente ao fenótipo quando o SNP é a mutação causal que modifica a regulação ou expressão do gene, ou associados indiretamente, quando o marcador está ligado a mutação causal por LD, o qual ocorre em estudos de associação genômica (BUSH; MOORE, 2012). O SNP associado ao QTL pode significar tanto um efeito direto ao fenótipo quanto uma ligação ao SNP causal, o qual deve-se investigar posteriormente os resultados encontrados. Apesar da associação direta não ser observada, o GWAS pode indicar SNPs que estejam em forte LD com a variante causal por meio do alto r^2 entre dois *loci*, considerando que a medida r^2 descreve a proporção da variância do QTL que seria explicada pelo marcador, se este fosse verdadeiramente um QTL (BALDING, 2006; HILL; ROBERTSON, 1968).

O uso de painéis contendo número significativo de marcadores possibilitou o uso de GWAS para identificação da base genética de características complexas, ao identificar SNPs em LD com genes ou QTL associados a característica de interesse. Em características complexas, muitas regiões contribuem para a variação genética, de modo que muitos marcadores influenciam o fenótipo e assim cada variante explica pequena proporção de variância, o que implica diretamente a necessidade de maiores tamanhos da população para identificá-las. No estudo de associação se pressupõe as diferenças das frequências alélicas na população, o qual implica associar um marcador ao SNP causal da característica, de acordo com aumento na frequência de um alelo ou genótipo, evidenciando diferenças genéticas na população.

Uma das vantagens do GWAS justamente não exigir conhecimento prévio profundo dos processos biológicos, fisiológicos e bioquímicos relacionados ao fenótipo testado (WANG et al., 2005). Outra vantagem do GWAS é adquirir a

arquitetura genética da característica, isto significa que avaliando os indivíduos, cada um possui números de alelos que aumentam e outros que diminuem a expressão da característica. Assim, marcadores que explicam 1% da variação fenotípica já se tornam efetivos para elucidação das bases biológicas da característica (CANTOR; LANGE; SINSHEIMER, 2010). Neste contexto, a identificação de genes, proteínas e vias metabólicas relacionado na manifestação fenotípica auxilia na compreensão dos mecanismos genéticos que controlam características de interesse.

Em gado de corte, diversos trabalhos têm identificado variantes genéticas associadas com característica de eficiência alimentar. Barendse *et al.* (2007) publicaram o primeiro estudo utilizando associação genômica ampla (GWAS) em bovinos e identificaram 161 SNPs, sendo apenas 20 SNPs associados com características de eficiência alimentar em diferentes raças taurinas, explicando cerca de 76% da variação genética do CAR quando consideradas todas raças. Isto mostra que poucos SNPs foram associados a característica, mas muitos contribuem para a variância da mesma. Sherman *et al.* (2008) identificaram SNPs que explicaram 6,9% da variação do CAR fenotípico e 5,7% da variação do CAR genético nos cromossomos 2, 5, 10, 20 e 29, sendo que apenas um SNP coincidiu com o estudo anterior citado, mas sem significância para o CAR neste estudo. Rolf *et al.* (2012) identificaram 66 SNPs associados ao CAR em animais da raça Angus. Nota-se que em uma mesma espécie resultados diferentes são obtidos devido a população a qual pertencem e do painel de SNPs utilizado.

Considerando modelos que inclui ou não inclui efeitos poligênicos, Serão *et al.* (2011) encontraram respectivamente 69 e 131 SNPs associados ao CAR, utilizando Angus, Simental e seus cruzamentos. No entanto, estudos de GWAS identificam regiões genômicas divergentes, demonstrando que efeitos de QTL podem ser diferentes nas raças e populações, ou ainda, o QTL pode não segregar entre as raças (HAYES *et al.*, 2009a). Em animais Nelore, Santana *et al.* (2014) utilizando dois painéis de SNPs de densidades diferentes, obtiveram dois SNPs associados ao CAR, próximos de genes controladores de apetite, localizados nos cromossomos 8 e 21 próximos a dois QTL associados a característica. Também em estudo com Nelore, Olivieri *et al.* (2016) obtiveram 12 janelas de regiões

genômicas que juntas explicaram mais de 19% da variância genética aditiva do CAR, com grande parte dos genes mapeados nessas regiões relacionados com o transporte de íons. Este sistema em ruminantes é associado por consumir aproximadamente 10% da energia gasta, portanto, animais que possuem menor gasto de energia na manutenção para uso neste sistema, encaminhará energia para outros processos, influenciando a eficiência alimentar (SANTANA et al., 2014).

A compreensão dos mecanismos biológicos, por meio da avaliação das diferenças na expressão dos genes que afetam o CAR, podem ajudar na explicação das bases genéticas e possibilitar seus resultados na utilização da seleção genômica. Tizioto et al. (2015) e Tizioto et al. (2016) estimaram o mérito genético aditivo para o CAR em animais Nelore e selecionaram animais com fenótipo divergente, utilizaram dados RNA-seq de biópsias para fígado e músculo *longissimus dorsi*, e descreveram a expressão gênica diferencial para os grupos contrastantes. Estes mesmo autores, encontraram 112 e 73 genes diferencialmente expressos em fígado e *longissimus dorsi*, respectivamente, e relataram mecanismos biológicos comuns detectados previamente em estudos de GWAS.

Oliveira et al. (2014) analisaram em bovinos Nelore, 9 características de eficiência alimentar e identificaram 36 regiões genômicas associadas, sendo que apenas alguns destes QTL apresentaram sobreposição de regiões previamente relatadas em *Bos taurus*, indicando que diferentes regiões e genes atuam no controle da eficiência alimentar nas duas espécies.

A limitação de poucos dados fenótipos identificados dificulta abordagem de seleção para a eficiência alimentar, sendo ainda escassos trabalhos relacionando a caracterização genética e genômica das medidas de eficiência alimentar em bovinos, especialmente na raça Nelore (SANTANA et al., 2014). A maioria dos estudos de associação com característica de eficiência alimentar são realizadas em animais *Bos taurus*, conseqüentemente identificar regiões genômicas e genes em *Bos indicus* é de grande importância e justificável, contribuindo para melhoramento desta raça. Assim, compreender vias moleculares que interfiram em mudanças metabólicas e fisiológicos que regulam a eficiência alimentar é essencial, uma vez que divergências nestes fatores podem ser explicados, em

parte, pelas variações encontradas no genoma dos animais, diferenciando-os em eficientes e ineficientes.

2.5 Metodologia ssGBLUP ponderado (WssGBLUP)

A metodologia proposta por Misztal, Legarra e Aguilar (2009) denominada *single-step genomic BLUP* (ssGBLUP) é capaz de prever os valores genéticos genômicos integrando informações genômicas, fenotípicas e de pedigree, podendo ser usada por qualquer modelo BLUP para obtenção de GEBVs, possibilitando maior abrangência de uso. As informações são integradas em um único procedimento denominado *single-step* e é uma modificação do BLUP (*Best Linear Unbiased Predictor*) com substituição da matriz de parentesco tradicional (A) pela matriz que combina informações de pedigree e genômica (H). É um método com aplicabilidade mais simples comparado a métodos *multi-step* e podem ser mais precisos por ponderar adequadamente informações de animais genotipados, evitando dupla contagem de informações fenotípicas devido ao pedigree e/ou próprios dados (LOURENCO et al., 2015; VANRADEN, 2008). Essa metodologia utiliza informações dos fenótipos observados como variável resposta (MISZTAL et al., 2013).

O ssGBLUP possui modelo infinitesimal, assumindo variância igual para todos efeitos de SNPs, limitando o uso da metodologia por não representar a real situação das características complexas. Para superar essa limitação, Wang *et al.* (2012) propuseram o ssGBLUP ponderado (*weighted single-step genomic BLUP* - WssGBLUP), atribuindo ponderações para os marcadores genéticos em processos iterativo, e, a partir disso, obter GEBVs e os efeitos individuais dos marcadores. Nesse método, o GEBV é convertido em soluções de SNPs e usado para estimar as variâncias dos marcadores, e estas usadas para criar matriz G ponderada (LOURENCO et al., 2014). Ao ponderar pesos e conseqüentemente, ponderações diferentes para os marcadores, o processo obtém ganhos na detecção de regiões genômicas ao evidenciar presença de QTL de maior ou menor efeito de acordo com sua importância para característica estudada.

3. OBJETIVO GERAL

Realizar associação genômica ampla e avaliar as diferenças genômicas entre grupos contrastantes que ocorrem como resultado da seleção para consumo alimentar residual em animais Nelore.

3.1 Objetivos específicos

- Prospecção de regiões genômicas e genes associados ao consumo alimentar residual;
- Identificar genes de acordo com suas funções metabólicas e biológicas que influenciam o consumo alimentar residual;
- Estimar os efeitos da substituição alélica de SNPs que apresentarem frequências alélicas diferentes entre grupos contrastantes.
- Aplicar índice de fixação (F_{st}) para identificação de regiões sujeitas a diferenciação genética entre grupos contrastantes.

4. MATERIAL E MÉTODOS

4.1 Dados fenotípicos e pedigree

Os dados foram disponibilizados pelo programa de melhoramento genético Nelore Qualitas®, referente a 946 touros Nelore, provenientes de rebanhos localizados nas regiões sudeste, centro-oeste e norte do Brasil criados em sistema de produção extensivo. Os dados consistem dos testes de eficiência alimentar efetuados entre os anos de 2010 e 2017, sendo até 2015 executados no confinamento experimental de bovinos de corte da EVZ/UFG em sistema de baias individuais, e nos anos de 2016 e 2017 no Centro de inovação em genética e nutrição animal – CIGNA da Unesp/Botucatu por meio de avaliação individual automatizada utilizando o sistema eletrônico da empresa Intergado® em cochos coletivos.

Os touros ingressaram no teste com idade média de $21,7 \pm 1,3$ meses, permanecendo 21 dias em adaptação e 56 dias em prova propriamente dita. Os animais foram pesados ao início da adaptação e início, meio (pesagem intermediária) e fim do teste. Os touros participantes do teste foram escolhidos com base no índice Qualitas dentro de cada safra, integrado por 40% de características para produção de matrizes, composto pelo peso a desmama e perímetro escrotal, e 60% de características de maior relevância para produção de carne a pasto, incluindo musculosidade e ganho pós-desmama. O pedigree considerou animais relacionados aos animais com fenótipo até a sexta geração, totalizando 4.682 animais.

Análise dos dados foi realizada retirando animais que não adaptaram, apresentaram problemas locomotores ou sanitários. Quanto aos dados de consumo, foram mantidos dados de parâmetros produtivos, fisiológicos e comportamentais do rebanho processados no sistema. A descrição dos testes de eficiência alimentar é apresentada na Tabela 1.

Tabela 1. Número de animais por ano de teste seguido pela média e desvio-padrão (DP) do peso inicial, idade e consumo de matéria seca (CMS)

Ano teste	Número de animais (n)	Peso inicial \pm DP (kg)	Idade \pm DP (dias)	CMS \pm DP (kg MS/dia)
2010	118	483,22 \pm 37,17	705,11 \pm 29,64	9,57 \pm 0,87
2011	120	393,46 \pm 37,87	633,10 \pm 30,17	10,32 \pm 1,18
2012	120	398,53 \pm 47,41	632,40 \pm 32,57	10,56 \pm 1,14
2013	119	435,38 \pm 53,11	641,05 \pm 31,01	10,42 \pm 1,22
2014	119	510,42 \pm 74,26	651,29 \pm 26,67	12,37 \pm 1,27
2015	119	483,58 \pm 67,90	665,21 \pm 34,18	12,08 \pm 1,22
2016	118	474,25 \pm 58,74	641,37 \pm 32,13	11,69 \pm 1,15
2017	113	469,13 \pm 61,48	646,76 \pm 37,38	11,35 \pm 1,23
Média Geral	118,25	455,75 \pm 68,69	651,98 \pm 38,81	11,04 \pm 1,48

4.2 Estimação do consumo alimentar residual

O consumo alimentar residual (CAR) foi calculado no pacote estatístico SAS 9.3 por meio da regressão fenotípica do GMD e $PM^{0,75}$ em CMS, conforme citado por Crews (2005):

$$CMS_i = \beta_0 + \beta_1(GPD) + \beta_2(PVM^{0,75}) + \varepsilon_i$$

em que o CMS_i é Consumo de matéria seca predita para animal i ; β_0 é intercepto da regressão; β_1 é coeficiente de regressão parcial sobre GPD ; β_2 é coeficiente de regressão parcial sobre $PVM^{0,75}$; ε_i é erro residual do consumo predito do animal i , definindo o CAR.

4.3 Dados genotípicos – Controle de qualidade e imputação de genótipos

Foi totalizado 956 animais genotipados, que ao comparar com os animais que possuíam somente dados fenotípicos, apresentou 10 animais a mais com dados genômicos. Essa diferença foi ocasionada pelo fato de que animais sem fenótipo foram genotipados por serem pais de animais com relação genealógica com os participantes dos testes. Destes animais com dados genômicos, 604 foram genotipados com o painel de alta densidade de 777.962 SNPs da *Illumina® BovineHD BeadChip* (Illumina Inc., San Diego, CA, USA). Com painéis de baixa densidade, 121 animais foram genotipados com o painel de 74.677 SNPs da NeoGen®, e 231 animais genotipados com painel Clarifide de 56.169 SNPs da Zoetis®.

Para enriquecimento dos dados genotípicos foi realizada análise de imputação por meio do programa Fimpute v.2.2 (SARGOLZAEI; CHESNAIS; SCHENKEL, 2014). Este programa possui processamento rápido e metodologia determinística, considerando informações de família, trabalhando com genealogia e o parentesco entre indivíduos. A imputação acurada dos genótipos se inicia entre indivíduos aparentados, quando há informação de pedigree disponível, e, em seguida, procura haplótipos longos e curtos no grupo de referência, o qual representa parentesco próximo ou distante, respectivamente (SARGOLZAEI;

CHESNAIS; SCHENKEL, 2014).

A imputação dos painéis de baixa densidade foi realizada independentemente. Primeiramente foi imputado o painel de SNP de 74k para o HD seguido da imputação do painel de 56k para o de alta densidade, combinando informação de animais genotipados em painel de baixa e alta densidade. O grupo de referência utilizado foi o dos animais genotipados em HD, determinando imputação mais precisa por ser o maior grupo e com parentesco próximo. A estatística foi calculada com base nos dados genotípicos que foram codificados como 1 para heterozigoto (AB), 0 e 2 para homozigotos dominante (AA) e recessivo (BB), respectivamente e 5 para não identificados (NA).

Visando aumentar a precisão das análises genômicas foi feito controle de qualidade (CQ) dos genótipos com os seguintes parâmetros de exclusão: SNPs de cromossomos não-autossômicos, SNPs na mesma posição, alelos de menor frequência (MAF) menor que 0,05, taxa de genotipagem de SNPs (*call rate*) menor que 0,90 e, quando a proporção observada de marcadores heterozigotos diferiu da proporção esperada (equilíbrio de Hardy-Weinberg) em valores maiores que 0,15. Para amostras aplicou-se taxa de genotipagem (*call rate*) menor que 0,90 para exclusão.

Após imputação, os 956 genótipos continham informação de 730.422 SNPs, e após o CQ dos dados, 466.315 marcadores foram efetivados e utilizados nas análises subsequentes.

O LD foi estimado pelo quadrado da correlação das frequências alélicas (coeficiente de correlação - r^2) entre todos pares de SNPs, considerando a seguinte equação (HILL; ROBERTSON, 1968):

$$r_{ij}^2 = \frac{p_{ij} - p_i \times p_j}{p_i(1 - p_i) \times p_j(1 - p_j)}$$

onde p_{ij} é a frequência do haplótipo de dois marcadores, e p_i e p_j são as frequências alélicas marginais no i e j SNP, respectivamente. A frequência dos haplótipos é calculada como a diferença entre a frequência observada e a esperada. Se estas frequências são maiores ou menores do que o esperado, então esses dois *loci* tendem a segregar juntos e estão em LD, em que r^2 pode

variar de 0 a 1, onde zero significa não correlação entre os pares de SNPs, e um significa completa correlação entre os pares de SNP. Devido a quantidade significativa de comparações de todos pares SNPs, foi considerado para o cálculo de r^2 SNPs dentro de distâncias máximas de 100 kb um do outro, uma vez que o r^2 diminui em função da distância entre os marcadores e distâncias maiores apresentam baixos valores de LD (PRITCHARD; PRZEWORSKI, 2001). Para obter o LD foi utilizado o programa Plink v1.90 (CHANG et al., 2015) e resultado analisado no programa R (R DEVELOPMENT CORE TEAM, 2009).

O objetivo da análise de LD foi capturar a variação genética da população, verificando o nível presente na população estudada, visto que os padrões de LD são específicos para cada população e os SNPs selecionados podem diferir entre populações diferentes.

Para análise de componentes principais (PCA – do inglês *Principal component analysis*), as amostras e SNPs que permaneceram após o controle de qualidade foram submetidos a *pruning* (poda) dos SNPs, retirando marcadores que não estavam em LD (considerando $r^2 < 0,2$), obtendo SNPs não correlacionados dentro de uma janela de 50 marcadores, restando 50.410 SNPs que foram utilizados para a análise.

