

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS DE JABOTICABAL

PARÂMETROS GENÉTICOS E IDENTIFICAÇÃO DE
REGIÕES REGULATÓRIAS PARA CARACTERÍSTICAS
PONDERAIS EM BOVINOS NELORE

Miller de Jesus Teodoro

Zootecnista

2023

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CAMPUS DE JABOTICABAL

PARÂMETROS GENÉTICOS E IDENTIFICAÇÃO DE
REGIÕES REGULATÓRIAS PARA CARACTERÍSTICAS
PONDERAIS EM BOVINOS NELORE

Discente: Miller de Jesus Teodoro

Orientador: Prof. Dr. Henrique Nunes de Oliveira

Coorientadora: Profa. Dra. Amanda Marchi Maiorano

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Campus de Jaboticabal, como parte das exigências para obtenção do título de Mestre em Zootecnia.

2023

T314p Teodoro, Miller de Jesus
 Parâmetros genéticos e identificação de regiões regulatórias
 para características ponderais em bovinos Nelore / Miller de Jesus
 Teodoro. -- Jaboticabal, 2023
 63 f. : il., tabs.

 Dissertação (mestrado) - Universidade Estadual Paulista
 (Unesp),
 Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
 Orientador: Henrique Nunes Oliveira
 Coorientadora: Amanda Marchi Maiorano

 1. Bovinos de corte. 2. Genes. 3. Peso corporal. 4. Polimorfismo
 de nucleotídeo único. I. Título.

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

Essa ficha não pode ser modificada.

Impacto potencial desta pesquisa

Esta dissertação aborda regiões regulatórias do cromossomo 14 do genoma de bovinos Nelore associadas com características ponderais. Almeja-se que seus resultados auxiliem a compreensão da contribuição dessas regiões na variância explicada dos pesos ao nascimento, desmama e sobreano, assim como na identificação de variantes para serem priorizadas em estudos de predição genômica nessas características.

Potencial impact of this research

This research addresses regulatory regions of chromosome 14 of the genome of Nellore cattle associated with weight traits. It is hoped that its results will help understanding the contribution of these regions in the explained variance of weights at birth, weaning and yearling, as well as identifying variants to be prioritized in genomic prediction studies on these traits.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: PARÂMETROS GENÉTICOS E IDENTIFICAÇÃO DE REGIÕES REGULATÓRIAS PARA CARACTERÍSTICAS PONDERAIS EM BOVINOS NELORE

AUTOR: MILLER DE JESUS TEODORO

ORIENTADOR: HENRIQUE NUNES DE OLIVEIRA

COORIENTADORA: AMANDA MARCHI MAIORANO

Aprovado como parte das exigências para obtenção do Título de Mestre em Zootecnia, especialidade: Genética e Melhoramento dos Animais Domésticos pela Comissão Examinadora:

Prof. Dr. HENRIQUE NUNES DE OLIVEIRA (Participação Virtual)
Departamento de Zootecnia / FCAV UNESP Jaboticabal

Documento assinado digitalmente
gov.br HENRIQUE NUNES DE OLIVEIRA
Data: 08/03/2023 15:22:47-0300
Verifique em <https://verificador.itl.br>

Prof.Dr. FERNANDO SEBASTIAN BALDI REY (Participação Virtual)
Departamento de Zootecnia / FCAV UNESP Jaboticabal

Documento assinado digitalmente
gov.br FERNANDO SEBASTIAN BALDI REY
Data: 08/03/2023 07:50:53-0300
Verifique em <https://verificador.itl.br>

Dr. GABRIEL SOARES CAMPOS (Participação Virtual)
Purdue University / West Lafayette/Indiana/EUA

Documento assinado digitalmente
gov.br GABRIEL SOARES CAMPOS
Data: 08/03/2023 12:03:35-0300
Verifique em <https://verificador.itl.br>

Jaboticabal, 28 de fevereiro de 2023

DADOS CURRICULARES DO AUTOR

Miller de Jesus Teodoro – nascido na cidade de São Paulo, São Paulo, no dia cinco de abril de 1995, filho de Antônio Teodoro e Nailde Carolina de Jesus Teodoro. Em 2014 ingressou no curso de Bacharelado em Zootecnia no Instituto Federal de Ciência, Educação e Tecnologia de Mato Grosso – IFMT, Campus São Vicente. Realizou estágio extracurricular no Setor de Ovinocultura e na Fábrica de Ração do campus, entre 2015 e 2017. Participou do Programa Institucional de Bolsas de Iniciação Científica, com um projeto de alimentação alternativa para ovinos confinados, com fomento da Fundação de Amparo à Pesquisa do Estado de Mato Grosso - FAPEMAT, entre 2017 e 2018. Em 2018 cumpriu o estágio obrigatório na fazenda Zohar, na cidade de Vila Bela da Santíssima Trindade, Mato Grosso, atuando na cria, recria e engorda de Bovinos de Corte. Em 2019 obteve o título de Bacharel em Zootecnia. Em 2020, iniciou o Mestrado Acadêmico no Programa de Pós-Graduação em Zootecnia na Universidade Estadual Paulista - UNESP, no Campus de Jaboticabal, como bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, sob orientação do Prof. Dr. Henrique Nunes de Oliveira.

“New perceptions keep us inspired.”

“The Quantum Enigma – Kingdom of Heaven Part II” - Epica

AGRADECIMENTOS

Agradeço aos meus pais, por entenderem e darem suporte à minha jornada. Sem o apoio deles, eu não poderia ter seguido em frente.

Agradeço a todos os professores que cruzaram meu caminho, seja na graduação ou na pós-graduação, e que fizeram parte da minha formação. Em especial aos professores Doutor Henrique Leal Perez e Doutor Affonso Amaral Dalla Libera, por todos os ensinamentos de vida e pela amizade que construímos.

Agradeço ao Professor Doutor Henrique Nunes de Oliveira, por ter me aceitado como seu orientado, depositando sua confiança em mim e compartilhando seu conhecimento. Agradeço a orientação, paciência, gentileza e empatia que teve comigo desde que o conheci.

Agradeço à minha coorientadora Professora Doutora Amanda Marchi Maiorano, que me auxiliou durante toda minha trajetória na pós-graduação, me aconselhando na dissertação e nos meus passos futuros.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pelo fomento e auxílio que tornou possível a dedicação exclusiva a este estudo.

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

SUMÁRIO

	Página
RESUMO	iii
ABSTRACT	iv
1. CAPÍTULO 1 – Considerações gerais	1
1.1 Introdução	1
1.2 Objetivos	2
1.2.1 Objetivo geral	2
1.2.2 Objetivos específicos	2
1.3 Revisão de literatura	3
1.3.1 Seleção artificial para características de crescimento	3
1.3.2 Marcadores moleculares	5
1.3.3 QTLs e o cromossomo 14	6
1.3.4 A importância das classes funcionais nos estudos de GWAS	7
1.4 Referências bibliográficas	10
2. CAPÍTULO 2 – Estimação dos parâmetros genéticos, estudo de associação genômica ampla para características ponderais e variância explicada por regiões codificantes e regulatórias do cromossomo 14 de bovinos Nelore	15
2.1. Introdução	16
2.2 Metodologia	17
2.2.1 Animais e fenótipos	17
2.2.2 Dados de genótipos e controle de qualidade	19
2.2.3 Estimativas de parâmetros genéticos	20
2.2.4 Estudo de associação genômica ampla conforme abordagem WssGWAS ...	21
2.2.5 Anotação de genes e mapeamento de QTLs	24
2.2.6 Classes funcionais	25
2.2.7 Estimação da variância explicada pelas classes funcionais pelo método GBLUP com abordagem REML	27
2.3. Resultados e discussão	28
2.3.1 Parâmetros genéticos	28
2.3.2 GWAS	30
2.3.3 Detalhamento dos genes candidatos e QTLs	34
2.3.4 Variâncias explicadas pela classe funcional	37
2.3.4.1 Similaridades entre as classes anotadas	37

2.3.4.2 Frequências alélicas	39
2.3.4.3 Classes funcionais	42
2.4 Conclusão.....	45
2.5 Referências bibliográficas	46
APÊNDICE	51

PARÂMETROS GENÉTICOS E IDENTIFICAÇÃO DE REGIÕES REGULATÓRIAS PARA CARACTERÍSTICAS PONDERAIS EM BOVINOS NELORE

RESUMO – Características ponderais têm sido amplamente usadas na seleção de bovinos Nelore no Brasil, estando diretamente associadas com a produtividade e lucratividade dos rebanhos. Estudos de associação genômica ampla (GWAS) tem sido utilizados para investigar variantes presentes no genoma bovino que estão associadas a fenótipos de interesse. Entretanto, devido algumas limitações do GWAS, a anotação de classes funcionais para identificação de regiões candidatas pode ser usada para aumentar o poder de detecção do GWAS. Considerando os poucos estudos de classes funcionais associadas com características ponderais na raça Nelore, os objetivos deste estudo foram realizar GWAS para identificar regiões do cromossomo 14 associadas com características ponderais de bovinos da raça Nelore, e verificar quais classes funcionais explicam maior proporção da variância, buscando auxiliar o processo de seleção em características de crescimento. Para isso, foi empregado um banco de dados de genótipos e fenótipos para as características peso ao nascimento (PN), peso a desmama (PD) e peso ao sobreano (PS). Duas abordagens do GWAS foram conduzidas, associação genômica ampla em passo único ponderada (WssGWAS) e outra com base na matriz genômica de parentesco (GRM) única para cada classe funcional. As herdabilidades aditivas diretas obtidas para PN e PD foram de moderada magnitude, sendo 0,24 e 0,20, respectivamente, e para PS foi de alta magnitude, sendo 0,48. Já as herdabilidades materna obtidas foram de baixa magnitude para PN (0,06) e PD (0,12). Os resultados mostram que existem variantes associadas com as características ponderais distribuídas por diversas regiões do cromossomo 14, reafirmando a importância deste cromossomo na seleção destas características. Os genes *CSMD3*, *PCMTD1*, *PLAG1* e *STK3* estão associados com crescimento e desenvolvimento corporal. Os resultados das classes funcionais demonstram a importância de classes regulatórias e *splice sites* na proporção da variância capturada e explicada pelas variantes nessas classes. As análises das classes funcionais complementam o GWAS, contribuindo para o entendimento da atuação das variantes nas características ponderais da raça Nelore.

Palavras-chave: bovinos de corte, genes, peso corporal, polimorfismo de nucleotídeo único

GENETIC PARAMETERS AND IDENTIFICATION OF REGULATORY REGIONS FOR WEIGHT TRAITS IN NELORE CATTLE

ABSTRACT – Weight traits have been widely used in the selection of Nelore cattle in Brazil, being directly associated with herd productivity and profitability. Genome-wide association studies (GWAS) have been used to investigate variants in the bovine genome that are associated with target phenotypes. However, due to some limitations of GWAS, functional classes annotation for identifying candidate regions can be used to increase GWAS detection power. As few studies were published regarding functional classes associated with weight traits in the Nelore cattle, the objectives of this study were to perform GWAS to identify chromosome 14 regions associated with weight traits in Nelore cattle, and to verify which functional classes explain the greater proportion of variance, seeking to assist the selection process on growth traits. For this, a database of genotypes and phenotypes was analyzed for birth weight (BW), weaning weight (WW) and yearling weight (YW). Two GWAS approaches were conducted, weighted single-step genome-wide association (WssGWAS) and another based on the genomic relationship matrix (GRM) unique to each functional class. The direct additive heritabilities obtained for BW and WW were moderate (0.24 and 0.20, respectively) and for YW was high (0.48). The maternal heritabilities obtained were of low magnitude for BW (0.06) and WW (0.12). The results show that there are variants associated with weight traits distributed across different regions of chromosome 14, reaffirming the importance of this chromosome in the selection of these traits. The *CSMD3*, *PCMTD1*, *PLAG1* and *STK3* genes are associated with growth and development. The results of the functional classes demonstrate the importance of regulatory classes and splice sites in the proportion of the variance captured and explained by the variants in these classes. The analyzes of the functional classes complement the GWAS, contributing to the understanding of the performance of the variants in the weight traits in Nelore cattle.

Keywords: beef cattle, body weight, genes, single nucleotide polymorphism

1. CAPÍTULO 1 – Considerações gerais

1.1 Introdução

A população mundial atingiu aproximadamente 7,6 bilhões de pessoas em 2021 (Dohman et al., 2022). Se o ritmo de crescimento populacional for mantido estima-se que até 2031 alcance 8,5 bilhões de pessoas. Estudos apontam que para atender as futuras demandas de alimentos, a produção agropecuária mundial deverá aumentar cerca de 60% até 2028 (Resilience, 2017). Para atingir a produtividade esperada e atender as demandas futuras por carne bovina, é imprescindível a aplicação de pacotes tecnológicos mais eficientes, produtivos lucrativos e alinhados à sustentabilidade.

Há bastante tempo o melhoramento genético tem sido utilizado, principalmente por meio da seleção, por criadores de bovinos de corte e leite. Recentemente, o desenvolvimento de processos estatísticos para a identificação de animais melhoradores tornou disponível aos produtores ferramentas para aumentar a eficiência e lucratividade de seus rebanhos (Cardoso, 2009). As características ponderais, como peso corporal, estão diretamente ligadas à produção dos rebanhos e tem fundamental importância na pecuária de corte (Nieto et al., 2013).

Buscando incrementar a eficiência dos processos de seleção, ferramentas genômicas têm sido implementadas nos programas de melhoramento genético modernos. Com uso dessas ferramentas, é possível selecionar animais mais jovens, podendo até dobrar as taxas de ganho genético por ano em relação aos métodos tradicionais de seleção (Blasco e Toro, 2014).

Estudos de GWAS possibilitam a realização de análises de associação simultânea de SNPs e fenótipos e a identificação da proporção da variância genética explicada pelos SNPs, permitindo detectar regiões do genoma que estão associadas com características de interesse (Visscher et al., 2012). No entanto, segundo Hill et al. (2008) o poder de detecção de mutações causais é limitado uma vez que a variância gerada pela mutação depende do efeito e da frequência e assim, variantes mais raras ainda que tenham efeitos importantes nas características podem ser de difícil detecção.

Anotar as variantes em classes funcionais de acordo com suas informações biológicas têm sido uma das estratégias para ampliar o poder de detecção do GWAS (Koufariotis et al., 2014), possibilitando maior precisão na identificação de variantes que explicam a variação genética. Dentre as classes funcionais, regiões codificadoras de proteína, pontos de cortes (*Splice sites*) e regiões regulatórias devem ser consideradas como boas candidatas para priorização (Koufariotis et al., 2018). A análise de variantes dessas regiões em características ponderais pode auxiliar na seleção genômica, aumentando os ganhos genéticos e facilitando o melhoramento genético de bovinos da raça Nelore.

1.2 Objetivos

1.2.1 Objetivo geral

Estimar parâmetros genéticos e realizar estudos de associação genômica ampla para identificar regiões do cromossomo 14 associadas com características ponderais de bovinos da raça Nelore, e verificar quais classes funcionais explicam maior proporção da variância.

1.2.2 Objetivos específicos

- i. Estimar parâmetros genéticos para as características de peso ao nascimento, peso a desmama e peso ao sobreano.
- ii. Identificar a localização das classes funcionais, definidas com base em SNPs localizados em regiões codificadoras e regulatórias do cromossomo 14 do genoma bovino.
- iii. Conduzir estudos de associação genômica ampla considerando a abordagem de associação genômica ampla em passo único ponderada (WssGWAS) e outra abordagem com base na matriz genômica de parentesco (GRM) única para cada classe funcional, buscando identificar regiões genômicas que explicam maior proporção da variância nas características ponderais estudadas.

1.3 Revisão de literatura

O Brasil se encontra em crescente desenvolvimento agrícola, sendo um dos principais fornecedores mundiais de produtos de origem animal, e o maior exportador de carne bovina do mundo em 2022 (Dohlman et al., 2022), movimentando cerca de 913,14 bilhões de reais, em toda cadeia produtiva da pecuária de corte, no ano de 2021 (ABIEC, 2022). Além disso, apresenta grande potencial de expansão, em área e produção de carne bovina, sendo o país com um dos maiores rebanhos comerciais do mundo, com população de aproximadamente 196,47 milhões de cabeças, em 2021 (ABIEC, 2022). Entretanto, na maioria dos estados brasileiros, a pecuária de corte se caracteriza de baixo investimento tecnológico, tendo o gado zebuino (*Bos taurus indicus*) como a principal sub-espécie do rebanho comercial.

Os zebuínos importados para o Brasil no final do século 19 e princípio do século 20 constituem a maior parte do rebanho bovino brasileiro. Dentre estes, o Nelore, criado como raça pura ou utilizado em cruzamentos é a principal raça do rebanho comercial brasileiro. Este fato ocorre principalmente devido aos zeus possuírem boa adaptabilidade ao clima tropical e maior resistência aos ecto-parasitas (Sousa et al., 2012).

Ainda que o Brasil esteja presente no topo dos rankings mundiais de produção e exportação de carne bovina, para atender a demandas projetadas para 2031 pelo USDA (Dohlman et al., 2022) faz-se necessário a intensificação de sistemas de produção de carne bovina, sendo vital otimizar o uso da terra e maximizar o desempenho animal de forma consciente e atendendo as exigências do mercado atual, como a priorização do bem-estar animal, controle sanitário rigoroso e combate as mudanças climáticas.

1.3.1 Seleção artificial para características de crescimento

A seleção de animais geneticamente superiores aumenta a frequência de genes desejáveis e conseqüentemente modifica médias e frequências de fenótipos de interesse nos rebanhos (Koury Filho et al., 2010), aumentando a produtividade e lucratividade da produção. Tradicionalmente, a seleção de animais na bovinocultura

de corte é realizada com base nos fenótipos e em informações de parentesco (Meuwissen et al., 2001). Desse modo a seleção é eficaz na melhoria ou fixação de características de importância produtiva e econômica (Euclides Filho, 1999).

A obtenção de dados de peso corporal em diferentes idades tem fundamental importância na pecuária de corte (Nieto et al., 2013) e os pesos padronizados para determinadas idades tem sido amplamente utilizados como critério de seleção em diversos programas de melhoramento genético, principalmente porque essas características apresentam capacidade de resposta à seleção (Gunski et al., 2001; Gonçalves et al., 2011; Barbosa et al., 2017), além de serem de fácil mensuração e de baixo custo para obtenção dos dados.

A herdabilidade do peso corporal é um parâmetro importante para estimar a resposta da seleção e predizer o valor genético dos animais (Gunski et al., 2001). Dentre as características ponderais estudadas, as mais utilizadas em análises e programas de melhoramento genético são: peso ao nascimento, peso a desmama (205 dias), peso ao sobreano (450 ou 550 dias). Essas características são moderadamente herdáveis, apresentando progresso genético através de seleção direta e indireta (Grigoletto et al., 2019).

A utilização do peso ao nascimento como critério de seleção tem como objetivo reduzir ou eliminar partos distócicos (Nieto et al., 2013). Na literatura encontra-se com mais frequência valores de estimativas de herdabilidade direta entre 0,17 a 0,27 (Kamei et al., 2017; Correia et al., 2020). O peso à desmama pode ser considerado como um dos principais critérios de seleção das características de crescimento (Lindholm e Stonaker, 1957; Vesely e Robison, 1971) especialmente porque muitos criadores dedicam-se exclusivamente a fase de cria e assim esse peso tem relação direta com o valor do produto. As estimativas de herdabilidade dos efeitos genéticos diretos do peso à desmama na raça Nelore situam-se com mais frequência entre 0,17 e 0,32 (Pires et al., 2016; Kamei et al., 2017; Lopes et al., 2017). O sobreano é, em geral, a última fase em que é economicamente viável realizar a pesagem de uma grande proporção do rebanho visando a seleção dos animais. Assim, o peso ao sobreano é o mais próximo da reprodução e abate (Nieto et al., 2013) e apresenta estimativas de herdabilidade direta entre 0,21 e 0,61 (Lopes et al., 2017; Kamei et al., 2017).

Os pesos ao nascimento e a desmama são determinados pelo potencial genético do animal e por efeitos maternos. O efeito materno pode ser definido como qualquer influência da mãe no fenótipo do bezerro e reflete-se em diferenças no peso ao nascimento e ganho de peso até a desmama, e podendo ser determinados genética e ambientalmente. Assim, a mãe contribui para o fenótipo do bezerro transmitindo metade dos seus genes e por ambiente materno durante desenvolvimento intrauterino, nascimento, amamentação, defesa imunológica e habilidade materna (Muniz et al., 2005; Perotto 2008; Sarmiento et al., 2011).

Em bovinos de corte, geralmente a habilidade materna é medida a partir do desempenho pré-desmame, sendo a soma do potencial genético do animal para ganhar peso e do potencial genético da matriz para cuidar do bezerro e produzir leite (Silva, 2016). As estimativas de herdabilidade dos efeitos genéticos maternos do peso ao nascimento na raça Nelore situam-se entre 0,08 e 0,09 (Lira et al., 2008; Fernandes, 2022), e o peso a desmama apresenta estimativas de herdabilidade materna entre 0,09 – 0,13 (Laureano et al., 2011; Barros et al., 2018; Evangelista et al., 2020) na raça Nelore. Entretanto, observa-se na literatura antagonismo entre efeitos genéticos direto e materno para características de crescimento em bovinos de corte, apresentando correlações entre -0,28 e -0,37 para PN e entre -0,26 e -0,60 para PD (Grigoletto et al., 2019; Toro et al., 2020; Pardo et al., 2020; Scholtz et al., 2022). Ainda que essas correlações sejam negativas, a adição dos efeitos maternos na seleção resultaria em maior resposta econômica à seleção de longo prazo do que a seleção baseada apenas em efeitos genéticos diretos (Eler et al., 1995).

1.3.2 Marcadores moleculares

Com a evolução da genética molecular, tornou-se possível a obtenção dos genótipos dos animais, isto é, de informações sobre a variabilidade, identificação e localização de genes específicos, via marcadores moleculares. Marcadores moleculares são características herdáveis presentes no DNA, sendo estas, alterações na sequência de nucleotídeos, denominados polimorfismos de DNA, onde cada indivíduo apresenta um padrão único (Dias-Salman et al., 2009). Dentre os diversos

tipos de marcadores moleculares, um dos mais conhecidos e utilizados é o do tipo polimorfismo de nucleotídeo único (SNP).