O arquivo *prune* foi utilizado para a PCA, aplicando os dois primeiros componentes principais para visualização de possíveis subestruturas na população e a relação entre os indivíduos estudados. A análise foi realizada com base na matriz de parentesco genômico pelo software Plink v1.90 (CHANG et al., 2015).

4.4 Estimação dos componentes de variância

Os componentes de variância foram estimados por máxima verossimilhança restrita (REML) utilizando o algoritmo AI (*average information*), cujo permite o cálculo dos erros-padrão (Tabela 2) (JENSEN et al., 1997). Foi utilizado o programa AIREMLF90 (MISZTAL et al., 2015) sob modelo animal, o qual calcula o desvio-padrão para variâncias conforme descrito por Meyer & Houle. (2013). Foi incluída informações genômicas por meio da matriz (H) para

calcular a herdabilidade estimada para o CAR. O modelo matricial pode ser representado como: $Y = X\beta + Z_a a + e$, em que: Y é o vetor de fenótipos; β é o vetor dos efeitos fixos e covariável; X é matriz de incidência que relaciona fenótipos com efeitos fixos; a é o vetor dos valores genéticos aditivos; Z_a é a matriz de incidência que relaciona animais com fenótipos e e é o vetor dos efeitos aleatórios residuais. O modelo estatístico considerou os efeitos sistemáticos de grupo contemporâneo (ano da prova e baía) como efeito fixo e idade no início da prova como covariável linear, além dos efeitos aleatórios de animal e resíduo.

O teste de parentesco apontou 3,66% de erros no pedigree dos animais genotipados, os quais foram tratados como pais desconhecidos, mantendo todos animais na matriz de parentesco, totalizando 3.643 animais, de 424 pais e 2.314 mães.

Tabela 2. Estimativas de componente de variância e herdabilidade e estatística descritiva seguidas do desvio-padrão da característica consumo alimentar residual em bovinos Nelore

Estatística Descritiva do CAR	Média
Média	9,210E-05
Desvio-padrão	0,796
Mínimo	-2,967
Máximo	4,067
Parâmetros genéticos	
σ_a^2	0,153 ± 0,049 *
σ_e^2	0,478 ± 0,045 *
h^2	0,242 ± 0,075

* Médias das variâncias seguidas por erro-padrão.

4.5 Estudo de associação ampla do genoma

Com o objetivo de detectar regiões que explicam a variância genética do CAR, foram estimados os efeitos dos SNPs e as variâncias explicadas, considerando janelas genômicas de 1Mb não sobrepostas, pelo método *weighted single-step* GBLUP (WssGBLUP) (WANG et al., 2012). A metodologia difere do ssGBLUP proposto por Misztal, Legarra e Aguilar (2009) por atribuir pesos diferentes aos marcadores, ponderando o efeito do marcador de acordo com sua importância na característica estudada.

Para o WssGBLUP foi utilizado mesmo modelo estatístico do item 4.4. A covariância entre a e e foi assumida igual a zero e as variâncias conforme descrição de Aguilar *et al.* (2010):

$$\text{var} \begin{pmatrix} a \\ e \end{pmatrix} = \begin{pmatrix} H\sigma_a^2 & 0 \\ 0 & I\sigma_e^2 \end{pmatrix}$$

em que σ_a^2 e σ_e^2 são, respectivamente, variância genética aditiva e variância residual; H é a matriz de parentesco baseada tanto em informação genômica quanto em informação de pedigree; I é a matriz de identidade. A matriz (H) substitui a matriz de parentesco para todos animais (A) e sua inversa é dada por:

$$H^{-1} = A^{-1} + \begin{pmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{pmatrix}$$

em que A_{22} é a matriz dos numeradores do coeficiente de parentesco entre animais genotipados e G é a matriz de parentesco genômico, construída de acordo com VanRaden (2008), que assume frequências alélicas da população estudada, ajustada para compatibilidade com a A_{22} (CHEN *et al.*, 2011; VITEZICA *et al.*, 2011).

As soluções dos efeitos dos SNPs foram obtidos segundo Strandén e Garrick (2009) proposta por Wang *et al.* (2014) pela equação $\hat{u} = DZ'[ZDZ']^{-1}\hat{a}_g$, onde D é a diagonal da matriz de ponderação das variâncias dos SNPs, Z é a matriz que relaciona os genótipos de cada *locus*, e \hat{a}_g é o vetor dos valores genéticos dos animais (EBV). A matriz D e os efeitos dos SNPs e dos animais são calculadas seguidamente por n iterações, onde a cada iteração se atualiza os valores com base na saída anterior da variância de cada marcador, conforme metodologia ssGBLUP/S2 descrita por Wang *et al.* (2012).

O algoritmo para estimação dos efeitos de SNPs e animal (Wang *et al.* (2012) segue os seguintes passos:

1. $t = 0$, $D_{(t)} = I$; $G_{(t)} * = ZD_{(t)}Z' \lambda$;
2. Computar \hat{a}_g por ssGBLUP;
3. Cálculo de: $\hat{u}_{(t)} = \lambda D_{(t)}Z' G_{(t)} *^{-1} \hat{a}_g$;
4. Cálculo dos pesos para cada SNP: $G_{(t+1)} * = \hat{u}_{(t)}^2 2p_i (1 - p_i)$ para todos i como em Zhang *et al.* (2010);
5. Normalização dos pesos para variância genética aditiva permanecer

constante:

$$D(t+1) = \frac{\text{tr}(D(0))}{\text{tr}(D^*(t+1))} D^*(t+1).$$

6. Cálculo de $G^*_{(t+1)} = ZD_{(t+1)}Z' \lambda$;

7. $t = t + 1$;

8. Parar, ou voltar ao passo 2.

em que t representa o número de iterações, i o i -th SNP e λ é uma razão de variâncias ou uma constante de normalização. Ao retornar para o passo 2 são reestimados os efeitos de SNPs e animal, a partir do GEBV (ssGBLUP/S2). Desta maneira, obtém-se os valores genéticos genômicos preditos e os efeitos dos marcadores com maior acurácia na detecção das regiões genômicas, diferenciando QTL com maior ou menor efeito.

A porcentagem da variância genética explicada pela i -th região foi calculada conforme:

$$\frac{\text{Var}(S_i)}{\sigma_a^2} \times 100\% = \frac{\text{Var}(\sum_{j=1} Z_j \hat{u}_j)}{\sigma_a^2} \times 100\%$$

onde S_i é o valor da i -th região, que consiste em janelas de 1Mb com diferentes números de SNPs, σ_a^2 é a variância genética total, Z_j é o vetor de conteúdo genético do j -th SNP para todos indivíduos e \hat{u}_j é o efeito de marcador da j -th SNP na região i -th.

Foram realizadas duas iterações, na primeira o mesmo peso (iteração 1 = 1) foi dado para todos SNPs e serviu de base para obter a proporção de variância explicada de cada SNP, que foi utilizado na iteração seguinte, e utilizado janela de SNPs com tamanho de 1Mb. Wang et al. (2012), descreveram que com o processo iterativo obtiveram maior acurácia da estimação dos efeitos dos SNPs após uma rodada de iteração, em relação ao algoritmo ssGBLUP/S2, o qual foi utilizado no presente estudo.

Características complexas, como o CAR, possuem muitas variantes comuns de pequeno efeito influenciando o fenótipo, abrangendo grande número de genes. Desta maneira, a metodologia WssGBLUP, que atribui ponderações a marcadores genéticos via processo iterativo, representa a real situação de características poligênicas. A ponderação é importante por evidenciar regiões de maior efeito sobre a característica estudada, em que as iterações aumentam o

peso atribuído a SNPs de maior efeito e diminui o peso daqueles de menor efeito (WANG et al., 2012).

Como as ponderações dos SNPs com grandes efeitos são aumentadas e SNPs com efeitos menores são diminuídos, foi observado aumento da proporção da variância explicada pelos marcadores de acordo com as subseqüentes iterações realizadas. As análises foram realizadas pela família de programas BLUPF90 (MISZTAL et al., 2015).

4.6 Anotação de genes e análise funcional

Foram computadas 3.197 janelas genômicas espalhadas pelos cromossomos autossômicos, sendo que oito (8) regiões genômicas de 1Mb explicaram mais que 1% da variância genética aditiva, com maior probabilidade de estarem associadas com regiões QTL. Foram identificados genes presentes nestas janelas utilizando a ferramenta *Map Viewer* disponível no NCBI (*National Center for Biotechnology Information*), usando a montagem da sequência do genoma bovino UMD 3.1 (ZIMIN et al., 2009) como referência. A presença de QTL bovinos previamente relatados sobrepondo as janelas identificadas neste estudo, foi verificado na base de dados CattleQTL (HU; PARK; REECY, 2016). Os marcadores significativos para o CAR foram enriquecidos em genes que constituem determinada via metabólica, visando melhor entendimento dos mecanismos biológicos (WENG et al., 2011). A análise de enriquecimento foi realizada utilizando teste exato de Fisher e algoritmo *weight01* implementado pelo pacote topGO para o R (ALEXA; RAHNENFUHRER, 2016). Foram considerados os genes contidos nas oito (8) janelas, considerando significância estatística do valor $p < 0,01$ e analisado os três domínios: Processo biológico (BP), componente celular (CC) e função molecular (FM).

4.7 Diferenciação genética entre grupos divergentes para CAR

Dois grupos contrastantes foram definidos de acordo com o fenótipo ajustado. Os animais foram separados em eficientes ou ineficientes para o CAR,

estabelecido como 20% das extremidades da característica. A distinção pelo fenótipo ajustado permitiu que animais sem genótipo fossem incluídos nos grupos, sendo um animal no grupo dos eficientes e três dos ineficientes, estes animais foram removidos e as análises realizadas com 187 animais no grupo eficiente (*cluster 1*) e 189 animais grupo ineficiente (*cluster 2*), totalizando 376 animais identificados como divergentes para o fenótipo.

Foi realizado teste exato de Fisher usando todos SNPs para avaliar a diferença nas frequências alélicas entre os grupos. Os SNPs com valor significativo para o teste de Fisher ($p < 0,05$) foram considerados discordantes entre os grupos. Estes marcadores foram submetidos a correção de taxa de falso positivo (FDR – *False Discovery rate*) por gene para corrigir os valores de p do teste de Fisher, considerado q -valor para FDR menor que 5% (BENJAMINI; HOCHBERG, 1995). O q -valor mede a proporção de falsos-positivos quando o p -valor é considerado significativo. Os SNPs que permaneceram após a correção, foram denominados SNPs diferenciados.

Os SNPs com MAF menores que 5% (variantes raras) foram removidos no controle de qualidade, por serem monomórficos ou apresentarem frequência muito baixa, com o objetivo de aumentar o poder dos testes estatísticos realizados na análise de associação, diminuindo erros de associações (TURNER et al., 2011). A MAF representa a proporção do alelo menos frequente em determinado locus, variando de 0 a 50%. Segundo VANRADEN et al. (2009), marcadores com MAF baixa (2 a 5%) tem consequentemente baixa heterozigosidade, portanto, são pouco informativos.

A MAF foi dividida em cinco classes de acordo com a frequência de variantes, sendo: 0,05-0,1, 0,1-0,2, 0,2-0,3, 0,3-0,4 e 0,4-0,5 e classificadas em baixa frequência (MAF entre 5 e 10%) e comum (MAF entre 10% e 50%). Dos 466.315 SNPs que permaneceram após o controle de qualidade, foi aplicado o teste exato de Fisher ($p < 0,05$) nas frequências alélicas entre os grupos divergentes de CAR para obtenção dos SNPs com frequências alélicas diferentes entre os grupos.

Índice de fixação – Fst

O índice de fixação (Fst) é amplamente utilizado para identificação da diferenciação genética na população. Foi abordada o índice de fixação de Wright (Fst - do inglês *Wright's fixation index*) para verificar diferenciação genética entre os grupos contrastantes. Este método fornece estimativa do quanto da variabilidade genética é dividida entre populações (COCKERHAM, 1969; WRIGHT, 1950). A seleção altera a frequência alélica entre subpopulações (ou grupos) em relação a variância total da população e pode ser mensurada pelo Fst, determinando como esta seleção pode afetar o padrão genético dos grupos avaliados, sendo utilizado para detectar assinaturas de seleção entre populações ao comparar os seus valores entre *loci* (WEIR; CAROLINA; HILL, 2002; ZHAO et al., 2015a).

Ao quantificar diferenças nas frequências alélicas entre os grupos se obtêm o valor da estatística Fst, determinando se as diferenças entre raças e/ou grupos são decorrentes de seleção recente. O Fst está diretamente relacionado com a variância na frequência alélica entre populações, e é inversamente relacionado ao grau de semelhança entre indivíduos dentro de populações, com valor variando de zero a um, significando sem diferença nas frequências alélicas, ou completa fixação de alelos diferentes nos grupos, respectivamente (HOLSINGER; WEIR, 2009).

A diferenciação genética entre os grupos foi mensurada utilizando o estimador de Fst proposto por Weir e Cockerham (1984). Valores de Fst foram estimados para todos os 466.315 SNPs, considerando animais eficientes vs. animais ineficientes. Valores negativos de Fst foram ajustados para zero, dado que estes não possuem interpretação biológica (AKEY et al., 2002). Médias de Fst foram calculadas em janelas de 100kb, com 50kb de sobreposição.

O Fst requer a utilização de limites de significância para refutar a hipótese nula (que não existe efeito da seleção sobre específica região genômica) (HAYES et al., 2009b). Com base nas janelas significativas de Fst foram selecionadas as 10 janelas com maiores valores e com mais de cinco SNPs, permitindo identificar

os SNPs que se destacaram dos demais. Abordagem similar foi utilizada em outros estudos para identificação da diferenciação genética (CARDOSO et al., 2018).

4.8 Estimativa do efeito da substituição alélica

Para quantificar os efeitos dos marcadores com diferentes frequências alélicas, foi realizada o cálculo da substituição alélica. O efeito do marcador foi inferido por modelo em que o valor genético do animal é convertido em soluções dos SNPs e usado para estimar as variâncias individuais dos efeitos destes, conforme (ZHANG et al., 2010): $\hat{\sigma}_{u,i} = \hat{u}_i^2 2p_i(1 - p_i)$, em que \hat{u}_i^2 é o quadrado do efeito do i th SNP, p_i é a frequência alélica observada para o segundo alelo do i th SNP da população (descrito no item 4.4). Os efeitos de substituição alélica podem ser calculados usando modelos de predição genômica, e representam o efeito que uma cópia de determinado alelo tem no fenótipo (BOUWMAN; HAYES; CALUS, 2017). A mudança média no valor genotípico pela substituição aleatória de um alelo “A” por um alelo “B” em uma população é definida como o efeito da substituição aditiva (LYNCH; WALSH, 1998).

O conteúdo genético (z) de um marcador é o número de cópias de um alelo de referência específico, em que $z = 0, 1$ ou 2 para AA, AB e BB (FALCONER; MACKAY, 1996). Para os três genótipos, o valor de z é observado como característica quantitativa quando o arranjo genotípico (0, 1 e 2) tem efeito fenotípico como (0, α e 2α), e o efeito aditivo α do alelo de referência, neste caso, alelo B, é justamente 1, portanto, sem dominância e epistasia (FORNERIS et al., 2015). A média de z na população é $2p$, em que p é a frequência alélica e a variância igual a $2pq$, onde $q = 1-p$, e, conseqüentemente, a variância observada em z é especificamente genética aditiva.

Este α pode ser entendido como efeito de substituição alélica, o qual inclui ação gênica não aditiva. Os valores aditivos (genéticos) abrangem grande parte dos efeitos de dominância, além disso, devido a mudanças nas frequências alélicas, os efeitos de substituição que capturam dominância e efeitos epistáticos não se mantêm estáveis nas gerações (VARONA et al., 2018). A seleção atua

nos efeitos de substituição aditiva, pois os alelos e não os genótipos são transmitidos as gerações (VITEZICA; VARONA; LEGARRA, 2013). Além disso, o efeito da substituição aditiva captura grande parte dos efeitos não-aditivos (dominância e epistasia) (FALCONER, 1987). Portanto, apenas os efeitos de substituição (efeitos aditivos) interferem nos valores genéticos e são passados nas gerações, o qual foi utilizado para cálculo da variância explicada por cada SNP diferenciado estatisticamente entre os grupos divergentes para o CAR, explicando sua contribuição para a característica.

A variância genética total foi calculada como a soma da variância de todos SNPs. A proporção da variância genética explicada pelos SNPs foi estimada:

$$VAR (\%) = \frac{2pq\beta^2}{S^2} \times 100$$

em que p e q são as frequências alélicas, β^2 é o quadrado do efeito da substituição alélica e S^2 é a variância genética total dos marcadores.

5. RESULTADOS E DISCUSSÃO

5.1 Estrutura populacional e desequilíbrio de ligação

Analisando a estrutura populacional nos dois primeiros componentes principais (Figura 1), não se consta clara a estratificação da população, com poucas diferenças de ancestralidade, além de indicar baixas diferenças nas frequências alélicas entre os animais (PRICE et al., 2006). A PCA permite visualizar e classificar os indivíduos em grupos de acordo com a similaridade genética, segundo suas variâncias e auxilia na interpretação e elucidação da estrutura de determinada população (PRICE et al., 2006).

A análise de PCA demonstra que os indivíduos pertencem a um grupo principal, indicando que os animais analisados pertencem a mesma população ancestral. Por ser um único grupo, não houve necessidade em ajustar os dados para a ancestralidade (PRICE et al., 2006).