Os marcadores do tipo SNP evidenciam alterações elementares da molécula de DNA, ou seja, mutações em bases únicas da cadeia de bases nitrogenadas (Adenina, Citosina, Timina e Guanina), podendo ocorrer em regiões codificadoras ou com função regulatória, ou em espaços intergênicos, sem função determinada (Caetano, 2009).

O sequenciamento do genoma teve início em 1975 com metodologia descrita por Sanger e Coulson (1975). No início do século XXI plataformas baseadas no sequenciamento de nova geração (*Next-Generation Sequencing* – NGS) passaram a ser incorporadas no sequenciamento para substituir o método de Sanger, uma vez que, o NGS, era mais eficiente e econômico, sendo utilizado para sequenciamento do genoma de animais domésticos de interesse econômico (Pereira et al., 2013).

Nos últimos anos, chips de genotipagem de alta densidade de marcadores SNP, como a *Illumina BovineHD*, com 777.000 SNPs, foram disponibilizados no mercado, mudando drasticamente as práticas e técnicas de mapeamento do genoma, gerando mapas extremamente informativos e, conseqüentemente, valores genômicos com alta acurácia de predição para algumas características, proporcionando grandes avanços para programas de melhoramento animal (Caetano, 2009; Pereira et al., 2013).

Ainda que seja uma tecnologia cara para ser usada em larga escala na produção animal, os avanços tecnológicos no sequenciamento do DNA reduziram os custos e o tempo necessário para o sequenciamento do genoma e facilitaram a identificação de SNPs em segmentos de DNA, promovendo a identificação e seleção de alelos favoráveis para características analisadas (Cruz et al., 2022), gerando novas possibilidades para o melhoramento genético animal.

1.3.3 QTLs e o cromossomo 14

Características quantitativas são controladas por muitas variantes e genes, resultando em complexas interações alélicas e não alélicas, e são influenciadas pelo ambiente (Regiano, 2004; Collard et al., 2005), compreender como as variantes

contribuem com a variabilidade fenotípica é fundamental para a seleção dos animais superiores. Regiões do genoma associadas a características quantitativas, são chamadas de locus de características quantitativas - QTL (Collard et al., 2005).

O mapeamento de QTLs é um método poderoso para identificar regiões do genoma associados com características de interesse (Korte e Farlow, 2013). Esses mapas permitem a dissecação das características complexas, identificando segmentos do cromossomo que contêm genes determinantes (Coutinho et al., 2010). Entretanto, existe um desafio em detectar variantes causais, devido ao fato de as características quantitativas apresentarem herança poligênica e também devido ao alto desequilíbrio de ligação (*Linkage Disequilibrium* – LD) entre dado QTL e marcadores adjacentes, o que leva a observação de pequenos efeitos para cada marcador e perda de poder estatístico para considerar qualquer destes marcadores como significativo na variação genética total (Hay e Roberts, 2018).

O cromossomo 14 (BTA14) vem adquirindo especial interesse na bovinocultura de corte após diversos estudos relatarem presença de QTLs para várias características inclusive características de crescimento. A maioria dos QTLs foi mapeada na região 15 Mb – 45 Mb do BTA14 e envolvem diversos genes identificados (Wibowo et al., 2008; Hay e Roberts, 2018), inclusive na região do centrômero (Miyata et al., 2007). Nesse cromossomo, estudos detectaram SNPs significativamente associados com características de carcaça, eficiência alimentar, peso corporal, e outras características de crescimento (Veneroni et al., 2010; Oliveira et al., 2014; Bhuiyan et al., 2018; Hay e Roberts, 2018). Outros estudos encontraram o gene *PLAG1* no BTA14, o qual foi confirmado desempenhar um papel importante em mudanças recentes na estatura de bovinos (Boitard et al., 2016). Assim, o cromossomo 14 é considerado de interesse nas associações do genoma e em estudos de genes candidatos (Peters et al., 2012).

1.3.4 A importância das classes funcionais nos estudos de GWAS

Os estudos de associação genômica ampla (GWAS) tem como objetivo detectar variantes em loci genômicos associados com características complexas em uma população (Visscher et al., 2012). Nos estudos de GWAS, há a suposição de que

determinadas mutações causais, ou variantes causais, relacionadas ao fenótipo estejam ligadas, ou seja, em LD com SNPs específicos (Menezes et al., 2013). LD é a associação não aleatória entre os alelos em diferentes loci, podendo ser criado por mutações, derivas, seleção e recombinação (Visscher et al., 2012). Desta forma, é possível realizar uma análise de associação simultânea dos SNPs com fenótipos e identificar a proporção da variância genética explicada pelos SNPs.

Para algumas características, o GWAS pode encontrar uma arquitetura genética simples, com poucos loci com grandes efeitos e, para outras características, pode encontrar uma arquitetura genética mais complexa, podendo ocorrer a presença de variantes raras com grandes efeitos no fenótipo, ou muitas variantes comuns com pequenos efeitos (Korte e Farlow, 2013).

Em alguns estudos, o GWAS tem poder de detecção de mutações causais limitado, além de, frequentemente, serem tendenciosas para mutações comuns, e podem não detectar variantes mais raras com efeitos importantes nas características de interesse (Koufariotis et al., 2018). Uma estratégia para ampliar o poder de detecção do GWAS tem sido filtrar o grande número de variantes em subconjuntos, usando as informações biológicas subjacentes para anotar as variantes em classes funcionais (Koufariotis et al., 2014).

Existem classes podem ser priorizadas nas análises e entre elas destacam-se as relacionadas a estrutura do gene, regiões codificantes, regiões regulatórias, regiões intergenicas (Knight et al., 2011). SNPs localizados em regiões codificantes podem resultar na substituição de aminoácidos que alteram a estrutura e estabilidade de proteínas, potencialmente mudando sua função molecular (Farashi et al., 2019). Conforme Koufariotis et al. (2018), variantes encontradas em *Splice sites* e regiões regulatórias devem ser consideradas como boas candidatas de priorização. Estudos na genética humana demonstram importância que regiões regulatórias tem nas características, sendo altamente complexas com muitas interações que contribuem para a expressão e regulação de genes (Koufariotis et al., 2014).

Diversas classes demonstram efeitos regulatórios na expressão de genes associados com fenótipos, dentre eles, elementos presentes dentro de Íntrons, UTR's (regiões não traduzidas), microRNA (pequenos RNA não codificantes), lncRNA (RNA não codificantes longos), *upstream* e *downstream* apresentam funções regulatórias

na expressão de genes (Koufariotis et al., 2014). Koufariotis et al. (2018), trabalhando com a raça Holandesa encontraram que classes regulatórias, como lncRNA, miRNA e TFBS capturaram boa parte da variância e que *synonymous* e *splice sites* explicaram maior variância por SNP para características de produção e leite e fertilidade. Segundo Liu et al. (2022), o crescimento e desenvolvimento do músculo esquelético em bovinos de corte, característica relacionada ao crescimento corporal, pode ser afetada por uma variedade de fatores regulatórios. Embora a natureza desses fatores regulatórios seja claramente epistática, dado influírem na expressão de outros genes, sugerindo presença de variância não aditiva, Hill et al. (2008) argumentam que frequentemente perto de 100% da variância detectada em função dessas interações epistáticas é de natureza aditiva.

Sendo assim, é mais provável que estas abordagens alternativas do GWAS permitam a identificação de regiões candidatas que podem ser usadas de forma mais acurada na seleção (Costa et al., 2015) para as características desejadas.

1.4 Referências bibliográficas

Associação das Indústrias Exportadoras de Carne – ABIEC (2022) **Beef REPORT - Perfil da Pecuária no Brasil**. ABIEC: Brasília. Disponível em: <<https://www.abiec.com.br/publicacoes/beef-report-2022>>. Acesso 15 de fev. de 2023.

Barbosa ACB, Carneiro PLS, Rezende MPG, Ramos IO, Martins Filho R, Malhado CHM (2017) Parâmetros genéticos para características de crescimento e reprodutivas em bovinos Nelore no Brasil. **Archivos de Zootecnia**, v. 66, n. 255, p. 449-452.

Barros IC, Mota RR, Silva LP, Carneiro PLS, Martins Filho R, Malhado CHM (2018) Avaliação genética do crescimento de bovinos Nelore Mocho, por meio de modelos de multicaracterísticas. **Revista Ceres**, v. 65, p. 402-406.

Bhuiyan MAS, Lim D, Park M, Lee S, Kim Y, Gondro C, Park B, Lee S (2018) Functional partitioning of genomic variance and genome-wide association study for carcass traits in Korean Hanwoo cattle using imputed sequence level SNP data. **Frontiers in Genetics**, v.9, p.217.

Blasco A, Toro MA (2014) A short critical history of the application of genomics to animal breeding. **Livestock Science**, v. 166, p. 4-9.

Boitard S, Boussaha M, Capitan A, Rocha D, Servin B (2016) Uncovering adaptation from sequence data: lessons from genome resequencing of four cattle breeds. **Genetics**, 203(1), 433-450.

Caetano AR (2009) Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, v.38, p.64-71.

Cardoso FF (2009) Ferramentas e estratégias para o melhoramento genético de bovinos de corte. Bagé: EMBRAPA **Pecuária Sul Documentos**, v.83. p.06, 2009.

Collard BC, Jahufer MZZ, Brouwer JB, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. **Euphytica**, 142(1), 169-196.

Correia LEC, Silva Faria RA, Oliveira MHV, Aspilcueta Borquis RR, Curi RA, Vasconcelos Silva JAI (2020) Estimation of (co) variance components for birth weight in Nellore cattle using alternative data editing criteria. **Reproduction in Domestic Animals**. 55(10), 1355-1361.

Coutinho LL, Rosário MF, Jorge EC (2010) Biotecnologia animal. **Estudos Avançados**, v. 24, p. 123-147.

Costa RB et al (2015) Genome-wide association study of reproductive traits in Nelore heifers using Bayesian inference. **Genetics Selection Evolution: GSE**, 47:67, doi: 10.1186/s12711-015-0146-0.

Cruz GS, Correa RMDS, Macedo DB, Emerick LL, Castilho EFD, Rodrigues MDN (2022) Associação entre polimorfismos de genes do eixo somatotrópico e eficiência reprodutiva de fêmeas bovinas: revisão de literatura. **Ciência Animal Brasileira**, v. 23.

Dias-Salman AK, Giachetto PF, Malago Júnior W (2009) Marcadores moleculares na bovinocultura de corte. **Embrapa Informática Agropecuária-Artigo em periódico indexado (ALICE)**.

Dohlman E, Hansen J, Boussios D (2022) **USDA Agricultural Projections to 2031** (No. 323859). United States Department of Agriculture.

Eler JP, Van Vleck LD, Ferraz JB, Lobo RB (1995) Estimation of variances due to direct and maternal effects for growth traits of Nelore cattle. **Journal of Animal Science**, 73(11), 3253-3258.

Euclides Filho K (1999) **Melhoramento genético animal no Brasil: fundamentos, história e importância**. Campo Grande: EMBRAPA/CNPGC. 63p.