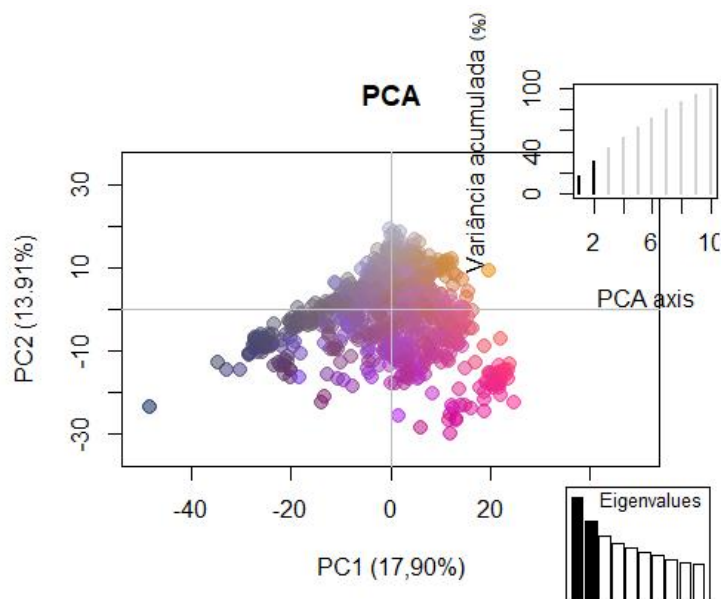


Figura 1. Distribuição dos animais em relação ao primeiro (PC1) e segundo (PC2) componentes principais estimados com base na matriz de parentesco.

Para análise de LD foi utilizado todos SNPs após CQ. Esses marcadores estavam distribuídos nos 29 cromossomos autossômicos, e a análise descritiva dos SNPs por cromossomo é apresentada na Tabela 3.

O r^2 médio para todos os autossomos foi de 0,11 em uma distância física de 100kb entre os marcadores, com amplitude de 0,087 a 0,127, para os cromossomos 27 e 14, respectivamente. O valor para mesma distância foi similar (0,11) ao obtido por Espigolan et al. (2013), em outro conjunto de dados de animais Nelore, indicando que o valor do LD observado é adequado a populações zebuínas.

As principais causas que diminuem a extensão do LD são os cruzamentos, altas taxas de recombinação e mutação. O valor médio de r^2 do presente trabalho pode ser explicado pela maior distância utilizada (100kb). De fato, as taxas de recombinação são proporcionais a distância entre os marcadores, sendo menor em distâncias reduzidas e aumentada em maiores distâncias. Marcadores genéticos próximos tem genealogia semelhante, induzindo dependência entre os alelos em diferentes marcadores, enquanto marcadores mais distantes fisicamente tendem a apresentar diferentes genealogias, devido a recombinação (PRITCHARD; PRZEWORSKI, 2001). Deste modo, o decaimento do LD tende a

ser mais rápido em regiões propensas a maior recombinação, como no caso de maiores distâncias (WATKINS et al., 1994).

Em estudos de associação genômica, o SNP é associado a mutação causal pelo desequilíbrio de ligação, ou seja, estão associados ao fenótipo de forma indireta (BUSH; MOORE, 2012). Geralmente se adota a medida r^2 para explorar as soluções de estudos associativos, em virtude de ser indicativa de como os SNPs se correlacionam com o QTL (DU; CLUTTER; LOHUIS, 2007). A associação do SNP com a variante causal por associação indireta é afetada pela extensão do LD, e, quanto maior frequência de marcadores esparsos pelo genoma, maior facilidade de detecção do QTL.

As associações entre polimorfismos ocorrem principalmente pela proximidade física, mas são influenciadas também pela recombinação, seleção artificial e pelos fatores evolutivos da população, como por exemplo deriva genética, gargalos populacionais, mutação, seleção natural e fluxo gênico. Desse modo, o padrão do LD auxilia na descrição dos eventos ocorridos na população.

Segundo Arias et al. (2009), a variabilidade nas taxas de recombinação nos autossomos provoca diversidade considerável no padrão do LD em diferentes regiões, resultado verificado com o obtido neste estudo na amplitude 0,087 a 0,127 entre os cromossomos.

Tabela 3. Descrição do desequilíbrio de ligação (r^2) para distância física de 100kb entre os marcadores, seguido por desvio padrão (SD) entre cada cromossomo

Cromossomo	SNP (n)	Média $r^2 \pm$ SD
1	29.770	0,108 \pm 0,184
2	25.840	0,114 \pm 0,190
3	23.242	0,110 \pm 0,187
4	21.475	0,100 \pm 0,178
5	20.644	0,120 \pm 0,198
6	24.686	0,124 \pm 0,195
7	21.011	0,120 \pm 0,197
8	21.932	0,119 \pm 0,191
9	21.078	0,108 \pm 0,185
10	17.909	0,101 \pm 0,183
11	19.816	0,126 \pm 0,207
12	16.036	0,110 \pm 0,186
13	14.831	0,100 \pm 0,178
14	17.309	0,127 \pm 0,202
15	15.421	0,101 \pm 0,177
16	14.951	0,104 \pm 0,179
17	15.057	0,113 \pm 0,189
18	12.323	0,106 \pm 0,188
19	10.454	0,099 \pm 0,174
20	13.080	0,107 \pm 0,185
21	13.292	0,105 \pm 0,184
22	10.901	0,098 \pm 0,175
23	10.083	0,096 \pm 0,167
24	11.761	0,110 \pm 0,188
25	8.028	0,101 \pm 0,180
26	10.165	0,088 \pm 0,165
27	8.317	0,087 \pm 0,159
28	8.403	0,108 \pm 0,184
29	8.470	0,102 \pm 0,178

Existe relação inversa entre o tamanho da população e o LD, conforme descrito por (MCKAY et al., 2007), o qual encontraram valores médio de r^2 de 0,37 e 0,56 para animais *Bos indicus* e *Bos taurus*, respectivamente. Animais *Bos taurus* apresentam maior LD devido a redução da população após a domesticação e processo intenso de seleção artificial da raça (GODDARD; HAYES, 2009). Os

valores de LD são maiores para taurinos, indicando que são animais mais homozigotos em comparação aos zebuínos, os quais possuem maior variabilidade na formação das raças, e, portanto, mais heterozigotos e com LD inferior. Por esta particularidade é recomendável a utilização de painéis de maiores densidades nos zebuínos (PÉREZ O'BRIEN et al., 2014). Outra característica que difere as raças é o fato dos zebuínos apresentarem proporção superior de alelos de menor frequência (MAF) em relação aos taurinos, o que influencia e determina o nível de LD.

De fato, a extensão do LD depende da história demográfica de determinada população por possuírem características específicas, o que justifica as diferenças dos valores da medida para diferentes populações e, também, de diferentes espécies.

5.2 Estudo de associação genômica e enriquecimento funcional

Os SNPs de menor efeito apresentaram efeitos reduzidos, enquanto os de maior efeito tornaram-se maiores com a recomputação dos pesos dos SNPs, e recálculo dos efeitos destes pelo processo iterativo. De acordo com Wang et al., (2012), a segunda iteração foi que possibilitou maior acurácia de estimação dos efeitos dos SNPs utilizando o WssGBLUP. Melo et al. (2016) apresentaram resultado similar, detectando com maior precisão QTL na segunda iteração, em estudo de simulação.

Das 3.197 janelas genômicas espalhadas pelos cromossomos autossômicos, oito (8) segmentos cromossômicos de 1Mb de extensão explicaram mais que 1% da variância genética aditiva, e consideradas de maior efeito, com maior probabilidade de estarem associadas com regiões QTL do CAR. As janelas que apresentaram maior proporção da variância genética explicada pelos marcadores para o CAR, obtidas na segunda iteração, são apresentadas via Manhattan Plot (Figura 2).

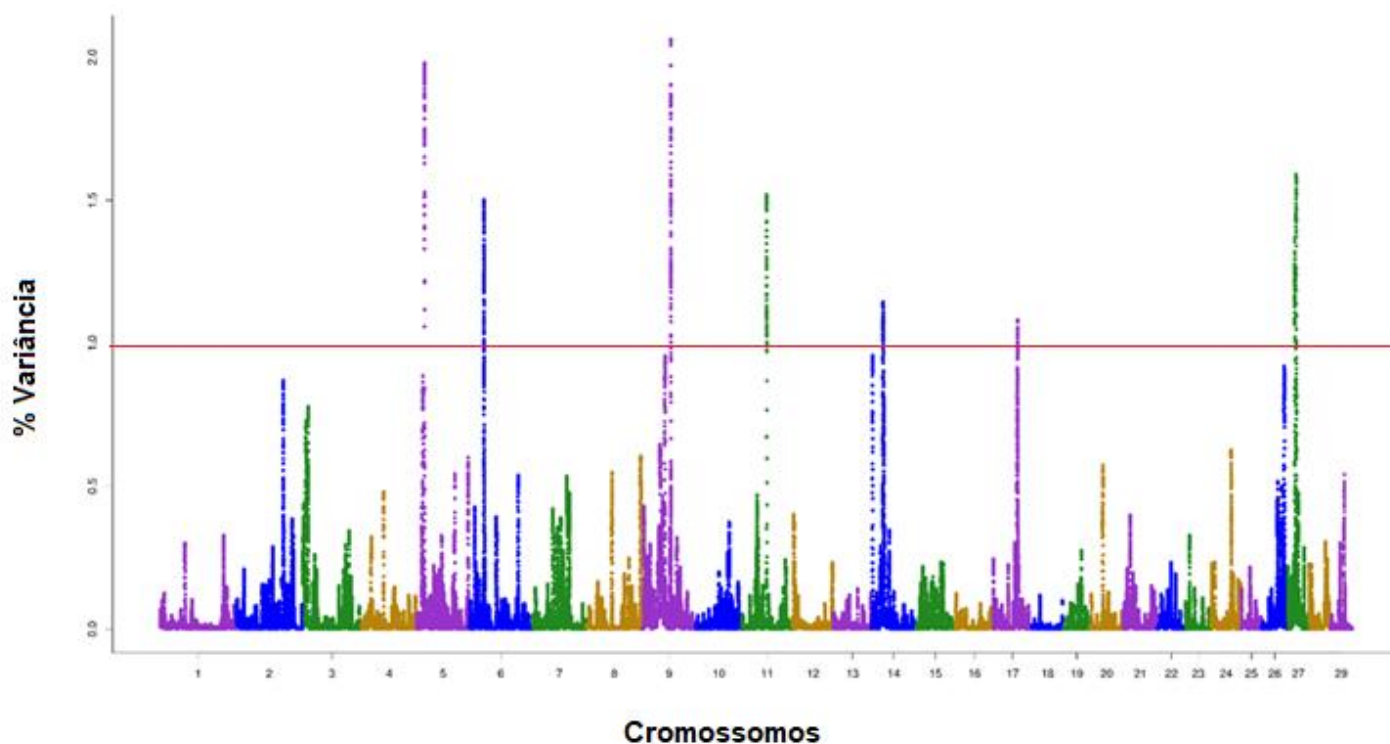


Figura 2. Manhattan Plot do consumo alimentar residual (CAR) de bovinos Nelore. O eixo y representa a proporção da variância explicada por janelas de 1Mb e o eixo x indica os cromossomos. A linha vermelha identifica as regiões cromossômicas com variância genética aditiva maior que 1%.

As oitos regiões genômicas com variância genética maior que 1% explicaram 12,25% da variância genética total para o CAR e estão localizadas nos cromossomos 5, 6, 9, 11, 14, 17 e 27. O cromossomo 27 apresentou duas janelas distintas de maior efeito, explicando 2,96% da variância total, com maior valor para o CAR. As variâncias genéticas e os genes conhecidos presentes nas janelas estão apresentados na Tabela 4.

Tabela 4. Localização cromossômica de regiões de importância para CAR, considerando janelas com 1Mb e variância genética maior que 1% em bovinos da raça Nelore

Cromossomo	Posição (pb)	Número de SNPs	Genes	Variância (%)	QTL*
5	15.517.762-16.508.110	244	RASSF9, MGAT4C, NTS, LOC107132461	1,98	6
6	28.332.475-29.328.534	256	LOC104972708, LOC517517, LOC104972709, LOC100301359, LOC782904, TRNAC-GCA, LOC104972710	1,50	18
9	55.124.063-56.120.878	213	MANEA	2,06	7
11	49.557.164-50.554.834	243	SUCLG1, LOC104973382, LOC101907084, TMSB10, LOC104973383, KCMF1, LOC101903367, LOC104973384, TCF7L1	1,52	30
14	22.823.936-23.823.912	259	ST18, LOC100141260, LOC101906592, LOC104974017, ALKAL1, RB1CC1, NPBWR1, OPRK1, ATP6V1H, RGS20, TCEA1, LYPLA1, MRPL15, LOC101906669, LOC107133113, TRNAG-UCC, TRNAC-GCA	1,14	19
17	48.981.534-49.976.674	320	TMEM132D, GLT1D1, LOC101907645, TRNAE-UUC, SLC15A4, TMEM132C,	1,08	19
27	14.043.813-15.028.998	202	CASP3, LOC104976051, PRIMPOL, LOC104976052, CENPU, TRNAC-GCA, ACSL1, HELT, LOC790015, SLC25A4, CFAP97, SNX25, LRP2BP, ANKRD37, UFSP2, C27H4orf47, CCDC110, PDLIM3, SORBS2	1,37	2
	16.362.786-17.362.110	223	LOC101905700, ZFP42, TRNAG-UCC, TRIML2, TRNAF-AAA, TRIML1, LOC507011	1,59	

* número de QTL/associações em bovinos previamente associados ao CAR consultando o banco de dados CattleQTL (HU; PARK; REECY, 2016)

Considerando todas as regiões foram encontrados 71 genes anotados, com alguns ainda não caracterizados (LOC). Foi constatado o caráter quantitativo do CAR, com influência de grande número de genes de pequeno efeito, observando as proporções das variâncias explicadas pelas janelas descritas. O CAR é característica poligênica e influenciada por muitos processos biológicos (BOLORMAA et al., 2013; KARISA; MOORE; PLASTOW, 2013).

A primeira janela genômica, localizada no cromossomo 5, nas posições 15Mb e 16Mb apresentou 3 genes caracterizados (RASFF9, MGAT4C, NTS) associados ao CAR. São genes citados como candidatos para o ganho de peso em animais Nelore (OLIVIERI et al., 2016). O gene RASSF9 foi associado com a espessura de gordura subcutânea na garupa em animais Brangus, localizado dentro de uma região QTL e está envolvido no processo de mitose (WENG et al., 2016).

Segundo a base de dados CattleQTL, há três QTL associados ao CAR localizados na posição entre 27.034.490-29.073.969Mb que sobrepueram a região do cromossomo 6 identificada neste estudo. Os QTL são os 21050, 21051 e 21052 identificados em animais Angus (ROLF et al., 2011). Outro QTL associado ao CAR, o 11449, foi identificado em animais cruzados e é localizado próximo a região associada deste estudo (27,7-27,9Mb) (MÁRQUEZ et al., 2009).

De maneira geral, o enriquecimento capturou 18 termos altamente significativos associados ao CAR (Tabela 5).

Tabela 5. Identificação (GO.ID) e descrição (termos) de termos GO (gene ontology) enriquecidos significativamente após GWAS em bovinos da raça Nelore

GO.ID	Genes	Termos	Anotados ¹	p-value
GO:0000221	ATP6V1H	<i>vacuolar proton-transporting V-type ATPa...</i>	3	0,0055
GO:0004569	MANEA	<i>glycoprotein endo-alpha-1,2-mannosidase ...</i>	1	0,0018
GO:0038048	OPRK1	<i>dynorphin receptor activity</i>	1	0,0018
GO:0071567	UFSP2	<i>UFM1 hydrolase activity</i>	2	0,0036
GO:0005471	SLC25A4	<i>ATP:ADP antiporter activity</i>	2	0,0036
GO:0004775	SUCLG1	<i>succinate-CoA ligase (ADP-forming) activ...</i>	2	0,0036
GO:0004776	SUCLG1	<i>succinate-CoA ligase (GDP-forming) activ...</i>	2	0,0036
GO:0003896	PRIMPOL	<i>DNA primase activity</i>	3	0,0054
GO:0021858	HELT	<i>GABAergic neuron differentiation in basa...</i>	1	0,0018
GO:0031635	OPRK1	<i>adenylate cyclase-inhibiting opioid rece...</i>	1	0,0018
GO:0046877	OPRK1	<i>regulation of saliva secretion</i>	4	0,0073
GO:0061709	RB1CC1	<i>reticulophagy</i>	2	0,0037
GO:0061723	RB1CC1	<i>glycophagy</i>	2	0,0037
GO:0045793	RB1CC1	<i>positive regulation of cell size</i>	5	0,0091
GO:0072734	CASP3	<i>cellular response to staurosporine</i>	2	0,0037
GO:0006269	PRIMPOL	<i>DNA replication, synthesis of RNA primer</i>	4	0,0073
GO:0019222	NPBWR1	<i>regulation of metabolic process</i>	5016	0,0081
GO:0033211	ACSL1	<i>adiponectin-activated signaling pathway</i>	5	0,0091

¹ Total de genes anotados para o termo descrito.

Os 18 termos GO (*gene ontology*) identificados apresentaram 13 genes associados com o consumo alimentar residual. As ontologias gênicas descrevem e categorizam genes de acordo com suas funções moleculares, envolvimento em processos biológicos e localização celular, com intuito de verificar a frequência dos termos funcionais associados aos genes. Dos termos GO identificados, 10 deles são referentes ao domínio de processos biológicos, 7 ao domínio de função molecular e 1 ao domínio de componente celular.

Incorporar informação funcional, a partir da análise de vias biológicas, é estratégia para potencializar os resultados do GWAS (YANG et al., 2017). As análises fundamentadas em vias consideram fatores que contribuem

simultaneamente para características poligênicas, não se limitando a SNP e gene mais significativo. O CAR possui muitas alterações genéticas de pequeno efeito que contribuem para diferenças na característica (SERÃO et al., 2011). Desse modo, o excesso de variedade de termos de ontologia genética e genes associados dificulta o enriquecimento funcional, o que representa baixa proporção em relação ao que é observado em características governadas por genes de efeito maior (PEREIRA et al., 2018).

As principais vias metabólicas associadas ao CAR possuem relação com a função mitocondrial e metabolismo energético. Destacam-se os genes OPRK1, SUCLG1 e RB1CC1 que foram enriquecidos e estão envolvidos em diferentes termos. Outros genes que não fazem parte de processos biológicos enriquecidos, mas que foram identificados no estudo de associação genômica, merecem atenção devido as suas funções conhecidas sobre processos que são relacionados ao CAR e foram descritos.

As anotações de GO relacionadas ao gene OPRK1 incluem atividade de receptores opióides do tipo kappa acoplado a proteína G, que funciona como receptor para opióides endógenos (como as dinorfinas), sinalização do receptor opióide inibidor da atividade da adenilato ciclase e desempenha papel na regulação da salivação. Este neuropeptídeo se liga preferencialmente ao receptor opióide kappa, estando distribuído no sistema nervoso central e relacionado com diferentes funções fisiológicas em ratos, incluindo regulação neuroendócrina, atividade motora, função cardiovascular, respiração, regulação da temperatura, resposta ao estresse e comportamento alimentar (FALLON; LESLIE, 1986; WATSON et al., 1982). Além disso, a manipulação da transmissão de opióide altera o consumo alimentar, que altera o nível de opióides no cérebro (GLASS; BILLINGTON; LEVINE, 2000).

O gene SUCLG1 está relacionado as reações de glicólise e ciclo de Krebs, com formação de ADP ou GPD. A respiração celular é processo que ocorre no interior das células por meio do qual se obtém energia, e constitui reações químicas distribuídas em três etapas, glicólise, ciclo de Krebs e fosforilação oxidativa, com as duas últimas ocorrendo na mitocôndria (HILL; WYSE; ANDERSON, 2012). A via metabólica de glicólise/glicogênese está relacionada

ao fornecimento de energia no organismo animal, sendo ativada após a alimentação para gerar ATP (RUI, 2014).

O gene SUCLG1 (succinil-CoA sintetase) está localizado no cromossomo 11. A enzima succinil-coA é constituída por duas subunidades, onde a subunidade α é codificada pelo gene SUCGL1, e expressa no coração, fígado, rim e cérebro (RIVERA et al., 2010; ROUZIER et al., 2010). Esta enzima atua catalisando a síntese reversível do succinato e ATP no ciclo dos ácidos tricarboxílicos (TCA), ou ciclo de Krebs (as vias que oxidam lipídios, carboidratos e aminoácidos convergem para o TCA), e está envolvida na replicação do DNA mitocondrial. As mitocôndrias produzem maior parte da energia utilizada pela célula, por meio da fosforilação oxidativa (LOWENSTEIN, 1969). A relação entre função mitocondrial e consumo alimentar foi observada em gado de corte, utilizando mitocôndrias musculares de animais Angus, apesar dos autores não encontrarem diferenças da função mitocondrial entre grupos divergentes para o CAR, a taxa de respiração mitocondrial foi maior em animais baixo CAR, indicando maior eficiência (KOLATH; KERLEY, 2006).

Em trabalho avaliando o perfil do transcriptoma epitelial ruminal em bovinos divergentes para o CAR, o gene candidato SUCLG1 estava presente no grupo de genes que foram supra-regulados, e como consequência, o aumento da expressão, sugerindo maior taxa de transcrição nas mitocôndrias epiteliais, indicando maior atividade mitocondrial e/ou maior número de cópias do genoma mitocondrial no epitélio ruminal de animais eficientes (KONG et al., 2016).

Neste mesmo trabalho, o gene TMSB10, com função na organização do citoesqueleto, foi um dos 122 genes diferencialmente expressos no epitélio ruminal entre animais baixo e alto CAR, sendo supra-regulado em animais eficientes. Mensurando diretamente a atividade mitocondrial, outros estudos (KOLATH; KERLEY, 2006; LANCASTER et al., 2014) relataram taxa mais rápida de fosforilação oxidativa nas mitocôndrias do músculo esquelético e no fígado em bovinos baixo CAR. Apoiando este resultado Iqbal et al. (2005), sugeriram que o estresse oxidativo é aumentado em animais menos eficientes devido a diferenças na função mitocondrial.

O gene RB1CC1 (Cromossomo 14/22.83-23.83Mb) está relacionado com

três vias, a glicofagia, a reticulogafia e a de regulação positiva do tamanho das células. Este gene possui somente SNPs não sinônimos identificados, ou seja, que resultam em trocas de aminoácidos, e é relacionado diretamente ao desenvolvimento muscular por ser associado com autofagia, apoptose, proliferação de células e crescimento (GODOY et al., 2015; NISHIMURA et al., 2011). Além disso, o gene foi diferencialmente expresso em três tecidos (cérebro, fígado e músculo esquelético) de frangos sob estresse térmico. A diferença na expressão deste gene em resposta ao estresse térmico pode ser relacionada a eficiência dos animais, uma vez que a exigência de manutenção em bovinos de corte varia de 10 a 12% e é afetada pela produção de calor (HOTOVY et al., 1991; SWANSON; MILLER, 2008).

Processos metabólicos afetam a produção de calor, causando diferenças na eficiência do uso de energia para manutenção dos animais, sendo que a energia de manutenção por unidade de peso metabólico foi associada com a variação genética do CAR (ARCHER et al., 1999; HERD; BISHOP, 2000). Segundo Nkrumah et al. (2006), em estudo de calorimetria indireta observou que animais baixo CAR produziram 21% menos calor, bem como menor produção de metano (HEGARTY et al., 2007). Assim, diferentes variantes do gene podem atuar na eficiência alimentar, resultando em animais mais ou menos eficientes.

Na mesma janela foram identificados mais dois genes caracterizados, sendo que o KCMF1 está relacionado a atividade nos canais de potássio, que estão envolvidos na regulação da homeostase celular, com função em diversas rotas metabólicas basais (OLIVEIRA et al., 2014). O bombeamento iônico é o processo metabólico no qual íons são movidos por meio das membranas celulares contra o gradiente de concentração, processo altamente exigente de energia (MONTANHOLI, 2009). No metabolismo energético, o bombeamento de íons com destaque é a bomba de sódio-potássio (Na^+/K^+), com diferentes funções fisiológicas e responsável por 20 a 30% da taxa metabólica basal (BALDWIN et al., 1980).

O gene MANEA, presente na região identificada do cromossomo 9, codifica a endo- α -1,2-manosidase, que atua no metabolismo de vias protéicas (KUKUSHKIN et al., 2012). Esta enzima é produzida no complexo de Golgi e está

envolvida na síntese de glicoproteínas. Esta região apresentou apenas este gene associado e com maior variância explicada (2,06%) para o CAR. Nesta região existem quatro QTL (22786, 22830, 22831, 22832) associados ao ganho médio diário, característica componente do CAR, o que sugere efeito pleiotrópico do gene MANEA devido sua atuação no metabolismo de vias protéicas (PETERS et al., 2012).

O CAR está relacionado as necessidades energéticas basais e nas diferenças da eficiência do crescimento, conseqüentemente, servindo como indicador da eficiência metabólica (RICHARDSON; HERD, 2004). Além disso, diferença na eficiência alimentar indica variação nos processos biológicos básicos, com evidências de que os requisitos de energia basal, principalmente na função de manutenção das células, são os principais causadores desta variação (BALDWIN et al., 1980; RICHARDSON; HERD, 2004).

Na região identificada no cromossomo 14, o gene NPBWR1 foi associado ao termo de regulação de processos metabólicos, o qual possui muitos genes anotados devido sua generalização. Este gene é relacionado com regulação homeostática da alimentação, comportamento alimentar e do metabolismo energético (TAKENOYA et al., 2010). Este gene, juntamente com o OPRK1, foram também correlacionados com idade ao primeiro parto, em animais Nelore, por participarem do feedback negativo da secreção do hormônio luteinizante (MOTA et al., 2017). O gene LYPLA1 que atua como um inibidor da grelina, hormônio envolvido na regulação do apetite, foi associado com consumo alimentar residual com dois SNPs significativos associados (LINDHOLM-PERRY et al., 2012; SHANADO et al., 2004). Ainda, segundo Lindholm-Perry et al. (2012), variações neste gene interferem na regulação da grelina e pode resultar em diferenças no apetite, e conseqüentemente, no ganho de peso em bovinos.

Na região do cromossomo 17, o gene TMEM132D foi identificado a partir de varredura seletiva como gene candidato durante a domesticação e formação da raça Fleckvieh, sendo que foi relacionado com fenótipos neurocomportamentais associados a ansiedade, diferindo os animais domesticados de seus ancestrais selvagens (ERHARDT et al., 2012; QANBARI et al., 2014).

Avaliando características de eficiência alimentar, Pasandideh; Rahimi-Mianji e Gholizadeh (2018), encontraram três SNPs significativos associados ao ganho médio diário em ovinos, e quando comparado a genes conhecidos em bovinos, o SLC15A4 foi um dos genes mais próximo da região associada a característica.

No cromossomo 27 foram identificadas duas regiões genômicas associadas ao CAR. A primeira região nas posições 14 e 15 Mb foi que apresentou maior número de genes associados ao CAR, com 19 genes candidatos. Um dos genes associados, o CASP3, é envolvido na resposta imune inata (JOHNSON et al., 2013). A variação do CAR pode ser causada por diferenças individuais na função do sistema imune, dado que é um sistema energeticamente caro (RAUW, 2012). Existe correlação negativa entre seleção para peso corporal e desempenho imunológico, sendo que há muitos exemplos em que animais selecionados para alta eficiência de produção correm maior risco de problemas comportamentais, fisiológicos e imunológicos (RAUW et al., 1998). Segundo estes mesmo autores, uma possível explicação é que a seleção genética causa desequilíbrio homeostático dos animais, podendo causar patologias e consequências comportamentais, fisiológicas e imunológicas.

O gene SLC25A4 codifica o translocador ADP/ATP (translocador de nucleotídeo de adenina - ANT), uma proteína mitocondrial que determina a taxa de fluxo de ADP/ATP entre mitocôndria e citosol (LI et al., 1989). O gene LRP2BP foi detectado por varredura de seleção em linhas divergentes de frangos selecionados para gordura abdominal, sendo importante para a deposição de gordura e presente em regiões de QTL associadas com a característica (ZHANG et al., 2012). Este gene participa de muitos processos biológicos, como a regulação do metabolismo lipídico, neurodesenvolvimento e transporte de nutrientes e vitaminas (MAY et al., 2007).

O gene PDLIM3 possui envolvimento na organização do citoesqueleto, tornando-o candidato envolvido na textura da carne, além de ser significativamente correlacionado com o músculo esquelético em bovinos (PIÓRKOWSKA et al., 2017; REVERTER et al., 2012). Outro gene encontrado na mesma região genômica foi o SORBS2, também conhecido como ArgBP2, o qual

interage como proteína adaptadora sendo componente das fibras de estresse contendo α -actinina, que realiza a união dos filamentos de actina a linha Z das células musculares esqueléticas, sendo importante componente da linha Z dos músculos (ANEKAL; YONG; MANSER, 2015). O metabolismo muscular esquelético é positivamente correlacionado ao gasto energético basal, representando um dos principais componentes da variação do gasto energético individual (ZURLO et al., 1990).

A segunda região genômica do cromossomo 27 (16,3-17,3Mb) identificou o gene ZPF42, que codifica uma forma de proteína “*zinc finger*”, caracterizada pela coordenação e estabilização de íons de zinco em processos de trocas iônicas (KLUG; RHODES, 1987). É elemento essencial para síntese de DNA, transcrição do RNA, divisão e ativação celular, além de ser elemento indispensável para uma resposta imune normal (JOHN et al., 2010). Os genes da mesma família, TRIML1 e TRIML2, estão associados a embriogênese precoce (XU et al., 2009).

Segundo a base de dados CattleQTL, três QTL localizados na posição entre 27034490-29073969Mb sobrepueram a região do cromossomo 6 identificada neste estudo, sendo os QTL 21050, 21051 e 21052 identificados em animais Angus (ROLF et al., 2011). Segundo Nkrumah et al. (2007), o QTL mais significativo para CAR dentre 17 cromossomos associados a característica foi identificado no cromossomo 5, o QTL 4413 (17,6-17,8Mb) e está localizado a 1Mb da região identificada neste estudo. Das regiões genômicas associadas ao CAR neste estudo, apenas o cromossomo 6 apresentou sobreposição de QTL previamente relatado na característica.

O presente estudo identificou vários genes candidatos envolvidos em diferentes mecanismos fisiológicos que podem ser importantes na compreensão da variação da eficiência alimentar, incluindo respiração celular (síntese celular de ATP), função mitocondrial, metabolismo energético, taxa metabólica basal, desenvolvimento e regulação do crescimento e controle homeostático da massa corporal. Segundo Barendse et al. (2007), diferentes variantes de genes podem afetar estes processos, que demandam energia constantemente, além de ajudar a elucidar quais variações são responsáveis por identificar QTL associados ao CAR. Estes resultados corroboram com Zulkifli (2016), indicando que variações

em genes envolvidos na função mitocondrial e metabolismo energético podem influenciar o CAR.

No entanto, diferentes regiões e genes encontrados no presente estudo diferiram dos relatados em outros estudos na mesma característica, indicando existência de diferenças genéticas entre as populações dentro da mesma raça (OLIVEIRA et al., 2014; OLIVIERI et al., 2016; SANTANA et al., 2014). As diferenças podem ser explicadas por variações nas frequências alélicas, LD, densidade do chip de SNPs utilizado, metodologia e número de animais de cada estudo. De acordo com, Lobo et al. (2003) apesar do reduzido número de genearcas responsáveis pela formação do Nelore no país, há evidências de diferentes perfis genéticos entre as linhagens, variando de acordo com o critério de seleção adotado durante a formação da raça. Programas de melhoramento genético levam em consideração a linhagem do animal para realização do acasalamento, formando linhagens ou famílias diferentes dentro da raça, gerando diferenças genéticas.

Selecionar animais eficientes tem implicação na melhor rentabilidade do rebanho, no entanto, a obtenção dos dados de eficiência alimentar em larga escala envolve logística e dispêndio. Animais da mesma raça e manejo semelhante apresentam diferentes gastos totais de energia, os quais sugerem diferenças nos processos fisiológicos como bombeamento de íons, vazamento de prótons mitocondrial, atividades de desacoplamento de proteínas, hormônios tireoidianos, leptina, IGF-1 e metabolismo lipídico, em que animais eficientes apresentam maior eficiência energética (JOHNSON; FERRELL; JENKINS, 2003).

A identificação de genes em GWAS é o primeiro passo na busca por variantes causais que exercem influência sobre características de interesse econômico e podem colaborar no futuro para subseqüentes análises genéticas. Regiões genômicas associadas ao CAR foram identificadas no presente estudo, proporcionando melhor entendimento biológico da característica em animais da raça Nelore.

5.3 Distribuição da frequência alélica, efeito da substituição alélica e índice de fixação de Wright (Fst).

Distribuição da frequência alélica

A maioria dos marcadores que causam variação no fenótipo são variantes comuns ($MAF > 5\%$), ou que possuem maior efeito na variância da característica, como por exemplo, características poligênicas (origem mendeliana), como o CAR (MANOLIO et al., 2009). Segundo estes autores, se torna improvável a detecção de associações com valores de MAF abaixo de 5%, exceto em características monogênicas onde este efeito pode ser considerado grande. Os SNPs apresentam ampla distribuição de MAF, de quase monomórfico ($MAF < 0.5\%$) a muito comum ($MAF \approx 50\%$) (TABANGIN; WOO; MARTIN, 2009).

A MAF dos 466.315 SNPs foi calculada e a distribuição é apresentada na Figura 3.