Evangelista AF, Cavalcante DH, Malhado CHM, Campelo JEG, Carvalho G, Barros Júnior CP, Sousa Junior SC (2020) Estimação de parâmetros genéticos para características de crescimento em bovinos Nelore Mocho da Região Norte do Brasil. **Livestock Research For Rural Development**, v. 32, n. 10.

Farashi S, Kryza T, Clements J, Batra J (2019) Post-GWAS in prostate cancer: from genetic association to biological contribution. **Nature Reviews Cancer**, v. 19, n. 1, p. 46-59.

Fernandes LS (2022) **Estimativas de parâmetros genéticos e estudo de associação genômica ampla para comportamento de proteção materna em bovinos da raça nelore**. 80 f. Tese (Doutorado em Genética e Melhoramento Animal) – Unesp, Jaboticabal.

Gonçalves FM, Pires AV, Pereira IG, Garcia DA, Farah MM, Meira CT, Cruz VAR (2011) Avaliação genética para peso corporal em um rebanho Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 63, p. 158-164.

Grigoletto L et al (2019) Genome-wide associations and detection of candidate genes for direct and maternal genetic effects influencing growth traits in the Montana Tropical® Composite population. **Livestock Science** 229: 64-76.

Gunski RJ, Garner AV, Borjas AR, Bezerra LAF, Lôbo RB (2001) Estimativas de parâmetros genéticos para características incluídas em critérios de seleção em gado Nelore. **Ciência Rural** 31:603-607.

Hay EH, Roberts A (2018) Genome-wide association study for carcass traits in a composite beef cattle breed. **Livestock Science**. 213, 35–43.

Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. **PLoS genetics**, v. 4, n. 2, p. e1000008.

Kamei LM, Azambuja Ribeiro EL, Fonseca NAN, Muniz CADSD, Camiloti TV, Koritiaki NA, Souza Fortaleza AP (2017) Genetic parameters of growth traits in Nellore cattle. **Semina: Ciências Agrárias** 38(3), 1503-1509.

Knight J, Barnes MR, Breen G, Weale M (2011) Using functional annotation for the empirical determination of Bayes factors for genome-wide association study analysis. **Plos One**, 6(4), e14808.

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. **Plant methods**, v. 9, n. 1, p. 1-9.

Koufariotis L, Chen YPP, Bolormaa S, Hayes BJ (2014) Regulatory and coding genome regions are enriched for trait associated variants in dairy and beef cattle. **BMC Genomics**, 15(1), 1-16.

Koufariotis LT, Chen YPP, Stothard P, Hayes BJ (2018) Variance explained by whole genome sequence variants in coding and regulatory genome annotations for six dairy traits. **BMC Genomics** 19, 237.

Koury Filho W, Albuquerque LG, Forni S, Silva JAV, Yokoo MJ, Alencar MM (2010) Estimativas de parâmetros genéticos para os escores visuais e suas associações com peso corporal em bovinos de corte. **Revista Brasileira de Zootecnia**, v. 39, p. 1015-1022.

Laureano MMM, Boligon AA, Costa RB, Forni S, Severo JLP, Albuquerque LGD (2011) Estimativas de herdabilidade e tendências genéticas para características de crescimento e reprodutivas em bovinos da raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, 63, 143-152.

Lindholm HB, Stonaker HH (1957) Economic importance of traits and selection indexes for beef cattle. **Journal of Animal Science**, v.16, n.3 p.998-1006.

Lira T, Rosa EM, Garnero ADV (2008) Parâmetros genéticos de características produtivas e reprodutivas em zebuínos de corte (revisão). **Ciência Animal Brasileira/Brazilian Animal Science**, 9(1), 1-22.

Liu R, Han M, Liu X, Yu K, Bai X, Dong Y (2022) Genome-wide identification and characterization of long non-coding RNAs in Longissimus dorsi skeletal muscle of Shandong black cattle and Luxi cattle. **Frontiers in Genetics**, 1010.

Lopes FB, Ferreira JL, Lobo RB, Rosa GJM (2017) Bayesian analyses of genetic parameters for growth traits in Nelore cattle raised in pasture. **Genetics and Molecular Research** 16(3).

Menezes GRO, Regitano LCA, Silva MVGB, Cardoso FF, Silva LOC, Siqueira F, Egito AA (2013) Genômica Aplicada ao Melhoramento Genético de Gado de Corte. In: Rosa NA, Martins EM, Menezes GRO, Silva LOC (Eds.). **Melhoramento genético aplicado em gado de corte: Programa Geneplus-Embrapa**. Brasília: Embrapa; Campo Grande: Embrapa Gado de Corte, 17, 214-255.

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, n. 4, p. 1819-1829.

Miyata M, Gasparin G, Coutinho LL, Martinez ML, Machado MA, Silva MVGB, Campos AL, Sonstegard TS, Rosário MF, Regitano LCA (2007) Quantitative trait loci (QTL) mapping for growth traits on bovine chromosome 14. **Genetics and Molecular Biology**, v. 30, p. 364-369.

Muniz CADSD, Carvalheiro R, Fries LA, Queiroz SAD (2005) Dois critérios de seleção na pré-desmama em bovinos da raça Gir. 1. Estimativas de parâmetros genéticos. **Revista Brasileira de Zootecnia**, 34, 807-815.

Nieto LM, Alencar MM, Rosa NA (2013) Critérios de seleção. In: Rosa NA, Martins EM, Menezes GRO, Silva LOC (Eds.). **Melhoramento genético aplicado em gado de corte: Programa Geneplus-Embrapa**. Brasília: Embrapa; Campo Grande: Embrapa Gado de Corte, 10, 109-122.

Oliveira PSN et al (2014) Identification of genomic regions associated with feed efficiency in Nelore cattle. **BMC Genetics**, v. 15, n. 1, p. 1-10.

Pardo AM, Elzo MA, Gama LT, Melucci LM (2020) Genetic parameters for growth and cow productivity traits in Angus, Hereford and crossbred cattle. **Livestock Science**, 233, 103952.

Pereira GL, Rosa KO, Curi RA, Regitano LCA, Mota MDS (2013) State of the art of genomic sequencing in livestock. **Ars Veterinaria**, v. 29, n. 3, p. 190-199.

Perotto D (2008) Habilidade materna em bovinos de corte. In: **Anais do 7º Congresso Brasileiro das Raças Zebuínas**; Uberaba/MG.

Peters SO, Kizilkaya K, Garrick DJ, Fernando RL, Reecy JM, Weaber RL, Silver GA, Thomar MG (2012) 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.

Pires BC et al (2016) Genetic analyses on bodyweight, reproductive, and carcass traits in composite beef cattle. **Animal Production Science**, 57(3), 415-421.

Regitano LCA (2004) Importância da genética molecular para o melhoramento de ruminantes. In: **Anais** do V Simpósio da Sociedade Brasileira de Melhoramento Animal; Pirassununga/SP.

Resilience B. (2017) The state of food security and nutrition in the world. **Rome: Building resilience for peace and food security.**

Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed Synthesis with DNA polymerase. **Journal of Molecular Biology**, v.94, n.3, p.441-448.

Sarmiento JLR, Torres RA, Sousa WH, Lôbo RNB, Sousa JER, Cavalcante Neto A, Do Ó AO (2011) Efeito materno sobre a curva de crescimento de ovinos Santa Inês por meio de modelos de regressão aleatória. **Communicata Scientiae**, 2, 133-121.

Scholtz MM, MacNeil MD, Chadyiwa MC, Naser FWC (2022) Genetic analysis of traits needed to optimize improvement in beef cow-calf efficiency in Afrikaner cattle. In **Proceeding of 12th World Congress on Genetics Applied to Livestock Production (WCGALP) Technical and species orientated innovations in animal breeding, and contribution of genetics to solving societal challenges** (pp. 1713-1716). Wageningen Academic Publishers.

Silva RP (2016) **Estudo genético quantitativo do efeito materno em bovinos da raça Nelore**. 47 f – Dissertação (Mestrado em Veterinária) – UFU, Uberlândia.

Sousa GGT, Junior, SCS, Santos KR, Guimarães JEC, Luz CSM, Junior CPB, Fonseca WJL (2012) Características reprodutivas de bovinos da raça Nelore do meio Norte do Brasil. **PUBVET**, Londrina, V. 6, N. 21, Ed. 208, Art. 1390.

Toro EJ, Burgos Paz WO, Elzo MA, Martínez Sarmiento RA, Cerón-Muñoz MF (2020) Genetic parameters and trends for growth traits in Blanco Orejinegro cattle. **Translational Animal Science**, 4(3), txaa174.

Veneroni GB, Meirelles SL, Grossi DA, Gasparin G, Ibelli AMG, Tizioto PC, Oliveira HN, Alencar MM, REGITANO LDA (2010) Prospecting candidate SNPs for backfat in Canchim beef cattle. **Genetics and Molecular Research** 9:1997-2003.

Vesely JA, Robison OW (1971) Conventional selection indexes for birth and weaning traits in beef calves. **Journal of Animal Science**, v.33, n.3, p.537-541.

Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. **The American Journal of Human Genetics**, v. 90, n. 1, p. 7-24.

Wibowo TA, Gaskins CT, Newberry RC, Thorgaard GH, Michal JJ, Jiang Z (2008) Genome assembly anchored QTL map of bovine chromosome 14. **International Journal of Biological Sciences**, 4(6), 406.

2. CAPÍTULO 2 – Parâmetros genéticos e identificação de regiões regulatórias para características ponderais em bovinos Nelore

RESUMO – A raça Nelore tem fundamental importância econômica na pecuária de corte brasileira. Características ponderais geralmente são utilizadas como critérios de seleção nos programas de melhoramento de Nelores. Estudos de associação genômica ampla (GWAS) com o uso de classes funcionais auxiliam na identificação de regiões do genoma bovino associadas com características ponderais. O objetivo deste estudo foi realizar duas abordagens de GWAS, sendo uma de GWAS em passo único ponderada (WssGWAS) com base nos animais genotipados com chips de alta densidade e outra abordagem com base na matriz genômica de parentesco (GRM) única para cada classe funcional com base nos animais sequenciados, buscando identificar regiões genômicas no cromossomo 14 que explicam maior proporção da variância nas características ponderais estudadas de bovinos da raça Nelore. Para isso, um banco de dados de fenótipos, contendo 1.106.848 animais, e dois bancos de dados de genótipo, um formado com chips de alta densidade, e outro com formado por informações de animais sequenciados, foram criados. As características consideradas foram peso ao nascimento (PN), peso a desmama (PD) e peso ao sobreano (PS). As herdabilidades aditivas diretas obtidas para PN e PD foram de moderada magnitude, sendo 0,24 e 0,20, respectivamente, e para PS foi de alta magnitude, sendo 0,48. Já as herdabilidades materna obtidas foram de baixa magnitude para PN e PD, sendo 0,06 e 0,12, respectivamente. Os resultados encontrados pelo WssGWAS demonstram que existem variantes associadas com as características ponderais distribuídas por diversas regiões do cromossomo 14. Janelas adjacentes de SNPs encontraram um total de 89 genes nas 3 características. Os genes *CSMD3*, *PCMTD1*, *PLAG1* e *STK3*, estão associados com crescimento e desenvolvimento corporal. As classes funcionais Non, 5UTR e Splice explicaram maior proporção da variância para PN, PD e PS, respectivamente. Os resultados das classes funcionais demonstram a importância de classes regulatórias e *Splice sites* na proporção da variância capturada e explicada pelas variantes nessas classes. Essas variantes devem ser consideradas em trabalhos futuros para entender as associações dessa classe com características ponderais na raça Nelore. Também são necessários trabalhos futuros para verificar se o uso das classes funcionais pode contribuir para aumentar a acurácia dos GEBV, além da sugestão de validar os resultados obtidos para o genoma bovino completo.