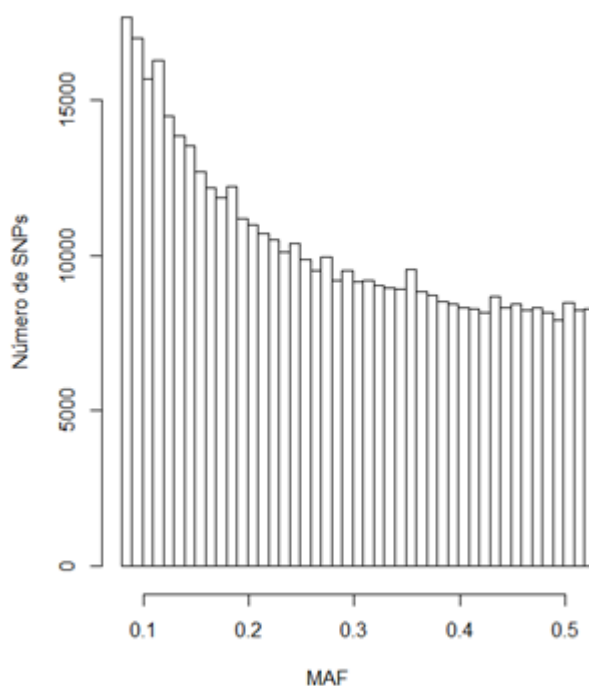


Figura 3. Histograma do alelo de frequência menor (MAF) de todos SNPs.

A média de todos marcadores revelou MAF médio de 0,25, resultado similar

foi descrito por Espigolan et al. (2013), em animais da mesma raça. No entanto, foi maior que o valor de 0,19 para animais Nelore, e comparável a valor obtido em animais taurinos, Charolês (0,26), Holstein (0,26) e Angus (0,25) (MATUKUMALLI et al., 2009). Este resultado de MAF próxima ao de animais taurinos, contraria ao descrito pelos autores citados, que relatam maior proporção de marcadores com alta MAF dos taurinos com relação aos zebuínos. Resultado semelhante também foi relatado por Khatkar et al. (2008), que encontraram valor médio de MAF de 0,28 em animais Holstein. O maior valor de MAF encontrado neste trabalho pode ser devido a utilização de chip HD (770k), que possui maior quantidade de SNPs significativos para zebu, comparado ao chip de 50k utilizado nos dois trabalhos citados.

No presente estudo, 36% dos SNPs apresentaram MAF superior a 0,3, porcentagem acima do valor de 19% descrito em zebuínos pelo THE BOVINE HAPMAP CONSORTIUM (2009). Esta diferença pode ser devido ao fato da MAF média do Consórcio considerar três raças distintas (Brahman, Nelore e Gir), e os dados somente avaliou animais Nelore. De acordo com Khatkar et al. (2008), conforme a medida do limiar da MAF aumenta, há aumento do número de pares de SNPs com frequências alélicas semelhantes, e conseqüentemente, ocorre aumento do LD.

Dos 466.315 SNPs que permaneceram após o controle de qualidade, foi aplicado o teste exato de Fisher ($p < 0,05$) nas frequências alélicas entre os grupos divergentes de CAR e obtido 104 SNPs com diferença significativa. Como regiões gênicas longas são mais propensas a conter mais SNPs em comparação com regiões mais curtas, os valores de significância do teste de Fisher foram corrigidos para o número de SNPs por gene, com base na taxa de descoberta falsa (FDR) e todos os 104 SNPs obtiveram valor q para FDR significativo ($p < 0,05$). Desta forma, foram definidos os SNPs (104) diferenciados entre os grupos contrastantes.

As frequências do menor alelo destes SNPs estão apresentadas na Figura 4.

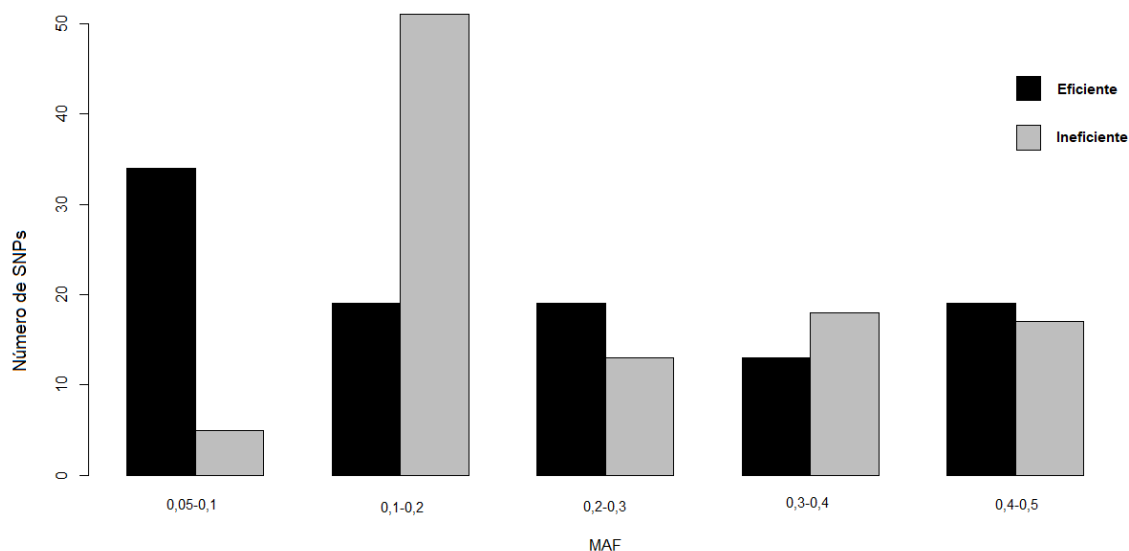


Figura 4. Distribuição da MAF dos SNPs estatisticamente diferenciados entre os grupos do CAR.

Apesar dos grupos divergentes apresentarem diferenças nas distribuições para classes de MAF, não houve diferença na média da MAF entre os grupos, que apresentaram valor igual a 0,23.

Identificar variantes menos frequentes ($MAF < 0,1$) em estudos de GWAS, auxiliam na caracterização das diferenças na arquitetura genética (CIRULLI; GOLDSTEIN, 2010). A MAF das variantes causais possui menor valor quando comparado com alelos comuns (Tabela 6). Os animais eficientes apresentaram maior frequência de $MAF < 10\%$, enquanto animais ineficientes apresentaram menor proporção da MAF na mesma classe. Frequências alélicas tendem a diferir em *loci* aleatórios quando amostras de grupos contrastantes são utilizadas (ARDLIE; LUNETTA; SEIELSTAD, 2002).

Tabela 6. Porcentagem dos SNPs diferenciados divididos em 5 classes da frequência do menor alelo (MAF) entre grupos do CAR

Grupo	MAF (%)					Média
	< 0,1	0,1-0,2	0,2-0,3	0,3-0,4	0,4-0,5	
Eficientes (187)	0,30	0,17	0,17	0,11	0,17	0,23
Ineficientes (189)	0,04	0,45	0,11	0,12	0,15	0,23
Todos (956)	0,17	0,26	0,21	0,19	0,18	0,25

Os animais dos grupos contrastantes apresentaram diferentes perfis de marcadores comuns e de baixa frequência. As diferenças nas frequências alélicas devido a ancestralidade são mais evidentes em variantes de baixa frequências (ALTSHULER et al., 2012). Variantes de baixa frequência tendem a ser de origem recente, portanto, exibem maiores níveis de diferenciação na população e mostram estratificação diferente de variantes comuns, que são tipicamente mais antigas (MATHIESON; MCVEAN, 2012).

Dos 104 SNPs que apresentaram diferenças significativas nas frequências alélicas entre os grupos contrastantes, 76 marcadores estão localizados em regiões sobrepostas de maior variância genética (maior que 1%) associadas ao CAR. Grupos de animais que apresentam diferenças genéticas exibem padrão consistentes nas frequências alélicas em muitos locais do genoma, por outro lado, se os grupos forem semelhantes geneticamente, as diferenças estão localizadas perto de regiões de suscetibilidade do fenótipo (PRITCHARD; ROSENBERG, 1999).

Apesar de 73% dos SNPs com frequências alélicas diferenciadas estarem em regiões associadas ao CAR, os outros SNPs estão localizados adjacentes a estas regiões. A relação dos SNPs está no Apêndice (Tabela A1). Apesar dos grupos divergentes apresentarem heterogeneidade alélica, os SNPs localizados adjacentes as regiões associadas ao CAR indicam que os animais são semelhantes geneticamente e pertencem a mesma estrutura populacional. Fato que ajuda nesta semelhança possivelmente se deve ao programa de melhoramento aplicar índice de seleção para escolha dos melhores reprodutores, os quais participam do teste de eficiência para obtenção do fenótipo do CAR, e finalizado o teste, os melhores animais para CAR tem o sêmen coletado e

disseminado nas fazendas participantes em programa de teste de progênie.

O número alto de alelos com menor MAF indica expansão populacional (crescimento do tamanho efetivo em relação a um ancestral comum), aumento do número de mutações (MARTH et al., 2004), e fornece informações sobre as taxas de mutação e fluxo gênico (potencial de diferenciação genética) (SLATKIN, 1985), e, ainda, demonstra existência de subdivisão da população (BARTON; SLATKIN, 1986). Deste modo, a distribuição das frequências alélicas reflete a combinação dos fatores citados, que alteram (diminuem ou aumentam) os alelos menos frequentes na população.

A distribuição da frequência da MAF dos polimorfismos foi diferente, evidenciando divergências entre os grupos contrastantes, e por serem mais associados a determinado grupo do que outro, há possibilidade de serem utilizados como marcadores informativos do CAR.

Efeito da substituição alélica

Considerando somente os 104 SNPs diferenciados ($p < 0,05$), a soma dos valores da substituição alélica representou 0,51% da variância genética aditiva do consumo alimentar residual. Segundo Visscher et al. (2017), SNPs identificados em GWAS explicam pequena fração da variância genética de características complexas, por ser análise que possui pequena capacidade de detecção dos efeitos, associando marcadores frequentes e excluindo aqueles com efeitos menores. SNPs com pequenos efeitos não são estatisticamente significativos e não são associados ao fenótipo, além disto, muitos SNPs não são detectados por possuírem baixo nível de LD entre variantes causais, dado que apresentam menor MAF (YANG et al., 2010).

Os SNPs com frequências diferentes entre os grupos estão presentes nos cromossomos 5, 11, 14, 17 e 27, estando associados a estes SNPs, 15 genes (Tabela 7).

Tabela 7. Descrição dos SNPs diferenciados entre grupos divergentes para o CAR

Cromossomo	Genes associados	Número de SNPs	%gVAR ¹
5	RASSF9	1	0,2942
11	SUCLG1	12	0,0087
14	RB1CC1, MRPL15, ATP6V1H, RGS20, TCEA1, LYPLA1, ST18	57	0,0643
17	GLT1D1	25	0,0721
27	TRIML2, TRIML1, SNX25, SORBS2, UFSP2	9	0,0744
TOTAL	-	104	0,51

¹ %gVAR, proporção da variância genética aditiva explicada pelo SNP

As diferentes proporções de variância explicada por cromossomo apresentaram variabilidade, sugerindo diferente distribuição dos marcadores que influenciam a característica. O SNP no cromossomo 5 está associado ao gene RASSF9 e explica grande proporção da variância total, demonstrando que o valor do efeito do SNP é independente do número de SNPs dentro do segmento, auxiliando a caracterizar a arquitetura genética do CAR. Este resultado tem implicações no que diz respeito a detecção de variantes adicionais e é informativo em relação a natureza das variantes que influenciam o CAR.

Os resultados indicam que há muitas regiões genômicas que influenciam o CAR, mas explicam pequena proporção da variância genética total. Identificar o particionamento da variância genética de SNPs que apresentaram diferentes frequências alélicas entre grupos divergentes auxilia compreender a relação entre a frequência dos alelos de *loci* associados ao CAR e caracterizaram melhor a arquitetura genética da característica. Os SNPs identificados em regiões de maior proporção da variância genética para o CAR são marcadores significativos e indicam presença de QTL e genes próximos envolvidos na característica em bovinos Nelore.

Índice de fixação (Fst)

A média de Fst na comparação dos grupos contrastes foi baixa, valor de 0,0004. Segundo Hartl e Clark (2006), valores de Fst variando de 0 a 0,05 indicam pequena diferenciação genética. Este baixo valor se deve possivelmente por serem animais originados da mesma população ancestral e possuir seleção recente direcional para CAR, característica utilizada na diferenciação dos grupos divergentes. As mudanças das frequências alélicas ocorrem ao longo das gerações na direção da característica desejada, sendo que a população analisada iniciou o processo de seleção para eficiência alimentar em 2010, o que explica a proximidade genética e baixo valor do Fst entre os grupos. Este resultado demonstra que provavelmente muitos alelos foram fixados em comum nos dois grupos, sendo consistente com a estrutura da população.

Foram computadas 46.911 janelas para todos os cromossomos autossômicos, apresentando média de 20 SNPs por janela, variando de 6 a 76 marcadores. O valor máximo de Fst observado para janelas de 100kb foi de 0,025. As 10 janelas com maiores valores de Fst foi utilizada como limite de significância, apresentando valores acima de 0,019. Estas janelas estavam presentes nos cromossomos 4, 6, 9, 14, 23 e 25, detectando 186 SNPs como relevantes para a seleção divergente. Embora os animais comparados sejam da mesma raça, foi possível identificar alterações nas frequências alélicas, mesmo que baixas. A média das 10 janelas foi igual a 0,022. As mudanças observadas nas frequências alélicas podem ser utilizadas para avaliar a seleção (LINNEN; HOEKSTRA, 2009). Desse modo, as alterações nas frequências alélicas observadas eventualmente podem ser originadas por eventos não relatados no passado e pela recente seleção, o qual reflete a estrutura da população estudada.

Na Figura 5, as 10 janelas com maiores valores de Fst são representadas na cor verde.

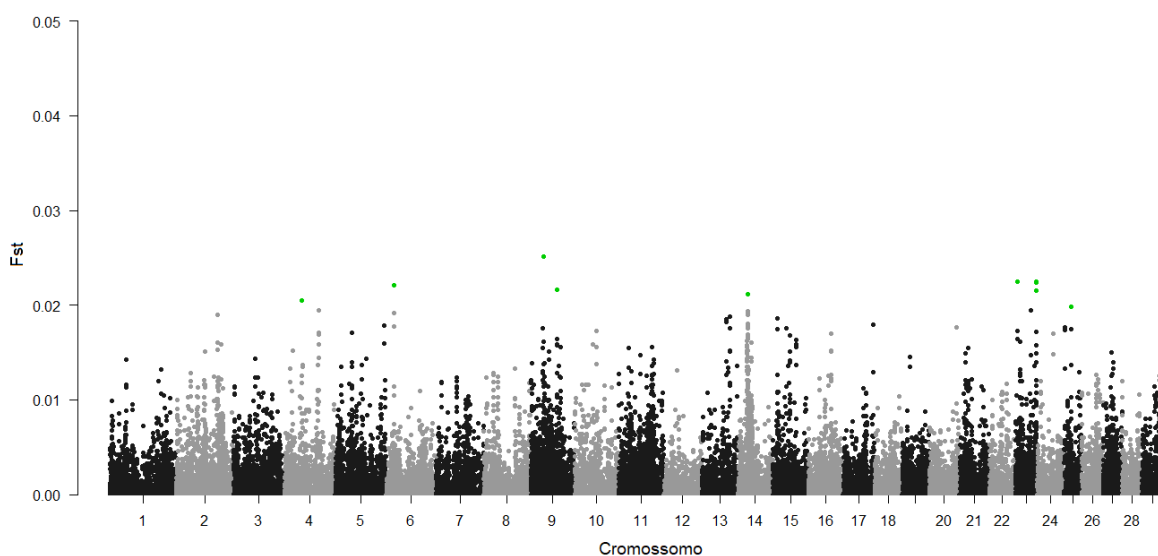


Figura 5. Gráfico de Manhattan dos valores de Fst por janelas de 100kb em 29 cromossomos autossômicos. As 10 janelas com maiores valores de Fst são representadas por pontos verdes.

Foram identificados 8 genes sobrepostos as regiões de maior Fst, e presentes em 4 cromossomos (Tabela 8). Dentre as regiões analisadas, o cromossomo 14 foi o único que apresentou SNPs com diferentes frequências alélicas ($q < 0.05$) e janela de maior valor para Fst.

Tabela 8 – Descrição das 10 janelas com maiores valores para Fst e genes sobrepostos as regiões genômicas

Cromossomo	Posição no cromossomo (bp)	Número de SNPs	Fst ¹	Genes
4	41.452.823-41.535.687	7	0,0204	-
6	18.106.489-18.192.536	25	0,0221	-
9	30.254.212-30.347.735	20	0,0251	-
	63.552.323-63.644.862	22	0,0216	ZNF292
14	25.154.132-25.246.448	21	0,0211	PENK, SDR16C6
23	5.404.541-5.499.102	13	0,0225	FAM83B
	49.959.092-50.049.075	26	0,0224	PRPF4B
	50.450.850-50.541.830	9	0,0215	NQO2
	50.503.994-50.569.741	6	0,0225	-
25	17.558.975-17.649.730	37	0,0198	IQCK, GPCR5B

¹ Valor de Fst para janelas de 100kb.

O gene ZNF292 codifica um fator de transcrição dependente do hormônio do crescimento, sendo expresso também em órgãos não endócrinos, onde sua função é desconhecida (FLYNN; HURLEY, 2006; TAKEDA et al., 2015). Este gene foi supra-regulado em folículos atrésicos de bovinos taurinos, o qual influencia a regulação do número de oócitos ovulados (HATZIRODOS et al., 2014).

O gene PENK foi localizado na região do cromossomo 14 e faz parte do sistema opióide, que influencia os eixos compostos do hipotálamo, glândula pituitária e gônada adrenal (HAWKEN et al., 2012; PETRAGLIA; IMPERATORE; CHALLIS, 2010; VUONG et al., 2010). Este gene foi associado a características de fertilidade em animais Brangus e Brahman e foi regulado positivamente em cinco tecidos (útero, endométrio, hipotálamo, hipófise e fígado) de novilhas pós-púberes (CÁNOVAS et al., 2014; FORTES et al., 2013; PETERS et al., 2013).

No cromossomo 25 o gene IQCK foi associado a ovário cístico, que é um tipo de distúrbio reprodutivo e afeta a rentabilidade no sistema de produção (GUARINI, 2017). O ovário cístico foi associado com disfunções no ciclo de Krebs, em que foi relacionado com a perda precoce de prenhez em ovelhas (ZHAO et al., 2015b).

Os genes identificados nas janelas com maiores valores de Fst estão relacionados a fertilidade e/ou desempenho reprodutivo. O cálculo do CAR é definido como a diferença entre o consumo observado e o consumo predito, sendo que o consumo predito para vacas é calculado em função do peso corporal, alterações no peso, gasto de energia na produção de leite e na ordem de parição (POTTS et al., 2015). Isto significa que crescimento, produção de leite e manutenção são incluídos para obtenção do CAR, portanto, não devem ser indiretamente afetadas pela seleção do CAR. No entanto, apesar do CAR revelar variação nos processos metabólicos básicos entre os animais, a reprodução é outra área que requer energia das vacas e não é capturada pelo modelo do CAR.

A reprodução requer pouca energia em relação a outros gastos energéticos das vacas, podendo ser mais afetada pela condição corporal do que pelo consumo alimentar e, conseqüentemente, pela eficiência alimentar (BUTLER,

2003). A associação de genes relacionadas ao desempenho reprodutivo em regiões de maior valores de *Fst* pode ser relacionado a alterações no metabolismo intermediário associado a diferentes condições corporais (deposição de gordura) e influenciar a secreção de LH e a puberdade (SCHILLO; HALL; HILEMAN, 1992). Como o CAR é relacionado a medidas de gordura, animais menos eficientes podem amadurecer sexualmente mais cedo, ou possuir sistema metabólico que inicie a função reprodutiva em uma idade mais precoce, uma vez que animais baixo CAR podem metabolizar hormônios reprodutivos em taxas diferentes e atrasar a capacidade de induzir a puberdade (SHAFFER et al., 2011).

Analisando a relação entre eficiência alimentar e desempenho reprodutivo, Arthur et al. (2005) e Blair et al. (2013), não relataram diferenças entre animais baixo e alto CAR para taxa de prenhez. Em contrapartida, Randel e Welsh (2013), em revisão, relataram que a seleção de novilhas baixo CAR pode prejudicar a eficiência reprodutiva, resultado apoiado por Damiran et al. (2018), que associaram desempenho reprodutivo reduzido com novilhas de maior eficiência alimentar. Apesar dos divergentes resultados citados entre a relação do CAR e características reprodutivas, a compreensão a longo prazo da seleção para CAR deve avaliar efeitos de seleção indireta, possivelmente negativas sobre outras características economicamente importantes, como as reprodutivas.

A estrutura da população visualizada por PCA sustenta a similaridade genética e demonstra que os indivíduos pertencem a mesma população. Os resultados não mostraram estratificação populacional, e o baixo valor de *Fst* informou sobre a baixa diferenciação entre os grupos contrastantes. O baixo valor de *Fst* indica que a frequência dos alelos dentro de cada grupo é semelhante.

Os genes localizados estão dentro das janelas com maiores valores e estão associados ao maior nível de diferenciação entre os grupos contrastantes. Considerando as janelas de *Fst* e resultados do GWAS para CAR, nenhum gene em comum foi identificado. No entanto, enquanto a média do *Fst* apresentou baixa diferenciação entre os grupos, as frequências alélicas exibiram diferenças entre os grupos. Esta diferença é a principal ressalva de que os 104 SNPs identificados estão associados mais em grupo que outro. A análise da diferenciação da frequência alélica demonstra tanto os efeitos de seleção em *loci* relevantes no

genoma e a relação com o histórico da população, como também resulta em diferenciação substancial na região, mesmo quando esta apresenta valores baixos de F_{st} (ALTSHULER et al., 2012).

O F_{st} é um dos principais estimadores para detectar assinaturas de seleção em animais. No entanto, é mais adequado para detecção de evidência de seleção que ocorreram em um passado distante (CADZOW et al., 2014). A escala de tempo a qual ocorre a seleção possui grande impacto sobre a capacidade do método em identificar a sua presença. O menor tempo de seleção para CAR da população A baixa diferenciação entre os grupos pode ser uma das razões para o baixo valor encontrado do F_{st} neste estudo.

Embora o uso de animais não relacionados permita alta diferenciação genética, os animais de programas de melhoramento genético são escolhidos para serem genotipados de acordo com sua importância no rebanho. Desse modo, estes animais podem estar relacionados e apresentarem perfil genético semelhante dentro da população. Adicionalmente, os padrões capturados da variação genotípica não são necessariamente resultado da seleção. A variação presente pode ser consequência de outros eventos ancestrais, que podem modificar as variantes e frequências haplotípicas de maneira semelhante à seleção, como a deriva genética, gargalos e expansão populacional (CADZOW et al., 2014).

O diferente padrão das frequências alélicas encontrado entre os grupos divergentes para o CAR indica que há evidências de eventos genéticos em direção ao fenótipo selecionado. Para detectar padrões de seleção deve-se levar em consideração o histórico da população, assim como o tempo da aplicação da seleção, uma vez que é necessário tempo para que a variante selecionada atinja frequência detectável na população. As regiões identificadas nas janelas de F_{st} foram diferentes das encontradas em GWAS, demonstrando diferenças de genes associados a característica e dos genes identificados pela diferenciação genética entre os grupos contrastantes. Estas regiões, associadas ao desempenho reprodutivo, demonstram quais genes possivelmente sofrem modificações para adaptação a seleção.

Considerações finais

Os resultados deste estudo fornecem informações preliminares sobre a arquitetura genética do CAR. Os genes candidatos encontrados em GWAS e em janelas de F_{st} contribuem para compreensão de como podem afetar o fenótipo e como a população se adapta a pressão seletiva. Além disto, esta população possui recente seleção para o CAR, sendo que as frequências alélicas provavelmente serão alteradas ao longo do tempo na direção da seleção para a característica desejada, diferenciando os grupos para futuramente produzir assinaturas de seleção na população e investigações na raça.

6. CONCLUSÃO

O estudo permitiu a identificação de SNPs, regiões cromossômicas e genes candidatos associados ao consumo alimentar residual (CAR) em bovinos Nelore. Os genes identificados possuem função biológica que influenciam a eficiência alimentar e devem contribuir para a formação de uma base genética em bovinos Nelore. O padrão da diferenciação genética entre os grupos contrastantes para CAR foi consistente com a estrutura da população, indicando regiões e genes que possivelmente sofrem modificações para adaptação a seleção. Os genes identificados podem fornecer informações adicionais sobre regiões QTL, e devem ser considerados em estudos de associação, genômica funcional e utilizados em programas de melhoramento genético, como ferramenta para aumentar a eficiência da seleção do CAR.

7. REFERÊNCIAS

- AGUILAR, I. et al. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score¹. **Journal of Dairy Science**, v. 93, n. 2, p. 743–752, 2010.
- AKEY, J. M. et al. Interrogating a high-density SNP map for signatures of natural selection. **Genome Research**, v. 12, p. 1805–1814, 2002.
- ALEXA, A.; RAHNENFUHRER, J. **topGO: Enrichment Analysis for Gene Ontology**. Disponível em: <<http://bioconductor.org/packages/release/bioc/html/topGO.html>>.
- ALTSHULER, D. M. et al. An integrated map of genetic variation from 1,092 human genomes. **Nature**, v. 491, n. 7422, p. 56–65, 2012.
- ANEKAL, P. V.; YONG, J.; MANSER, E. Arg kinase-binding protein 2 (ArgBP2) interaction with α -actinin and actin stress fibers inhibits cell migration. **Journal of Biological Chemistry**, v. 290, n. 4, p. 2112–2125, 2015.
- ARCHER, J. A. et al. Potential for selection to improve efficiency of feed use in beef cattle: a review. **Australian Journal of Agricultural Research**, v. 50, p. 147–161, 1999.
- ARCHER, J. A. et al. Optimum postweaning test for measurement of growth rate, feed intake, and feed efficiency in British breed cattle. The online version of this article, along with updated information and services, is located on the World Wide Web at: Optimum Postweanin. n. SEPTEMBER, p. 2024–2032, 1997.
- ARDLIE, K. G.; KRUGLYAK, L.; SEIELSTAD, M. Patterns of linkage disequilibrium in the human genome. **Nature Reviews Genetics**, v. 3, n. 4, p. 299–309, 2002.
- ARDLIE, K. G.; LUNETTA, K. L.; SEIELSTAD, M. Testing for population subdivision and association in four case-control studies. **American journal of human genetics**, v. 71, n. 2, p. 304–11, 2002.
- ARIAS, J. A. et al. A high density linkage map of the bovine genome. **BMC Genetics**, v. 10, p. 1–12, 2009.
- ARTHUR, J. P. F.; HERD, R. M. Residual feed intake in beef cattle. **Revista Brasileira de Zootecnia**, v. 37, n. spe, p. 269–279, 2008.
- ARTHUR, P. F. et al. **Maternal productivity of Angus cows divergently selected for post-weaning residual feed intake**. Australian Journal of Experimental Agriculture. **Anais...**2005
- ARTHUR, P. F.; ARCHER, J. A.; HERD, R. M. Feed intake and efficiency in beef cattle: overview of recent Australian research and challenges for the future. **Australian Journal of Experimental Agriculture**, v. 44, n. 5, p. 361, 2004.

ARTHUR, P. F.; RENAND, G.; KRAUSS, D. Genetic and phenotypic relationships among different measures of growth and feed efficiency in young Charolais bulls. **Livestock Production Science**, v. 68, n. 2–3, p. 131–139, 2001.

BALDING, D. J. A tutorial on statistical methods for population association studies. **Nature Reviews Genetics**, v. 7, n. 10, p. 781–791, 2006.

BALDWIN, R. L. et al. Manipulating metabolic parameters to improve growth rate and milk secretion. **Journal of animal science**, v. 51, n. 6, p. 1416–1428, 1980.

BALDWIN, R. L.; SAINZ, R. D. Energy Partitioning and Modeling in Animal Nutrition. **Annual Review of Nutrition**, v. 15, p. 191–211, 1995.

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

BARENDSE, W. et al. A validated whole-genome association study of efficient food conversion in cattle. **Genetics**, v. 176, n. 3, p. 1893–1905, 2007.

BARENDSE, W. et al. Genome wide signatures of positive selection: The comparison of independent samples and the identification of regions associated to traits. **BMC Genomics**, v. 10, p. 1–15, 2009.

BARTHA, T.; SAYED-AHMED, A.; RUDAS, P. Expression of leptin and its receptors in various tissues of ruminants. **Domestic Animal Endocrinology**, v. 29, n. 1, p. 193–202, 2005.

BARTON, N. H.; SLATKINT, M. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. **Heredity**, v. 56, n. 3, p. 409–415, 1986.

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

BELLMANN, O. et al. Beef versus dairy cattle: A comparison of metabolically relevant hormones, enzymes, and metabolites. **Livestock Production Science**, v. 89, n. 1, p. 41–54, 2004.

BENJAMINI, Y.; HOCHBERG, Y. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society B**, v. 57, n. 1, p. 289–300, 1995.

BIF. **Beef Improvement Federation. Guidelines For Uniform Beef Improvement Programs.(9th. Edition)**Athens, GA, 2010.

BLAIR, E. E. et al. Relationship between residual feed intake and female reproductive measurements in heifers sired by high–or low–residual feed intake Angus bulls¹. **The Professional Animal Scientist**, v. 29, n. 1, p. 46–50, 2013.

BOLORMAA, S. et al. Genome-wide association studies for feedlot and growth traits in cattle. **Journal of Animal Science**, v. 89, n. 6, p. 1684–1697, 2011.

BOLORMAA, S. et al. Detection of quantitative trait loci in *Bos indicus* and *Bos taurus* cattle using genome-wide association studies. **Genetics Selection Evolution**, v. 45, n. 1, p. 43, 2013.

BOUWMAN, A. C.; HAYES, B. J.; CALUS, M. P. L. Estimated allele substitution effects underlying genomic evaluation models depend on the scaling of allele counts. **Genetics Selection Evolution**, v. 49, n. 1, p. 1–13, 2017.

BROOKES, A. J. The essence of SNPs. **Gene**, v. 234, n. 2, p. 177–186, 1999.

BUSH, W. S.; MOORE, J. H. Chapter 11: Genome-Wide Association Studies. **PLoS Computational Biology**, v. 8, n. 12, 2012.

BUTLER, W. R. **Energy balance relationships with follicular development ovulation and fertility in postpartum dairy cows**. Livestock Production Science. **Anais...**2003

CADZOW, M. et al. A bioinformatics workflow for detecting signatures of selection in genomic data. **Frontiers in Genetics**, v. 5, n. AUG, 2014.

CÁNOVAS, A. et al. Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. **PLoS ONE**, v. 9, n. 7, 2014.

CANTOR, R. M.; LANGE, K.; SINSHEIMER, J. S. Prioritizing GWAS Results: A Review of Statistical Methods and Recommendations for Their Application. **American Journal of Human Genetics**, v. 86, n. 1, p. 6–22, 2010.

CARDOSO, D. F. et al. Genome-wide scan reveals population stratification and footprints of recent selection in Nelore cattle. **Genetics Selection Evolution**, v. 50, n. 1, p. 1–12, 2018.

CASTILHOS, A. M. et al. Eficiência alimentar em bovinos Nelore selecionados para peso pósdesmame. **Revista Brasileira de Zootecnia**, v. 39, n. 11, p. 2486–2493, 2010.

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.

CHEN, M. et al. Identification of selective sweeps reveals divergent selection between Chinese Holstein and Simmental cattle populations. **Genetics Selection Evolution**, v. 48, n. 1, p. 1–12, 2016.

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

CIRULLI, E. T.; GOLDSTEIN, D. B. Uncovering the roles of rare variants in common disease through whole-genome sequencing. **Nature Reviews Genetics**, v. 11, n. 6, p. 415–425, 2010.

COCKERHAM, C. C. Variance of Gene Frequencies. **Evolution**, v. 23, n. 1, p. 72, 1969.

CREWS, D. H. Genetics of efficient feed utilization and national cattle evaluation: A review. **Genetics and Molecular Research**, v. 4, n. 2, p. 152–165, 2005.

CROWLEY, J. J. et al. Phenotypic and genetic parameters for different measures of feed efficiency in different breeds of Irish performance-tested beef bulls. **Journal of Animal Science**, v. 88, n. 3, p. 885–894, 2010.

CULBERTSON, M. M. et al. Optimum measurement period for evaluating feed intake traits in beef cattle. **Journal of Animal Science**, v. 93, n. 5, p. 2482–2487, 2015.

DAETWYLER, H. D. et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. **Nature Genetics**, v. 46, n. 8, p. 858–865, 2014.

DAMIRAN, D. et al. Use of residual feed intake as a selection criterion on the performance and relative development costs of replacement beef heifers. **Professional Animal Scientist**, v. 34, n. 2, p. 156–166, 2018.

DEKKERS, J. C. M. Prediction of response to marker-assisted and genomic selection using selection index theory. **Journal of Animal Breeding and Genetics**, v. 124, n. 6, p. 331–341, 2007.

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

DU, F. X.; CLUTTER, A. C.; LOHUIS, M. M. Characterizing linkage disequilibrium in pig populations. **International Journal of Biological Sciences**, v. 3, n. 3, p. 166–178, 2007.

ELSIK, C. G.; TELLAM, R. L.; WORLEY, K. C. The Genome Sequence of Taurine Cattle: A window to ruminant biology and evolution. **Science (New York, N.Y.)**, v. 324, n. 5926, p. 522–528, 2009.

ERHARDT, A. et al. Replication and meta-analysis of TMEM132D gene variants in panic disorder. **Translational Psychiatry**, v. 2, n. July, 2012.

ESPIGOLAN, R. et al. Study of whole genome linkage disequilibrium in Nellore cattle. **BMC Genomics**, v. 14, n. 1, 2013.

FALCONER, D. . **Introdução a genética quantitativa**. Viçosa: [s.n.].

FALCONER, D. .; MACKAY, T. F. C. **Introduction to quantitative genetics**. Pearson ed. New York: [s.n.].

FALLON, J. H.; LESLIE, F. M. Distribution of dynorphin and enkephalin peptides in the

rat brain. **Journal of Comparative Neurology**, v. 249, n. 3, p. 293–336, 1986.

FARJALLA, Y. B. **Desempenho , características de carcaça e qualidade de carne de bovinos Nelore estratificados pela eficiência através do consumo alimentar residual**. [s.l.] Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, 2009.

FLYNN, M. P.; HURLEY, D. L. Growth hormone transcription factor ZN-16 genomic coding regions are composed of a single exon and are evolutionarily conserved in mammals. **Gene**, v. 368, n. 1–2, p. 78–83, 2006.

FORNERIS, N. S. et al. Quality control of genotypes using heritability estimates of gene content at the marker. **Genetics**, v. 199, n. 3, p. 675–681, 2015.