Palavras-chave: *Bos indicus*, classes funcionais, GWAS, variantes

2.1. Introdução

A raça Nelore utilizada como raça pura ou em cruzamentos tem fundamental importância econômica na pecuária de corte brasileira e vem sendo selecionada em programas de melhoramento genético há mais de 50 anos (Euclides Filho, 2009). Um dos principais objetivos de seleção nessa raça é o aumento da velocidade de crescimento desses animais e os pesos ajustados em idades padrão são geralmente utilizados como critérios de seleção nos programas de melhoramento, devido estarem correlacionados com características reprodutivas e de abate. A obtenção das estimativas dos parâmetros genéticos para características ponderais tem impacto na definição de estratégias para seleção animal direta e indireta, sendo ainda a base para as avaliações genéticas (Evangelista et al., 2020).

Estudos de associação genômica ampla (GWAS) tem sido empregado na identificação de variantes genéticas associadas com as características de interesse (Marín-Garzón et al., 2021), possibilitando realizar análise de associação simultânea dos SNPs com fenótipos e identificar a proporção da variância genética explicada pelos SNPs. Entretanto, o GWAS pode ter limitações na detecção de mutações causais, não detectando variantes que podem ter efeitos importantes nas características de interesse (Koufariotis et al., 2018). Uma alternativa é combinar o GWAS com ferramentas de anotação funcional, usando as informações biológicas das variantes para anotá-las em classes funcionais com o objetivo de auxiliar na detecção do papel biológico das variantes e dos genes candidatos e possibilitar maior precisão na identificação de variantes que explicam a variação genética (Koufariotis et al., 2018; Marín-Garzón et al., 2021). Além de utilizar essas variantes de classes funcionais para aumentar as acurácias em predições genômicas, melhorando o progresso genético e permitindo aumentar a eficiência de painéis de SNPs quando prioriza essas variantes identificadas. Dentre as classes funcionais, *splice sites* e regiões regulatórias devem ser consideradas para análises de classes funcionais (Koufariotis et al., 2018).

Estudos de variância explicada por variantes identificadas por classes funcionais com o uso de GWAS podem ser encontradas na literatura para *Bos taurus*

taurus (Koufariotis et al., 2018), que destacaram que as classes lncRNA, miRNA, TFBS, *synonymous* e *splice sites* capturam boa parte da variância, sendo a *splice sites* a classe com maior variância explicada por SNP para características de produção de leite e fertilidade, e *Bos taurus coreanae* (Bhuiyan et al., 2018), onde observaram que SNPs intragenicos contribuem mais para a variância, sendo a classe *synonymous* a que explica maior variância por SNP para características de carne e carcaça. Porém, existem poucos estudos na literatura para identificação dessas variantes para características ponderais em *Bos indicus*, especificamente, a raça Nelore.

Dessa forma, o objetivo do estudo foi: i) estimar parâmetros genéticos para as características peso ao nascimento, peso a desmama e peso ao sobreano em bovinos da raças Nelore, e, ii) conduzir dois estudos de GWAS, com a abordagem de associação genômica ampla em passo único ponderada (WssGWAS) para informações de genótipos e com uma abordagem, para dados de sequência, que considera a matriz genômica de parentesco (GRM) única construída para cada classe funcional priorizada e a matriz GRM conjunta que engloba várias classes funcionais, no cromossomo 14. Com isso, espera-se que os resultados do emprego de diferentes abordagens de GWAS para características ponderais na raça Nelore, forneça bases científicas sobre a variância explicada por variantes de regiões codificantes e regulatórias, localizadas no cromossomo 14 do genoma bovino, podendo auxiliar os programas de melhoramento genético na seleção destas características.

2.2 Metodologia

2.2.1 Animais e fenótipos

O banco de dados dos fenótipos foi formado por um total de 1.106.848 animais, registrados entre 1995 e 2019, provenientes do banco de dados do Programa de Melhoramento Genético Aliança Nelore. Os animais foram criados em pastagens tropicais, em 544 fazendas localizadas no Brasil e no Paraguai.

As características consideradas no presente trabalho foram peso ao nascimento (PN), peso a desmama (PD) medido por volta dos 7 meses de idade (205 dias), e peso ao sobreano (PS) medido por volta dos 18 meses de idade (550 dias).

A estruturação do banco de dados foi realizada, inicialmente, no programa R, versão 4.2.1. O banco de dados de fenótipos foi dividido em três arquivos de dados, sendo um para cada das características analisadas. Os grupos contemporâneos (GC) foram estabelecidos por fazenda, ano de nascimento, estação de nascimento, sexo, data da desmama e grupo de manejo da desmama e sobreano. As estações de nascimento foram definidas em: 1 = janeiro a março, 2 = abril a junho, 3 = julho a setembro, e 4 = outubro a dezembro. Com a finalidade de eliminar informações discrepantes (*outliers*), os pesos a desmama e sobreano corrigidos para os efeitos de idade na medida (linear e quadrático) após análise por meio de modelo linear. Registros de pesos ao nascimento, e de pesos na desmama e sobreano corrigidos para idade de medida, que apresentavam desvio superior a 3,5 desvios padrão da média do GC foram removidos, assim como grupos contemporâneos com menos de 10 animais. Também foram removidos os registros dos animais sem pais conhecidos. Após aplicado o controle de qualidade dos dados, obteve-se 336.004 animais com peso ao nascimento, 659.542 animais com peso a desmama e, 652.864 animais com peso ao sobreano (Tabela 1).

O arquivo de *pedigree* inicial foi formado por 1.596.325 animais, de oito gerações. Após a eliminação dos dados de *pedigree* dos animais sem dados fenotípicos e que não fossem ancestrais de animais com fenótipos, foi gerado um novo arquivo de *pedigree* com dados de 741.272 animais, dos quais 9.669 eram touros e 433.862 matrizes.

Tabela 1. Quantidade de animais, touros e matrizes no *pedigree*, número de grupos contemporâneos (GC), médias, desvio padrão (DP), valores mínimos (MIN), máximos (MAX) e coeficiente de variação (CV) para as características de peso ao nascimento (PN), desmama (PD) e sobreano (PS).

	Animais	Touros	Matrizes	GC	Media ¹	DP ¹	MIN ¹	MAX ¹	CV (%)
PN	336.004	5.772	197.824	2.835	31,97	4,11	18	48	12,88
PD	659.542	8.671	390.395	19.829	193,7	31,75	67	300	16,39
PS	652.864	8.590	388.774	21.499	298,3	49,38	200	550	16,55

¹quilogramas (kg).

2.2.2 Dados de genótipos e controle de qualidade

Dois bancos de dados diferentes foram considerados a depender da abordagem: um arquivo de genótipos e um arquivo de informações de sequências. Uma filtragem foi realizada para escolher os mesmos animais com informação de fenótipos e genótipos nos dois conjuntos de dados, sendo 7.204, 15.541 e 14.732 animais para PN, PD e PS, respectivamente. Informações sobre painel de genotipagem bem como sobre a plataforma de sequenciamento e imputação são detalhadas a seguir.

O arquivo de genótipos continha aproximadamente 62.000 animais com genótipos imputados para painel de alta densidade. Os painéis de genotipagem usados foram de alta densidade da Illumina 777k BovineHD e de média densidade (35k SNPs) conforme detalhado em Fernandes Junior et al. (2021). O controle de qualidade dos genótipos foi realizado considerando SNPs localizados em cromossomos autossômicos com *GenCall score* maior que 0,80, excluindo SNPs localizados na mesma posição genômica, com frequência do alelo menor (MAF) menor que 0,05, *call rate* menor que 0,90 e teste de qui-quadrado para equilíbrio de Hardy-Weinberg (HWE) com $p_value < 10^{-5}$.

Para o arquivo do sequenciamento, o total de 151 touros com contribuição para a diversidade genética da população recente da raça Nelore no Brasil foram selecionados para o sequenciamento, conforme metodologia descrita por Maiorano et al. (2022). Ao todo, 52 touros foram sequenciados com o *Illumina HiSeq X™* e os outros 99 touros foram sequenciados com o *Illumina NovaSeq™* (Fernandes Junior et al., 2021). Os critérios usados para definir os 151 touros para representar a raça foram baseados no grande número de progênie presentes em uma amostra de 10.000 Nelores comerciais genotipados. Os critérios de seleção desses 151 touros, o controle de qualidade dos SNPs e posterior imputação de genótipos para sequência estão detalhados em Fernandes Júnior et al. (2021). Os 151 touros foram usados como população referência para imputação de genótipos, resultando em arquivo contendo informações de sequências imputadas de aproximadamente 62.000 animais (Fernandes Júnior et al., 2021). Basicamente, os parâmetros de controle de qualidade

aplicado após a imputação foram: quality by depth (QD) <2,0; Fisher Strand test (FS) >60,0; mapping quality score (MQ) <40,0; ReadPos-RankSum <8,0; MQRank-Sum <12,5 e SOR >3,0 (Fernandes Júnior et al., 2021).

Os arquivos de sequência foram filtrados e alinhados com o genoma bovino de referência mais recente, o ARS-UCD1.2, resultando em 991.358 SNPs bialélicos localizados no cromossomo 14. Os seguintes controles de qualidade foram aplicados no programa plink v1.9 (Chang et al., 2015) para verificar a qualidade dos bancos de dados de sequências imputadas: *call rate* ≤ 0,90 para as amostras, e *call rate* ≤ 0,90 e $P < 10^{-7}$ no teste de HWE, para os SNPs e $MAF > 0,01$.

O banco de dados dos genótipos foi utilizado para a estimação das variâncias fenotípicas e genéticas, análises de WssGWAS e para anotação dos genes e mapeamento dos QTLs. Já o banco de dados do sequenciamento foi utilizado para a anotação das classes funcionais e estimação da variância explicada pelas classes funcionais pelo método GBLUP, com abordagem de máxima verossimilhança restrita (REML).