FORTES, M. R. S. et al. Evidence for pleiotropism and recent selection in the PLAG1 region in Australian Beef cattle. **Animal Genetics**, v. 44, n. 6, p. 636–647, 2013.

GIBB, D. J.; MCALLISTER, T. A. **The impact of feed intake and feeding behaviour of cattle on feedlot and feedbunk management**. Proc. 20th Western Nutrition Conf. **Anais...**1999Disponível em: <http://www.balancedbeef.com/extension_doc/wnc_intake.pdf>

GLASS, M.; BILLINGTON, C.; LEVINE, A. Opioids, food reward, and macronutrient selection. In: BERTHOUD, H.; SEELEY, R. (Eds.). **Neural and Metabolic Control of Macronutrient Intake**. New York: CRC Press, 2000. p. 407–423.

GODDARD, M. E.; HAYES, B. J. Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics**, v. 10, n. 6, p. 381–391, 2009.

GODOY, T. F. et al. SNP and INDEL detection in a QTL region on chicken chromosome 2 associated with muscle deposition. **Animal Genetics**, v. 46, n. 2, p. 158–163, 2015.

GOMES, R. C. et al. Feedlot performance, feed efficiency reranking, carcass traits, body composition, energy requirements, meat quality and calpain system activity in Nellore steers with low and high residual feed intake. **Livestock Science**, v. 150, n. 1–3, p. 265–273, 2012.

GUARINI, A. R. **Combining Genotypic , Phenotypic and Pedigree Information to Analyze Functional Traits in Dairy Cattle**. [s.l.] University of Guelph, 2017.

GUNSETT, F. C. Linear index selection to improve traits defined as ratios. **Journal of Animal Science**, v. 59, n. 5, p. 1185–1193, 1984.

HAPMAP, T. B. C. Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. **Science (New York, N.Y.)**, v. 324, n. April, p. 528–532, 2009.

HARTL, D. L.; CLARK, A. G. **Principles of Population Genetics**. 4. ed. Sunderland,

MA, USA: Sinauer Associates, 2006.

HATZIRODOS, N. et al. Transcriptome profiling of granulosa cells from bovine ovarian follicles during atresia. 2014.

HAWKEN, R. J. et al. Genome-wide association studies of female reproduction in tropically adapted beef cattle 1 Genome-wide association with cattle reproduction. **J. Anim. Sci. Qld Ltd**, v. 90, p. 1398–1410, 2012.

HAYES, B. J. et al. Accuracy of marker-assisted selection with single markers and marker haplotypes in cattle. **Genetical Research**, v. 89, n. 4, p. 215–220, 2007.

HAYES, B. J. et al. Accuracy of genomic breeding values in multi-breed dairy cattle populations. **Genetics Selection Evolution**, v. 41, n. 1, p. 1–9, 2009a.

HAYES, B. J. et al. A genome map of divergent artificial selection between *Bos taurus* dairy cattle and *Bos taurus* beef cattle. **Animal Genetics**, v. 40, n. 2, p. 176–184, 2009b.

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

HERD, R. M.; ARTHUR, P. F. Physiological basis for residual feed intake. **Journal of animal science**, v. 87, n. 14 Suppl, p. 64–71, 2009.

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

HERD, R. M.; ODDY, V. H.; RICHARDSON, E. C. Biological basis for variation in residual feed intake in beef cattle. 1. Review of potential mechanisms. **Australian Journal of Experimental Agriculture**, v. 44, n. 4–5, p. 423–430, 2004.

HILL, R.; WYSE, G.; ANDERSON, M. **Animal physiology**. 3th. ed. Sunderland, MA, USA: Associates, Sinauer, 2012.

HILL, W. G.; ROBERTSON, A. Linkage disequilibrium in finite populations. **TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik**, v. 38, n. 6, p. 226–31, 1968.

HOLSINGER, K. E.; WEIR, B. S. **Genetics in geographically structured populations: Defining, estimating and interpreting FST** *Nature Reviews Genetics*, 2009.

HOTOVY, S. K. et al. Variation among twin beef cattle in maintenance energy requirements. **Journal of animal science**, v. 69, n. 3, p. 940–946, 1991.

HU, Z. L.; PARK, C. A.; REECY, J. M. Developmental progress and current status of the Animal QTLdb. **Nucleic Acids Research**, v. 44, n. D1, p. D827–D833, 2016.

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

JENSEN, J. et al. Residual maximum likelihood estimation of (co) variance components in multivariate mixed linear models using average information. **Jour. Ind. Soc. Ag. Statistics**, v. 49, p. 215–236, 1997.

JOHN, E. et al. Zinc in innate and adaptive tumor immunity. **Journal of Translational Medicine**, v. 8, n. 1, p. 118, 2010.

JOHNSON, D. E.; FERRELL, C. L.; JENKINS, T. G. The history of energetic efficiency research : Where have we been and where are we going? **Journal of animal science**, v. 81, n. (E. Suppl. 1), p. E27–E38, 2003.

JOHNSON, K. A. et al. Differential expression of innate immune system genes in liver of beef cattle with divergent phenotypes for RFI. In: Oltjen J.W., Kebreab E., Lapierre H. (eds). In: **Energy and protein metabolism and nutrition in sustainable animal production**. [s.l.] Wageningen Academic Publishers, Wageningen, 2013. p. 371–372.

KARISA, B.; MOORE, S.; PLASTOW, G. Analysis of biological networks and biological pathways associated with residual feed intake in beef cattle. **Animal Science Journal**, n. February 2013, p. 374–387, 2013.

KHATKAR, M. S. et al. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. **BMC Genomics**, v. 9, 2008.

KIJAS, J. W. et al. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. **PLoS Biology**, v. 10, n. 2, 2012.

KLUG, A.; RHODES, D. Zinc fingers: A novel protein fold for nucleic acid recognition. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 52, p. 473–482, 1987.

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

KOLATH, W.; KERLEY, M. The relationship between mitochondrial respiration and residual feed intake in Angus steers. **Journal of Animal Science**, v. 84, n. March, p. 131, 2006.

KONG, R. S. G. et al. Transcriptome profiling of the rumen epithelium of beef cattle differing in residual feed intake. **BMC Genomics**, v. 17, n. 1, p. 592, 2016.

KUKUSHKIN, N. V. et al. Restricted processing of glycans by endomannosidase in mammalian cells. **Glycobiology**, v. 22, n. 10, p. 1282–1288, 2012.

KWOK, P. Y.; GU, Z. Single nucleotide polymorphism libraries: Why and how are we building them? **Molecular Medicine Today**, v. 5, n. 12, p. 538–543, 1999.

LANCASTER, P. A. et al. Relationships between residual feed intake and hepatic mitochondrial function in growing beef cattle. **Journal of Animal Science**, v. 92, n. 7, p. 3134–3141, 2014.

LI, K. et al. A human muscle adenine nucleotide translocator gene has four exons, is located on chromosome 4, and is differentially expressed. **Journal of Biological Chemistry**, v. 264, n. 24, p. 13998–14004, 1989.

LINDHOLM-PERRY, A. K. et al. A region on BTA14 that includes the positional candidate genes LYPLA1, XKR4 and TMEM68 is associated with feed intake and growth phenotypes in cattle. **Animal Genetics**, v. 43, n. 2, p. 216–219, 2012.

LOBO, R. B. et al. **PERFIL GENÉTICO DOS PRINCIPAIS TOUROS FUNDADORES DA RAÇA NELORE NA BASE DE DADOS DO PMGRN-USP**. REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA. **Anais...**2003

LOURENCO, A. A. L. et al. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. **J Anim Sci**, v. 93, n. March, p. 2653–2662, 2015.

LOURENCO, D. A. L. et al. Methods for genomic evaluation of a relatively small genotyped dairy population and effect of genotyped cow information in multiparity analyses. **Journal of Dairy Science**, v. 97, n. 3, p. 1742–1752, 2014.

LOWE, M. et al. A Value Chain Analysis of the U.S. Beef and Dairy Industries. **Social Science Research**, n. 919, p. 1–55, 2009.

LOWENSTEIN, J. M. **Methods in Enzymology: Citric Acid Cycle**. Academic P ed. New York: [s.n.].

LYNCH, M.; WALSH, B. **Genetics and Analysis of Quantitative Traits**. Sunderland, MA, USA: Sinauer Associates, 1998.

MACEACHERN, S. et al. An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (*Bos taurus*) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in Domestic ca. **BMC Genomics**, v. 10, 2009.

MANAFIAZAR, G. et al. Optimizing feed intake recording and feed efficiency estimation to increase the rate of genetic gain for feed efficiency in beef cattle. **Canadian Journal of Animal Science**, v. 97, n. 3, p. 456–465, 2017.

MANOLIO, T. A. et al. Finding the missing heritability of complex diseases. **Nature**, v. 461, n. 7265, p. 747–753, 2009.

MÁRQUEZ, G. C. et al. Quantitative trait loci with effects on feed efficiency traits in Hereford x composite double backcross populations. **Animal Genetics**, v. 40, n. 6, p. 986–988, 2009.

MARTH, G. T. et al. The Allele Frequency Spectrum in Genome-Wide Human Variation

- Three Large World Populations. **Genetics**, v. 372, n. January, p. 351–372, 2004.
- MATHIESON, I.; MCVEAN, G. Differential confounding of rare and common variants in spatially structured populations. **Nature Genetics**, v. 44, n. 3, p. 243–246, 2012.
- MATUKUMALLI, L. K. et al. Development and Characterization of a High Density SNP Genotyping Assay for Cattle. **PLoS ONE**, v. 4, n. 4, 2009.
- MAY, P. et al. The LDL receptor-related protein (LRP) family: An old family of proteins with new physiological functions. **Annals of Medicine**, v. 39, n. 3, p. 219–228, 2007.
- MCKAY, S. D. et al. Whole genome linkage disequilibrium maps in cattle. **BMC Genetics**, v. 8, p. 1–12, 2007.
- MEALE, S. J. et al. Exploration of Biological Markers of Feed Efficiency in Young Bulls. **Journal of Agricultural and Food Chemistry**, v. 65, n. 45, p. 9817–9827, 2017.
- MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, n. 4, p. 1819–1829, 2001.
- MEYER, K.; HOULE, D. Sampling Based Approximation of Confidence Intervals for Functions of Genetic Covariance Matrices. **Association for the Advancement of Animal Breeding and Genetics Conference**, v. 20, p. 523–526, 2013.
- MISZTAL, I. et al. Methods to approximate reliabilities in single-step genomic evaluation. **Journal of Dairy Science**, v. 96, n. 1, p. 647–654, 2013.
- MISZTAL, I. et al. Manual for BLUPF90 family of programs. **university of Georgia, Athens, USA**, p. 125, 2015.
- MISZTAL, I.; LEGARRA, A.; AGUILAR, I. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. **Journal of Dairy Science**, v. 92, n. 9, p. 4648–4655, 2009.
- MONTANHOLI, Y. R. Investigations on Biological Indicators for Feed Efficiency in Cattle. n. September 2016, p. 164, 2009.
- MOORE, S. S.; MUJIBI, F. D.; SHERMAN, E. L. Molecular basis for residual feed intake in beef cattle. **Journal of animal science**, v. 87, n. 14 Suppl, p. E41-47, 2009.
- MOTA, R. R. et al. Genome-wide association study and annotating candidate gene networks affecting age at first calving in Nellore cattle. **Journal of Animal Breeding and Genetics**, v. 134, n. 6, p. 484–492, 2017.
- NIELSEN, M. K. et al. Review: Life-cycle, total-industry genetic improvement of feed efficiency in beef cattle: Blueprint for the Beef Improvement Federation¹¹The development of this commentary was supported by the Beef Improvement Federation. **The Professional Animal Scientist**, v. 29, n. 6, p. 559–565, 2013.

NISHIMURA, I. et al. RB1CC1 protein suppresses type II collagen synthesis in chondrocytes and causes dwarfism. **Journal of Biological Chemistry**, v. 286, n. 51, p. 43925–43932, 2011.

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

NKRUMAH, J. D. et al. Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. **Journal of Animal Science**, v. 85, n. 12, p. 3170–3181, 2007.

OLIVEIRA, P. S. N. et al. Identification of genomic regions associated with feed efficiency in Nelore cattle. **BMC genetics**, v. 15, n. 1, p. 100, 2014.

OLIVIERI, B. F. et al. Genomic Regions Associated with Feed Efficiency Indicator Traits in an Experimental Nellore Cattle Population. **Plos One**, v. 11, n. 10, p. e0164390, 2016.

PASANDIDEH, M.; RAHIMI-MIANJI, G.; GHOLIZADEH, M. A genome scan for quantitative trait loci affecting average daily gain and Kleiber ratio in Baluchi Sheep. **Journal of Genetics**, v. 97, n. 2, p. 493–503, 2018.

PEREIRA, G. L. et al. Genomic regions associated with performance in racing line of Quarter Horses. **Livestock Science**, 2018.

PÉREZ O'BRIEN, A. M. et al. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. **Genetics Selection Evolution**, v. 46, n. 1, p. 1–14, 2014.

PETERS, S. O. et al. Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in brangus heifers. **Journal of Animal Science**, v. 90, n. 10, p. 3398–3409, 2012.

PETERS, S. O. et al. Heritability and bayesian genome-wide association study of first service conception and pregnancy in Brangus heifers. **Journal of Animal Science**, v. 91, n. 2, p. 605–612, 2013.

PETRAGLIA, F.; IMPERATORE, A.; CHALLIS, J. R. G. **Neuroendocrine mechanisms in pregnancy and parturition** *Endocrine Reviews*, 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, 2017.

POTTS, S. B. et al. Residual feed intake is repeatable for lactating Holstein dairy cows fed high and low starch diets. **Journal of Dairy Science**, v. 98, n. 7, p. 4735–4747, 2015.

PRICE, A. L. et al. Principal components analysis corrects for stratification in genome-

- wide association studies. **Nature Genetics**, v. 38, n. 8, p. 904–909, 2006.
- PRITCHARD, J. K.; PRZEWORSKI, M. Linkage Disequilibrium in Humans: Models and Data. **The American Journal of Human Genetics**, v. 69, n. 1, p. 1–14, 2001.
- PRITCHARD, J. K.; ROSENBERG, N. A. Use of Unlinked Genetic Markers to Detect Population Stratification in Association Studies. **The American Journal of Human Genetics**, v. 65, n. 1, p. 220–228, 1999.
- PURCELL, S. et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. **The American Journal of Human Genetics**, v. 81, n. 3, p. 559–575, 2007.
- QANBARI, S. et al. Classic Selective Sweeps Revealed by Massive Sequencing in Cattle. **PLoS Genetics**, v. 10, n. 2, 2014.
- R DEVELOPMENT CORE TEAM. **R: A Language and Environment for Statistical Computing** Vienna, Austria, 2009.
- RANDEL, R. D.; WELSH, J. H. Joint Alpharma-beef Species Symposium: Interactions of feed efficiency with beef heifer reproductive development. **Journal of Animal Science**, v. 91, n. 3, p. 1323–1328, 2013.
- RAUW, W. M. et al. Undesirable side effects of selection for high production efficiency in farm animals: a review. **Livestock Production Science**, v. 56, n. 1, p. 15, 1998.
- RAUW, W. M. Immune response from a resource allocation perspective. **Frontiers in Genetics**, v. 3, n. DEC, p. 1–14, 2012.
- REEDS, P. J. et al. Amino acid metabolism and the energetics of growth. **Archives of Animal Nutrition**, v. 51, n. 2–3, p. 187–197, 1998.
- REVERTER, A. et al. A gene coexpression network for bovine skeletal muscle inferred from microarray data 2nd International Symposium on Animal Functional Genomics A gene coexpression network for bovine skeletal muscle inferred from microarray data. p. 76–83, 2012.
- RICHARDSON, E. C. et al. Australian Journal of Experimental Agriculture. **Australian Journal of Experimental Agriculture**, v. 41, p. 1065–1072, 2001.
- RICHARDSON, E. C. et al. Blood cell profiles of steer progeny from parents selected for and against residual feed intake. **Australian Journal of Experimental Agriculture**, v. 42, n. 7, p. 901–908, 2002.
- RICHARDSON, E. C. et al. Metabolic differences in Angus steers divergently selected for residual feed intake. **Australian Journal of Experimental Agriculture**, v. 44, n. 5, p. 441, 2004.
- RICHARDSON, E. C. A; HERD, R. M. B. Biological basis for variation in residual feed intake in beef cattle . 2 . Synthesis of results following divergent selection Cooperative

Research Centre for Cattle and Beef Quality . **Australian Journal of Experimental Agriculture**, v. 44, p. 431–440, 2004.

RIVERA, H. et al. Marked mitochondrial DNA depletion associated with a novel SUCLG1 gene mutation resulting in lethal neonatal acidosis, multi-organ failure, and interrupted aortic arch. **Mitochondrion**, v. 10, n. 4, p. 362–368, 2010.

ROLF, M. M. et al. Genome-wide association analysis for feed efficiency in Angus cattle. **Animal Genetics**, v. 43, n. 4, p. 367–374, 2011.

ROSA, A. DO N. F.; MENEZES, G. R. DE O.; SULEIMAN, K. **PECUÁRIA DE CORTE BRASILEIRA O VALOR DO ZEBU**, 2016. Disponível em: <<http://www.agroanalysis.com.br/index.php/7/2016/mercado-negocios/pecuaria-de-corte-brasileira-o-valor-do-zebu>>

ROUZIER, C. et al. The severity of phenotype linked to SUCLG1 mutations could be correlated with residual amount of SUCLG1 protein. **Journal of Medical Genetics**, v. 47, n. 10, p. 670–676, 2010.

RUI, L. Energy Metabolism in the Liver GS GAP. v. 4, n. January, p. 177–197, 2014.

SANTANA, M. H. et al. Genome-wide association analysis of feed intake and residual feed intake in Nellore cattle. **BMC Genetics**, v. 15, n. 1, p. 21, 2014.

SARGOLZAEI, M.; CHESNAIS, J. P.; SCHENKEL, F. S. A new approach for efficient genotype imputation using information from relatives. 2014.

SCHILLO, K. K.; HALL, J. B.; HILEMAN, S. M. **Effects of nutrition and season on the onset of puberty in the beef heifer.** **Journal of animal science**, 1992.

SEAL, C. J.; REYNOLDS, C. K. Nutritional implications of gastrointestinal and liver metabolism in ruminants. **Nutrition Research Reviews**, v. 6, p. 185–208, 1993.

SERÃO, N. V. L. et al. Integration of polygenic and individual SNP effects in genome-wide association analyses. **2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops, BIBMW 2011**, p. 985–987, 2011.

SHAFFER, K. S. et al. Residual feed intake, body composition, and fertility in yearling beef heifers. **Journal of Animal Science**, v. 89, n. 4, p. 1028–1034, 2011.

SHANADO, Y. et al. Lysophospholipase I identified as a ghrelin deacylation enzyme in rat stomach. **Biochemical and Biophysical Research Communications**, v. 325, n. 4, p. 1487–1494, 2004.

SHERMAN, E. L. et al. Identification of polymorphisms influencing feed intake and efficiency in beef cattle. **Animal Genetics**, v. 39, n. 3, p. 225–231, 2008.

SHERMAN, E. L.; NKRUHMAH, J. D.; MOORE, S. S. Whole genome single nucleotide polymorphism associations with feed intake and feed efficiency in beef cattle. **Journal of Animal Science**, v. 88, n. 1, p. 16–22, 2010.

SLATKIN, M. Rare alleles as indicators of gene flow. **Evolution**, v. 39, n. 1, p. 53–65, 1985.

SNELLING, W. M. et al. BREEDING AND GENETICS SYMPOSIUM: Networks and pathways to guide genomic selection. **Journal of Animal Science**, v. 91, p. 537–552, 2013.

SOBRINHO, T. L. et al. Residual feed intake and relationships with performance of Nelore cattle selected for post weaning weight 1 Consumo alimentar residual e relações com o desempenho de bovinos Nelore selecionados para peso pós-desmame. **Revista Brasileira de Zootecnia**, v. 40, n. 4, p. 929–937, 2011.

STRANDÉN, I.; GARRICK, D. J. Technical note: Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. **Journal of Dairy Science**, v. 92, n. 6, p. 2971–2975, 2009.

SUZUKI, K. et al. Genetic correlations among carcass cross-sectional fat area ratios, production traits, intramuscular fat, and serum leptin concentration in Duroc pigs. **Journal of Animal Science**, v. 87, n. 7, p. 2209–2215, 2009.

SWANSON, K.; MILLER, S. Factors regulating feed efficiency and nutrient utilization in beef cattle. In: FRANCE, J.; KEBREAB, E. (Eds.). . **Mathematical modelling in animal nutrition**. Cambridge: [s.n.]. p. 419–441.

TAKEDA, H. et al. Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression. **Nature Genetics**, v. 47, n. 2, p. 142–150, 2015.

TAKENOYA, F. et al. Neuropeptide W: A key player in the homeostatic regulation of feeding and energy metabolism? **Annals of the New York Academy of Sciences**, v. 1200, p. 162–169, 2010.

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. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC genomics**, v. 16, n. 1, p. 242, 2015.

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

USDA. Livestock and Poultry: World Markets and Trade. **United States Department of Agriculture. Foreign Agricultural Service**, p. 27, 2017.

VANRADEN, P. M. Efficient Methods to Compute Genomic Predictions. **Journal of Dairy Science**, v. 91, n. 11, p. 4414–4423, 2008.

VARONA, L. et al. Non-additive Effects in Genomic Selection. v. 9, n. March, p. 1–12, 2018.

VISSCHER, P. M. et al. 10 Years of GWAS Discovery : Biology , Function , and Translation. **The American Journal of Human Genetics**, v. 101, n. 1, p. 5–22, 2017.

VITEZICA, Z. G. et al. Bias in genomic predictions for populations under selection. **Genetics Research**, v. 93, n. 5, p. 357–366, 2011.

VITEZICA, Z. G.; VARONA, L.; LEGARRA, A. On the additive and dominant variance and covariance of individuals within the genomic selection scope. **Genetics**, v. 195, n. 4, p. 1223–1230, 2013.

VUONG, C. et al. **The effects of opioids and opioid analogs on animal and human endocrine systems** **Endocrine Reviews**, 2010.

WANG, H. et al. Genome-wide association mapping including phenotypes from relatives without genotypes. **Genetics Research**, v. 94, n. 2, p. 73–83, 2012.

WANG, H. et al. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. **Frontiers in Genetics**, v. 5, n. MAY, p. 1–10, 2014.

WANG, W. Y. S. et al. Genome-wide association studies: Theoretical and practical concerns. **Nature Reviews Genetics**, v. 6, n. 2, p. 109–118, 2005.

WANG, Z. et al. Test duration for growth, feed intake, and feed efficiency in beef cattle using the GrowSafe System. **Journal of Animal Science**, v. 84, n. 9, p. 2289–2298, 2006.

WATKINS, W. S. et al. Linkage disequilibrium patterns vary with chromosomal location: a case study from the von Willebrand factor region. **Am J Hum Genet**, v. 55, n. 2, p. 348–355, 1994.

WATSON, S. J. et al. Comparison of the distribution of dynorphin systems and enkephalin systems in brain. **Science**, v. 218, n. 4577, p. 1134–1136, 1982.

WEABER et al. National Program for Genetic Improvement of Feed Efficiency in Beef Cattle Methane Emissions. p. 2013, 2013.

WEBSTER, A. J. F. The energetic efficiency of growth. **Livestock Production Science**, v. 7, p. 243–252, 1980.

WEIR, B. .; COCKERHAM, C. C. Estimating F-Statistics for the Analysis of Population Structure Author (s): B . S . Weir and C . Clark Cockerham Reviewed work (s): Published by: Society for the Study of Evolution Stable URL : <http://www.jstor.org/stable/2408641> . **Evolution**, v. 38, n. 6, p. 1358–1370, 1984.

WEIR, B. S.; CAROLINA, N.; HILL, W. G. Estimating F-S Tatistics. **Annals of human genetics**, v. 36, p. 721–750, 2002.

WENG, L. et al. SNP-based pathway enrichment analysis for genome-wide association studies. **BMC Bioinformatics**, v. 12, p. 99, 2011.

WENG, Z. et al. Genome-wide association study of growth and body composition traits in Brangus beef cattle. **Livestock Science**, v. 183, p. 4–11, 2016.

WRIGHT, S. Genetical structure of populations. **Nature**, v. 166, n. 4215, p. 247–249, 1950.

XU, X. et al. DNaseI hypersensitivity at gene-poor, FSH dystrophy-linked 4q35.2. **Nucleic Acids Research**, v. 37, n. 22, p. 7381–7393, 2009.

YANG, J. et al. Common SNPs explain a large proportion of heritability for human height. **Nature genetics**, v. 42, n. 7, p. 565–569, 2010.

YANG, J. et al. A Scalable Bayesian Method for Integrating Functional Information in Genome-wide Association Studies. **American Journal of Human Genetics**, v. 101, n. 3, p. 404–416, 2017.

ZENGER, K. R. et al. Genome-wide genetic diversity of Holstein Friesian cattle reveals new insights into Australian and global population variability, including impact of selection. **Animal Genetics**, v. 38, n. 1, p. 7–14, 2006.

ZHANG, H. et al. A genome-wide scan of selective sweeps in two broiler chicken lines divergently selected for abdominal fat content. **BMC Genomics**, v. 13, n. 1, p. 704, 2012.

ZHANG, Z. et al. Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. **PLoS ONE**, v. 5, n. 9, p. 1–8, 2010.

ZHAO, F. et al. Detection of selection signatures in dairy and beef cattle using high-density genomic information. **Genetics Selection Evolution**, v. 47, n. 1, p. 1–12, 2015a.

ZHAO, H. et al. Comparative analysis between endometrial proteomes of pregnant and non-pregnant ewes during the peri-implantation period. **Journal of Animal Science and Biotechnology**, v. 6, n. 1, p. 1–14, 2015b.

ZIEBA, D. A.; AMSTALDEN, M.; WILLIAMS, G. L. Regulatory roles of leptin in reproduction and metabolism: A comparative review. **Domestic Animal Endocrinology**, v. 29, n. 1, p. 166–185, 2005.

ZIMIN, A. V. et al. A whole-genome assembly of the domestic cow, *Bos taurus*. **Genome Biology**, v. 10, n. 4, 2009.

ZULKIFLI, N. A. **Molecular Genetics of residual feed intake and mitochondrial** Adelaide, AUT, 2016.

ZURLO, F. et al. Skeletal muscle metabolism is a major determinant of resting energy expenditure. **Journal of Clinical Investigation**, v. 86, n. 5, p. 1423–1427, 1990.

8. APÊNDICE

Tabela A1. Descrição dos 104 SNPs com frequências alélicas diferentes entre grupos divergentes para o CAR em bovinos Nelore

SNPs	Posição (BP)	Gene associado	Cromossomo	%gVAR ¹	FDR ²
BovineHD0500004575	15.574.173	RASSF9	5	2,94E-01	0,034
BovineHD1100014513*, BovineHD1100014518*, BovineHD1100014519*, BovineHD1100014522*, BovineHD1100014523*, BovineHD1100014533*, BovineHD1100014543*, BovineHD1100014545*, BovineHD1100014875 , BovineHD1100014877 , BovineHD1100031620*, UA-IFASA-8854*	49.442.847 49.462.576 49.463.811 49.470.755 49.471.450 49.483.312 49.508.933 49.516.931 50.449.041 50.450.481 49.497.598 49.473.033	SUCLG1	11	1,22E+09 4,18E+09 4,18E+09 9,12E+06 1,33E+07 1,42E+07 1,00E-03 1,47E+09 3,81E-03 3,81E-03 6,15E+08 1,33E+07	0,032 0,012 0,012 0,032 0,032 0,016 0,049 0,018 0,049 0,049 0,035 0,032
ARS-BFGL-NGS-104268*, ARS-BFGL-NGS-34390 , BovineHD1400006532*, BovineHD1400006533*, BovineHD1400006534*, BovineHD1400006535*, BovineHD1400006540*, BovineHD1400006541*, BovineHD1400006544*, BovineHD1400006546*, BovineHD1400006547*, BovineHD1400006583 , BovineHD1400006584 , BovineHD1400006588 , BovineHD1400006589 , BovineHD1400006590 , BovineHD1400006591 , BovineHD1400006593 , BovineHD1400006594 , BovineHD1400006595 , BovineHD1400006597 , BovineHD1400006598 , BovineHD1400006599 , BovineHD1400006600 , BovineHD1400006601 , BovineHD1400006669 , BovineHD1400006706 , BovineHD1400006707 , BovineHD1400006708 , BovineHD1400006715 , BovineHD1400006716 , BovineHD1400006826 , BovineHD1400006833 , BovineHD1400006841 , BovineHD1400006842 , BovineHD1400006843 ,	24.057.354 22.823.936 22.670.121 22.671.890 22.673.753 22.680.469 22.697.330 22.700.051 22.710.693 22.713.285 22.714.764 22.826.411 22.828.573 22.842.406 22.845.413 22.846.824 22.848.436 22.853.710 22.856.194 22.859.987 22.866.120 22.868.017 22.869.024 22.869.966 22.870.488 23.067.943 23.153.952 23.155.836 23.158.403 23.171.958 23.174.074 23.516.712 23.541.960 23.566.600 23.569.113 23.574.109	RB1CC1, MRPL15, ATP6V1H, RGS20, TCEA1, LYPLA1, ST18	14	2,70E+08 1,03E-04 2,06E-04 6,97E+09 1,44E-04 3,13E-04 1,50E-04 4,88E+09 8,03E+09 9,99E+09 1,92E-04 1,84E-04 2,60E-04 5,03E+09 5,03E+09 5,03E+09 5,79E+09 5,79E+09 5,79E+09 1,31E-04 5,79E+09 5,79E+09 5,79E+09 4,87E-04 4,87E-04 2,72E-04 5,05E-04 5,05E-04 5,05E-04 4,61E-04 6,28E-04 6,62E-03 8,68E-03 1,08E-04 1,88E-04 7,98E-03	0,043 0,005 0,011 0,004 0,004 0,004 0,004 0,006 0,004 0,004 0,010 0,046 0,005 0,005 0,005 0,005 0,005 0,005 0,005 0,005 0,005 0,005 0,005 0,030 0,030 0,038 0,049 0,049 0,049 0,049 0,049 0,046 0,046 0,008 0,008 0,049

BovineHD1400006849,	23.596.595			1,90E-04	0,008
BovineHD1400006857,	23.623.076			1,90E-04	0,006
BovineHD1400006858,	23.625.695			1,29E-04	0,006
BovineHD1400006859,	23.628.265			1,29E-04	0,006
BovineHD1400006860,	23.629.963			2,21E-02	0,019
BovineHD1400006861,	23.636.639			1,04E-04	0,006
BovineHD1400006864,	23.653.415			1,79E-04	0,009
BovineHD1400006868,	23.661.927			1,90E-04	0,009
BovineHD1400006869,	23.664.267			3,53E-03	0,046
BovineHD1400006872,	23.678.510			1,79E-04	0,006
BovineHD1400006885,	23.725.226			1,37E-03	0,048
BovineHD1400006972*,	24.049.812			6,52E+09	0,012
BovineHD1400006975*,	24.053.137			2,70E+08	0,043
BovineHD1400006980*,	24.072.137			4,36E+09	0,012
BovineHD1400006981*,	24.074.220			4,36E+09	0,012
BovineHD1400006984*,	24.082.748			1,28E-04	0,012
BovineHD1400024431,	22.837.770			3,02E-03	0,011
BovineHD1400024432,	23.532.041			2,47E-04	0,046
BovineHD4100011284,	22.833.434			7,61E+09	0,006
BovineHD4100011291*,	23.996.288			2,43E-03	0,003
BovineHD4100011293*,	24.064.743			4,90E+09	0,012
BovineHD1700013907,	49.077.618			6,28E+08	0,018
BovineHD1700013909,	49.085.516			8,83E+08	0,018
BovineHD1700013958,	49.203.340			1,30E-04	0,010
BovineHD1700013959,	49.204.007			3,89E-03	0,010
BovineHD1700013962,	49.211.218			8,03E-03	0,015
BovineHD1700013963,	49.214.707			1,56E-03	0,012
BovineHD1700013966,	49.222.376			2,18E-02	0,010
BovineHD1700013970,	49.235.164			7,80E-03	0,015
BovineHD1700013986,	49.277.304			9,07E-04	0,019
BovineHD1700013988,	49.279.041			9,07E-04	0,019
BovineHD1700013989,	49.280.631			9,41E-04	0,019
BovineHD1700014028,	49.445.471	GLT1D1		3,28E-04	0,037
BovineHD1700014029,	49.449.632		17	3,34E-03	0,030
BovineHD1700014030,	49.454.158			2,83E-04	0,030
BovineHD1700014031,	49.459.648			2,82E-03	0,030
BovineHD1700014032,	49.465.580			2,83E-03	0,030
BovineHD1700014033,	49.467.264			7,56E-04	0,023
BovineHD1700014036,	49.473.035			7,56E-04	0,023
BovineHD1700014044,	49.503.308			1,20E-03	0,030
BovineHD1700014059,	49.547.925			1,87E+09	0,042
BovineHD1700014080,	49.591.613			8,56E-04	0,042
BovineHD1700014088,	49.629.443			5,57E-03	0,042
BovineHD1700014090,	49.634.201			5,57E-03	0,042
BovineHD1700014100,	49.648.148			1,73E-03	0,049
BovineHD1700014101,	49.652.622			3,03E+08	0,042
ARS-BFGL-NGS-102850,	14.931.898			5,55E-03	0,003
BovineHD2700004265,	14.660.311	TRIML2,		7,27E+04	0,027
BovineHD2700004277,	14.729.479	TRIML1,		1,10E+09	0,042
BovineHD2700004327,	14.944.280	SNX25,		6,56E-03	0,046
BovineHD2700004453*,	15.346.086	SORBS2,	27	1,19E+09	0,010
BovineHD2700004868,	16.708.041	UFSP2		1,66E-02	0,005
BovineHD2700004869,	16.708.959			1,59E-02	0,005
BovineHD2700013386,	16.737.116			1,54E-02	0,005
BovineHD4100018224,	16.710.960			1,44E-02	0,005

*SNPs que não estão localizados sobrepostos nas regiões identificadas em GWAS para CAR no presente estudo, ao contrário do que estão em negrito.