2.2.3 Estimativas de parâmetros genéticos

Para a obtenção dos componentes de variâncias, as análises foram conduzidas com os programas da família BLUPF90 (Misztal, 2002), incluindo registros de fenótipo e *pedigree*. Primeiro, foi usado o programa RENUMF90 para renumerar os animais, e criar os arquivos de fenótipos e *pedigree* no formato de entrada para os programas posteriores. Em seguida o programa GIBBS2F90 foi usado para aplicação de inferência bayesiana com o modelo descrito a seguir:

$$y = X\beta + Z\alpha + Mmat + Pmpe + \varepsilon$$

em que, y é vetor dos fenótipos, X é a matriz de incidência relacionando fenótipos a efeitos fixos, β é o vetor de efeitos fixos (GC e idade da mãe) e covariáveis (idade do animal na pesagem, linear e quadrática), Z é a matriz de incidência relacionando animais aos fenótipos, α é o vetor de efeitos genéticos aditivos diretos de cada animal,

M é a matriz de incidência relacionando fenótipos aos efeitos aditivos maternos, mat é o vetor de efeitos aditivos maternos, P é a matriz de incidência relacionando fenótipo ao ambiente materno permanente, mpe é o vetor de efeitos permanentes maternos, e ε é o vetor de resíduos. Foi suposto que $\alpha \sim N(0, G)$, $mat \sim N(0, M)$, $mpe \sim N(0, P)$, $\varepsilon \sim N(0, R)$, em que $G = A\sigma_{\alpha}^2$, $M = A\sigma_{mat}^2$, $P = I\sigma_{mpe}^2$, $R = I\sigma_{\varepsilon}^2$, e que:

$$y|\beta, \alpha, mat, mpe, G, P, R \sim N(X\beta + Z\alpha + Mmat + Pmpe, R)$$

em que $G = A\sigma_{\alpha}^2$, $M = A\sigma_{mat}^2$, $P = I\sigma_{mpe}^2$, $R = I\sigma_{\varepsilon}^2$, onde σ_{α}^2 , σ_{mat}^2 , σ_{mpe}^2 e σ_{ε}^2 correspondem as variâncias genética aditiva direta, materna, de ambiente permanente materno e residual, respectivamente, A é a matriz dos numeradores dos coeficientes de parentesco, e I é uma matriz identidade de ordem adequada. Para as variâncias foram utilizadas como prioris as distribuições qui-quadrados escalonadas invertidas com hiperparâmetros escolhidos para serem não informativas. Também para β foi utilizada uma priori não informativa. Possíveis correlações entre os efeitos genéticos diretos e maternos não foram consideradas.

Foram geradas 300.000 amostras das distribuições posteriores pelo algoritmo do amostrador de Gibbs, sendo descartadas as 4.000 iniciais e coletada uma a cada 10 iterações, totalizando 29.600 amostras armazenadas para os componentes de variância de cada característica. As convergências das cadeias foram verificadas com uso do pacote coda do R (Plummer *et al.*, 2006), com gráficos de traços e com o teste de Geweke, também foram obtidas as médias das herdabilidades, erros de Monte Carlo e intervalos de maior densidade dos componentes de (co)variância.

2.2.4 Estudo de associação genômica ampla conforme abordagem WssGWAS

O GWAS foi conduzido em análises univariadas utilizando a metodologia do WssGBLUP, proposta por Wang *et al.* (2012), com uso dos mesmos modelos utilizados para a obtenção das herdabilidades, considerando as mesmas pressuposições, exceto que $\alpha \sim N(0, H\sigma_{\alpha}^2)$ e $mat \sim N(0, H\sigma_{mat}^2)$. A matriz H combina as informações genômicas com as informações de *pedigree*, e sua inversa (Aguilar *et al.*, 2010), que possui estrutura simples para os cálculos, é descrita por:

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

em que A_{22} é a matriz de coeficientes de parentesco entre indivíduos genotipados, e G é a matriz de relacionamento genômico obtida segundo metodologia descrita por VanRaden (2008), descrita por:

$$G = \frac{ZZ'}{2 \sum_{i=1}^m p_i(1 - p_i)}$$

em que, Z é a matriz de conteúdo genético, ajustada para frequências alélicas (0, 1 ou 2 para AA, AB e BB, respectivamente), m é o número de SNPs e p_i é a menor frequência alélica do i -ésimo SNP.

Os programas da família BLUPF90 foram utilizados para obtenção dos valores genéticos preditos para os animais e, posteriormente, os efeitos dos SNPs. O programa preGSF90 foi utilizado para calcular a matriz H^{-1} . O programa BLUF90 foi utilizado para obtenção das soluções dos valores genéticos dos animais. O programa postGSF90 foi usado na obtenção das soluções dos efeitos dos SNPs a partir dos valores genéticos predito dos animais genotipados (GEBV), conforme proposto por Wang et al. (2012):

$$\hat{u}_{(t)} = \lambda D_{(t)} Z' G_{(t)}^{*-1} \hat{a}_g$$

em que: $\hat{u}$ é o vetor dos efeitos dos SNPs, λ é taxa de variância calculada de acordo com VanRaden et al. (2009), $\hat{a}_g$ é o GEBV, Z é a matriz que relaciona os genótipos de cada locus, G é a matriz de relação genômica e D é a matriz diagonal de ponderação das variâncias dos SNPs (inicialmente $D = 1$), obtido pelo algoritmo a seguir, onde t é o número da iteração e i refere-se ao i -ésimo SNP:

1. $t = 0$; $D_{(t)} = I$; $G_{(t)}^* = Z D_{(t)} Z' \lambda$.
2. Estimar $\hat{a}_g$ com o ssGBLUP.
3. Calcular $\hat{u}_{(t)} = \lambda D_{(t)} Z' G_{(t)}^{*-1} \hat{a}_g$.
4. Calcular $d_{i(t+1)}^* = \hat{u}_{i(t)}^2 2p_i(1 - p_i)$ para todos i como descrito por Zhang et al. (2010).
5. Normalizar $D_{(t+1)} = \frac{tr(D_{(0)})}{tr(D_{(t+1)}^*)} D_{(t+1)}^*$.

6. Calcular $G_{(t+1)}^* = ZD_{(t+1)}Z'\lambda$.

7. Voltar para o passo 2.

Na primeira iteração, os SNPs foram considerados igualmente importantes para a característica. Ao final da primeira iteração, juntamente com as soluções dos SNPs, foram geradas as variâncias dessas soluções, as quais foram utilizadas como ponderadores, para diferenciar os SNPs de acordo com a variância que esses explicam. O mesmo procedimento foi executado cinco vezes. Para a estimação das variâncias dos SNPs (utilizadas como ponderadores) foi utilizado o método Não-linear A, descrito por VanRaden (2008):

$$\sigma_{u,i}^2 = \frac{\sigma_a^2}{\sum_{j=1}^m 2p_jq_j} 1.25^{\frac{|\hat{u}_i|}{sd(\hat{u})}-2}$$

em que, $\sigma_{u,i}^2$ pode ser compreendido como variância prévia do SNP (Fragomeni et al., 2019), $|\hat{u}_i|$ é a estimação absoluta do efeito do SNP para o marcador i , e $sd(\hat{u})$ é o desvio padrão do vetor da estimação do efeito do SNP, o valor 1.25 é uma constante que determina o quanto a distribuição dos efeitos de SNP se afasta de uma distribuição normal.

Para selecionar qual iteração seria utilizada para análises seguintes, foi considerada a acurácia de cada iteração, calculada no programa BLUPF90, pela seguinte fórmula:

$$acurácia = \sqrt{1 - \frac{PEV}{(\sigma_u^2(1 + F))}}$$

em que: PEV representa o erro de predição da variância do GEBV, F é coeficiente de endogamia e σ_u^2 é a variância genética.

De acordo com a acurácia, foram escolhidas a quarta iteração para peso ao nascimento e peso à desmama, enquanto para o peso ao sobreano foi escolhida a segunda iteração. As variâncias de janelas deslizantes contendo dez SNPs adjacentes foram usadas como parâmetro nas análises de WssGWAS. Janelas sem sobreposição foram utilizadas para construir os *Manhattan plots*, as 20 janelas de maior variação, sem sobreposição de SNPs foram consideradas, com intuito de

identificar regiões do genoma que explicaram maior proporção da variância genética aditiva das características analisadas.

2.2.5 Anotação de genes e mapeamento de QTLs

Para complementar os resultados do GWAS, foi realizada análise funcional nas 20 janelas de maior variação para cada característica. Para cada janela, foram analisadas regiões de 500kb antes e depois da posição do primeiro e último SNP da janela. Para localizar genes candidatos em sobreposição com as janelas, o pacote do R Biomart (Durinck et al., 2005) foi usado em conjunto com o banco de dados Ensembl Genes 109, com o genoma bovino ARS–UCD1.2 como referência. A sobreposição das janelas de maior variância com locus de característica quantitativa (QTL), foi verificada considerando o banco de dados do CattleQTLdb (Hu et al., 2022).

O pacote GALLO (Fonseca et al., 2020) foi utilizado para obter as porcentagens dos QTLs mapeados, com o uso do banco de dados CattleQTLdb. Os pacotes “clusterProfiler” (Yu et al., 2012) e “org.Bt.eg.db” (Carlson, 2020) do programa R foram utilizados para obter o banco de dados biológicos *Gene Ontology Consortium* (GO). E assim, realizar uma análise de agrupamento de genes candidatos considerando a similaridade em termos associados com Componente Celular (CC), Função Molecular (FM) e Processo Biológico (PB), além de verificar a relação dos clusters com as características estudadas e determinar se há enriquecimento funcional ($p < 0,05$), ou seja, se os termos estão relacionados com os genes candidatos mais do que esperado ao acaso.

O pacote STRINGdb versão 11.5 do programa R (Szklarczyk, et al., 2021) foi usado para realizar análise funcional agrupando genes associados a termos GO, considerando como parâmetros o *score threshold* 200 e a espécie 9913 – *Bos taurus*. Genes encontrados que não tinham interações com outros genes da rede foram descartados desta análise.

2.2.6 Classes funcionais

A busca por informação sobre a classe funcional de cada variante nos dados de sequências foi realizada considerando informações disponíveis online, constantes nos bancos de dados Ensembl (Hubbard et al., 2009), RumimiR (Bourdon et al., 2019) e NONCODEV6 (Zhao et al., 2021), além de informações geradas e disponibilizadas em outros estudos (Zhou et al., 2020; Klees et al., 2021).

Com o uso do banco de dados do Ensembl, as seguintes classes foram consideradas para anotação das variantes das sequências: *Exon variants* (Exon), *Intron* (Intron), *Protein Coding Sequence* (CDS), *5 kb upstream* (Up), *5 kb Downstream* (Down), *frame-shift* (frame), *Splice sites* (Splice), *stop codon* (stop), *Missensesense* (Missense), *Non-coding transcript* (Non), *Synonymous* (Syn), *3' Untranslated Regions* (3UTR), e *5' 5' Untranslated Regions* (5UTR). Na classe Splice, foram consideradas variantes anotadas como *Splice acceptor variant*, *Splice donor variant*, encontradas, respectivamente, perto do fim do 3' e 5' dos introns, além de todas as outras variantes anotadas como Splice, seguindo a junção adotada por Koufariotis et al. (2018). A busca por essas classes foi realizada usando a versão do genoma bovino ARS-UCD1.2 juntamente com o pacote biomaRt do R (Durinck et al., 2005), considerando apenas o cromossomo 14.

Também foram usadas outras fontes de informações para as classes microRNA, *Transcription Factor Binding Sites* (TFBS), ilhas de CpG (CpG) e *long noncoding RNA* (lncRNA). Para anotação de microRNA, foi utilizado o banco de dados RumimiR (Bourdon et al., 2019), disponível em <http://rumimir.sigenae.org> (acessado em 15 de fevereiro de 2023). As informações de TFBS foram obtidas do banco de dados proveniente do estudo de Klees et al. (2021), disponível em <https://azifi.tz.agrar.uni-goettingen.de/agreg-snpdb> (acessado em 15 de fevereiro de 2023). Para anotação das ilhas de CpG foi utilizado dados gerados pelo artigo do Zhou et al. (2020). Por fim, a classe lncRNA foi obtida do banco de dados NONCODEV6 (Zhao et al., 2021), disponível em <http://www.noncode.org> (acessado em 15 de fevereiro de 2023), porém essas informações estavam baseadas na versão do genoma bovino anterior, sendo necessário o uso da ferramenta *Lift Genome*

Annotations (disponível em <https://genome.ucsc.edu/cgi-bin/hgLiftOver>, acessado em 15 de fevereiro de 2023) para conversão dos dados para a versão do genoma bovino mais recente. Como essa ferramenta não fazia a conversão direta do UMD3.1/bosTau6 para ARS-UCD1.2, primeiro foi realizada uma conversão dos dados de UMD3.1/bosTau6 para UMD3.1/bosTau8, e então, de UMD3.1/bosTau8 para ARS-UCD1.2. Classes com números pequenos de variantes foram excluídas das análises (Tabela 2).

Tabela 2. Classes funcionais anotadas e o número de variantes de cada classe no cromossomo 14

Classe	Abreviatura	Número total de variantes	Proporção (%)
<i>3 prime UTR variant</i>	3UTR	2305	0,59
<i>5 prime UTR variant</i>	5UTR	799	0,20
<i>Protein coding sequence</i>	CDS	4614	1,17
Ilhas de CpG	CpG	3807	0,97
<i>Downstream gene variant</i>	Down	29530	7,50
<i>Exon variant</i>	Exon	8891	2,26
<i>Frameshift variant</i> ¹	Frame ¹	6	0,00
<i>Intron variant</i>	Intron	281960	71,65
<i>Long noncoding RNA</i>	lncRNA	8263	2,10
microRNA ¹	microRNA ¹	34	0,01
<i>Missense variant</i>	Missense	1561	0,40
<i>Non-coding transcript variants</i>	Non	8752	2,22
<i>Splice site region variants</i>	Splice	390	0,10
<i>Stop codon variants</i> ²	Stop	36	0,01
<i>Synonymous variant</i>	Syn	2550	0,65
TFBS	TFBS	8295	2,11
<i>upstream gene variant</i>	Up	31849	8,09

¹classes removidas devida baixa quantidade de variantes.

Ao final, 14 classes funcionais foram usadas para as análises envolvendo os dados de sequências, além de uma análise conjunta das classes. Para compreender a influência da MAF nos resultados, uma análise separada foi realizada para seis categorias de MAF, sendo elas: maf1 (menor que 0,05), maf2 (entre 0,05 e 0,1), maf3 (entre 0,1 e 0,2), maf4 (entre 0,2 e 0,3), maf5 (entre 0,3 e 0,4), e maf6 (entre 0,4 e 0,5).

2.2.7 Estimação da variância explicada pelas classes funcionais pelo método GBLUP com abordagem REML

A variância explicada pela classe funcional foi determinada ajustando as GRM de todas as classes funcionais simultaneamente no modelo, com o uso do programa GCTA (Yang et al., 2011). Este método de ajustar classes conjuntamente foi descrito como confiável na obtenção das variâncias explicadas pelas classes (Koufariotis et al., 2018, Bhuiyan et al., 2018).

O plink foi usado para converter os arquivos “.vcf” em arquivos binários, os quais são solicitados no programa GCTA. Inicialmente uma matriz GRM foi calculada para os SNPs de cada classe funcional usando a opção “--make-grm”.

O programa PREDICTF90 da família BLUPF90 foi usado para ajustar os fenótipos de acordo com os efeitos fixos idade da mãe, idade do animal e idade animal quadrática antes de realizar as análises no GCTA. Após isso, um modelo misto foi usado para calcular a variância explicada para cada classe funcional de SNP, usando o método de máxima verossimilhança restrita (REML), descrito por Koufariotis et al. (2018):

$$y = X\beta + Zg + \varepsilon$$

em que, y é o vetor dos fenótipos ajustados, β é o vetor do efeito fixo (GC), X é a matriz de incidência dos efeitos fixos, g é o vetor dos valores genéticos, Z é a matriz de incidência dos valores genéticos e ε é o resíduo. Os valores genéticos foram obtidos com a seguinte pressuposição:

$$g \sim N(0, G\sigma_g^2)$$

em que g é a GRM da classe funcional e σ_g^2 é a variância genética de cada classe funcional.

As proporções da variância capturada e as variâncias explicadas para cada classe foram calculadas de acordo com as equações descritas por Koufariotis et al. (2018), da seguinte forma:

$$\text{proporção total da variancia} = \left(\frac{h^2}{\sum_{i+1}(h^2)} \right) \times 100$$

onde h^2 representa a variância capturada por cada classe funcional, e:

$$\text{VarPorSNP} = \frac{((h^2 \div n) \times 100)}{10^{-4}}$$

em que, neste caso, h^2 representa a herdabilidade da classe e é dividida por n , o número total de variantes encontradas na classe.

Para determinar as similaridades entre as GRM de cada classe funcional, foi calculada a distância Euclidiana entre elas, conforme metodologia adotada por Koufariotis et al. (2018), com a fórmula:

$$\text{Distância Euclidiana} = \sqrt{\sum_{i=1} (m_i - p_i)^2}$$

onde, m e p é a GRM correspondente a cada classe.

2.3. Resultados e discussão

2.3.1 Parâmetros genéticos

As análises para obtenção dos componentes de variância mostram que entre os efeitos considerados no modelo foi identificada maior contribuição da variação genética aditiva direta para as características em relação as outras fontes de variação consideradas no modelo animal (Tabela 3).

A estimativa de herdabilidade direta para PN e PD, foram de moderada magnitude, possibilitando progresso genético por meio da seleção direta, porém os efeitos não identificados no modelo, tem maior influência na característica dificultando evidenciar os valores genéticos dos animais. As estimativas estão próximas do relatado por Kamei et al. (2017) e Pedrosa et al. (2014), variando de 0,17 a 0,31 para PN e 0,17 a 0,22 para PD na raça. Os valores encontrados para herdabilidade dos efeitos maternos para PN e PD estão próximos com valores descritos na literatura, onde Fernandes (2022) obteve herdabilidade materna de 0,09 para PN, e Evangelista et al. (2020) obtiveram herdabilidade materna de 0,09 para PD, para bovinos Nelores. Apesar da baixa magnitude, esse efeito não deve ser desprezado no modelo utilizado para análises dessas características de peso.

Tabela 3. Médias e intervalo de 95% de maior densidade (HDI) das variâncias genética aditiva direta (σ_{α}^2) e materna (σ_{mat}^2), ambiente permanente materno (σ_{mpe}^2), fenotípica (σ_p^2), as herdabilidades direta (h_D^2) e materna (h_{mat}^2), e fração de variância devido matriz (c^2) obtidas para peso ao nascimento, peso a desmama e peso ao sobreano.

	σ_{α}^2	σ_{mat}^2	σ_{mpe}^2	σ_p^2	h_D^2	h_{mat}^2	c^2
Peso ao nascimento							
Média	2,83	0,66	0,31	11,43	0,25	0,06	0,03
HDI-I¹	2,68	0,58	0,24	11,36	0,23	0,05	0,02
HDI-S²	3,00	0,74	0,38	11,50	0,26	0,06	0,03
Peso à desmama							
Média	67,78	41,90	56,38	331,17	0,20	0,12	0,17
HDI-I¹	63,95	39,02	54,26	329,62	0,19	0,11	0,16
HDI-S²	71,60	44,61	58,83	332,71	0,21	0,13	0,18
Peso ao sobreano							
Média	313,48	-	-	646,34	0,48	-	-
HDI-I¹	307,10	-	-	643,20	0,47	-	-
HDI-S²	319,40	-	-	649,20	0,49	-	-

¹Limite inferior do HDI. ²Limite superior do HDI.

A estimativa da herdabilidade direta para PS foi de alta magnitude indicando possibilidade de ganho genético pela seleção. Nepomuceno et al. (2017), Felipe et al. (2021) e Rodrigues (2020) obtiveram estimativas de herdabilidade para peso ao sobreano entre 0,40 e 0,59. Essa variação das estimativas genéticas do PS pode estar ligada as diferenças ambientais e sistema de produção, bem como ao sistema de coleta de dados.

2.3.2 GWAS

Após a realização do WssGWAS com os programas da família BLUP90, foram identificadas as 20 janelas de maior contribuição à variabilidade genética para PN, PD e PS (Figura 1, 2 e 3). Os resultados do WssGWAS e os gráficos *Manhattan plot* evidenciam que as características ponderais analisadas são poligênicas, apresentando diversas regiões no cromossomo 14 que influenciam a expressão dos fenótipos. As posições das 20 maiores janelas adjacentes, bem como as variâncias explicadas para cada uma, estão presentes nas Tabelas Suplementares 1, 2 e 3, para PN, PD e PS, respectivamente.

Observa-se nos gráficos um padrão semelhante da variância genética aditiva explicada por regiões do cromossomo 14 das três características. Para variância genética materna, observa-se divergência entre as janelas encontradas para PN e PD, onde no PN há um pico evidenciado de janelas localizadas na região do início do cromossomo 14, enquanto no PD há três picos distribuídos no início, meio e final do cromossomo 14.

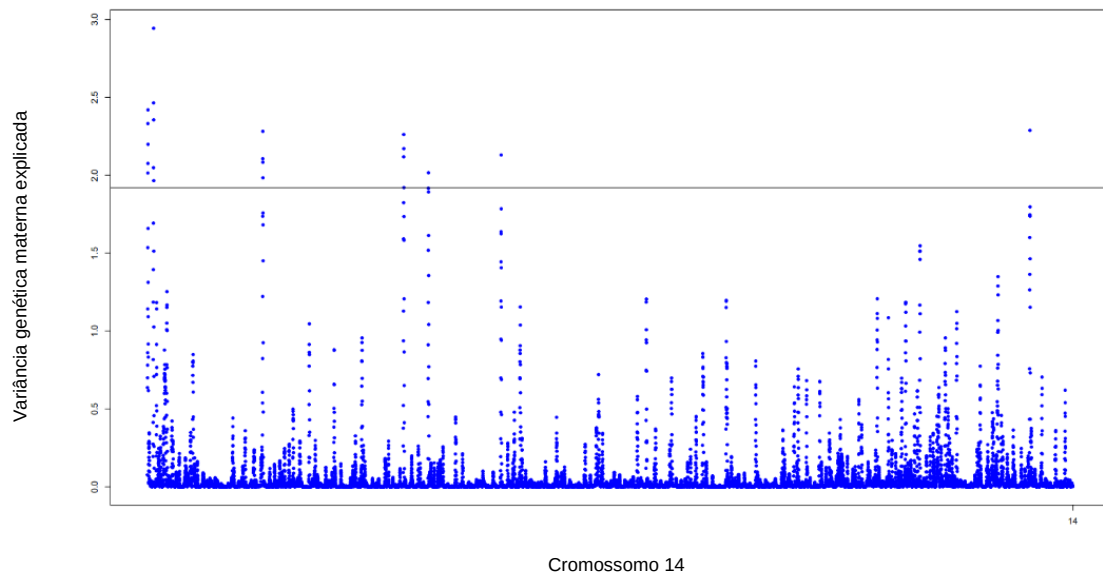
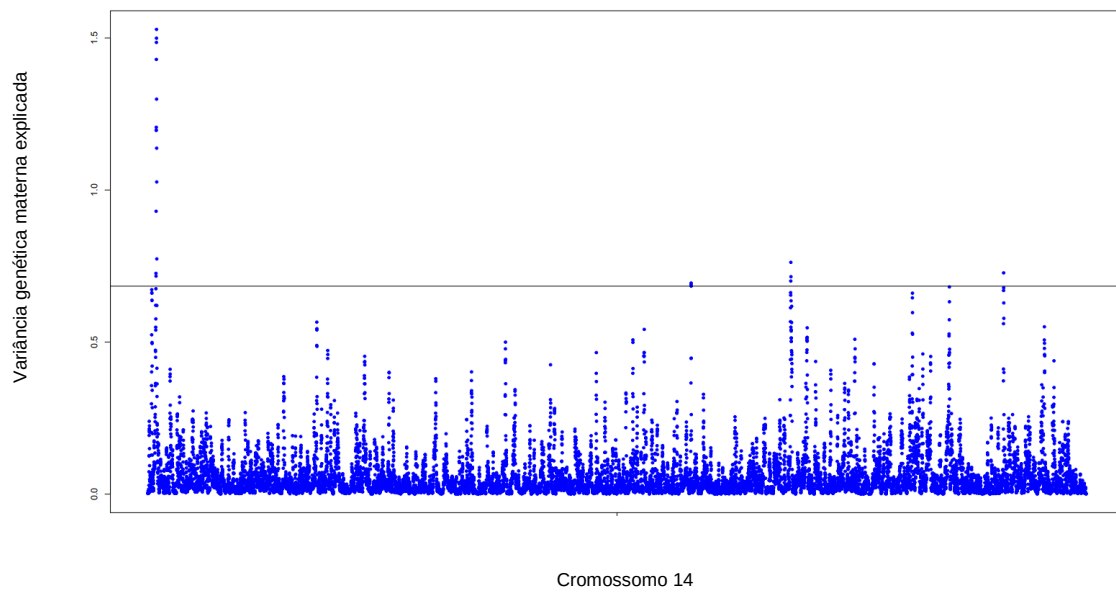
A**B**

Figura 1. *Manhattan plot* da variância genética explicada por janelas de 10 SNPs adjacentes do cromossomo 14 para peso ao nascimento. A linha horizontal representa o valor do limiar a partir do qual as 20 janelas com maior contribuição foram identificadas. **A)** variância genética aditiva direta. **B)** variância genética aditiva materna.

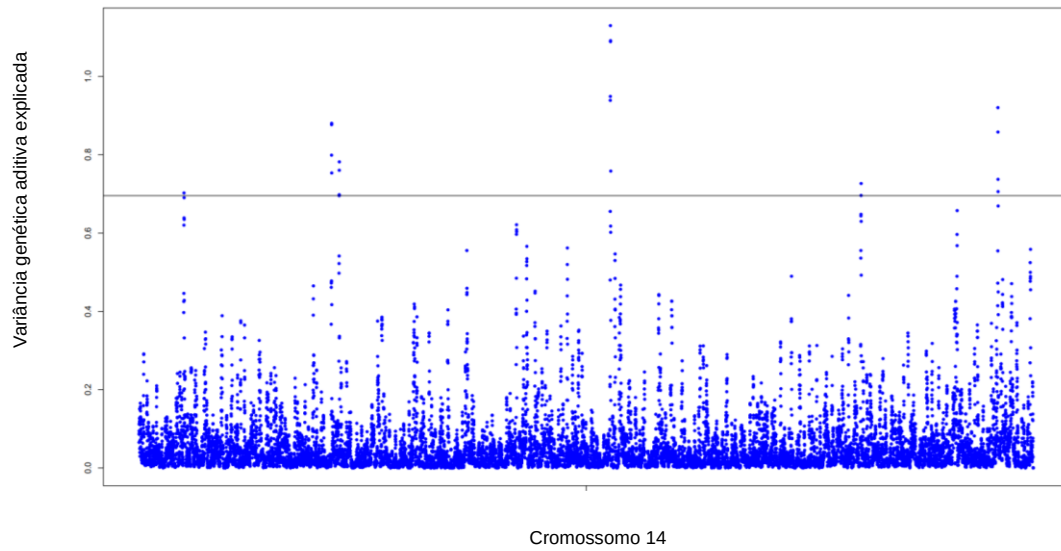
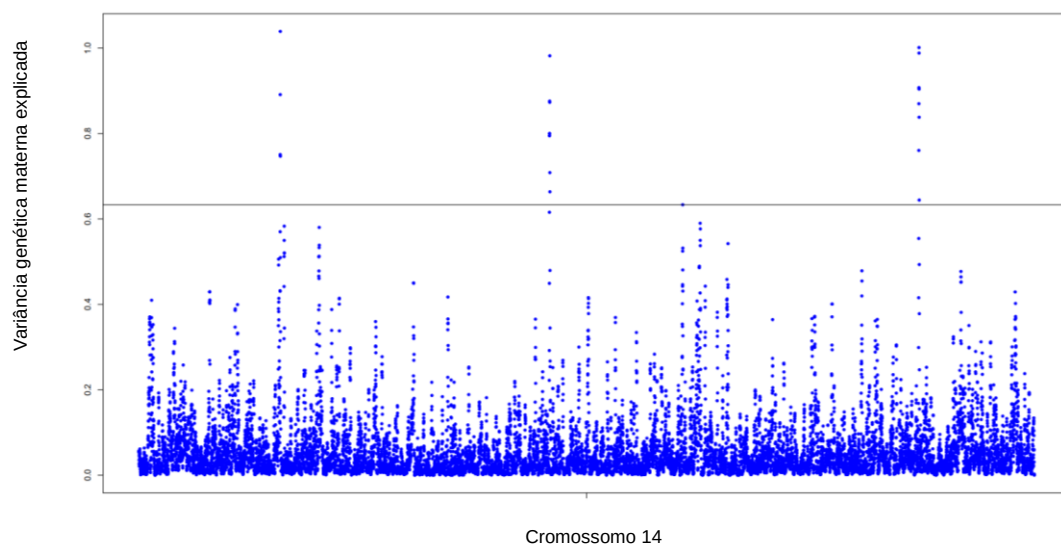
A**B**

Figura 2. *Manhattan plot* da variância genética explicada por janelas de 10 SNPs adjacentes do cromossomo 14 para peso a desmama. A linha horizontal representa o valor do limiar a partir do qual as 20 janelas com maior contribuição foram identificadas. **A)** variância genética aditiva direta. **B)** variância genética aditiva materna.

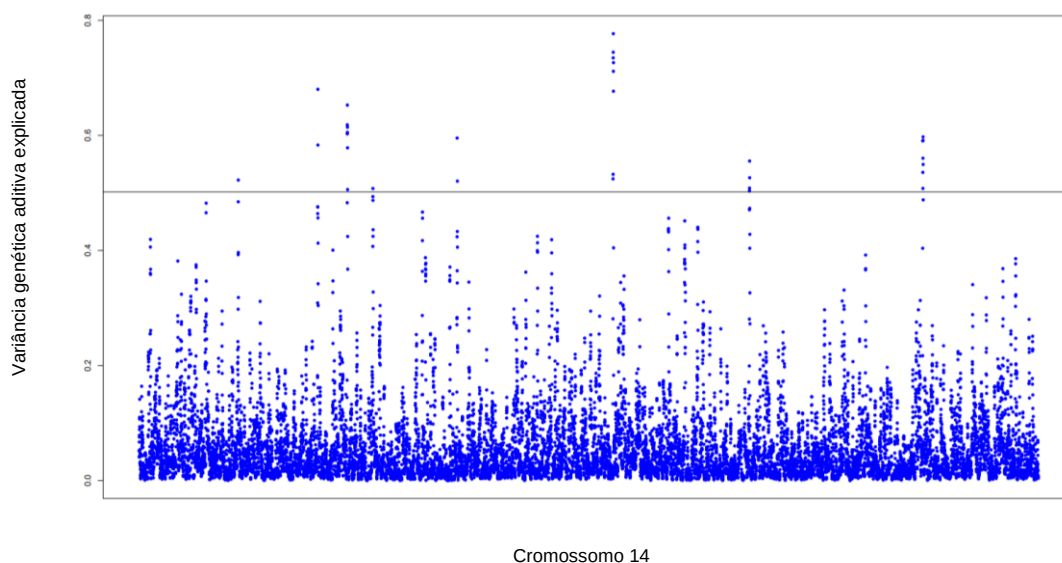


Figura 3. *Manhattan plot* da variância genética aditiva direta explicada por janelas de 10 SNPs adjacentes do cromossomo 14 para peso ao sobreano. A linha horizontal representa o valor do limiar a partir do qual as 20 janelas com maior contribuição foram identificadas.

É possível notar janelas em comum (Tabela 4) que explicam parte da variação para as três características, podendo-se inferir a presença de efeitos pleiotrópicos nessas regiões. Para janelas em comum entre PN e PD, a sobreposição (S) 2 identificou o gene *TBC1D31*. Enquanto para janelas em comum entre PD e PS, as S 1, 3, 4, 5 e 6 sobrepuseram os genes *TBC1D31*, *ENSBTAT00000081420*, *TRPA1*, *STMN2* e *CPQ*, respectivamente. O efeito pleiotrópico pode ser melhor investigado em uma abordagem de GWAS multicaracterísticas.

Estudos demonstram que o cromossomo 14 tem importância para características de crescimento e peso de carcaça (Veneroni et al., 2010; Bhuiyan et al., 2018; Hay e Roberts, 2018). Os resultados do presente trabalho evidenciam a importância do cromossomo 14, já que vários SNPs foram identificados com contribuição para as características de peso ao nascimento, peso a desmama e peso ao sobreano. Para compreender os genes e QTLs relacionados com estes SNPs, foram realizadas análises de enriquecimento funcional.

Tabela 4. Sobreposições (S) entre janelas identificadas nas análises WssGWAS para as características PN, PD e PS.

S	Janela	Início	Fim	Var ¹	Janela	Início	Fim	Var ¹
1	PD	16664647	16695012	0,88	PS	16664647	16695012	0,40
2	PN	17431022	17492989	0,55	PD	17453361	17523884	0,78
3	PD	35511046	35591127	0,79	PS	35511046	35591127	0,42
4	PS	35511046	35591127	0,42	PD	35568640	35627074	0,98
5	PD	42808565	42846993	1,13	PS	42822301	42853303	0,78
6	PD	67307041	67337501	0,73	PS	67307041	67337501	0,39
7	PN	71962685	72003733	0,46	PD	71962685	72003733	1,00
8	PN/D	71962685	72003733	0,46/1,0	PS	71977644	72029041	0,60

¹Variância explicada pela janela.

2.3.3 Detalhamento dos genes candidatos e QTLs

Como análise complementar do GWAS, as 20 maiores janelas adjacentes foram avaliadas nos bancos de dados do Ensembl Gene 104 e CattleQTLdb, para obtenção dos genes e QTLs associados com cada característica. Para PN, foram identificados 474 QTLs e 41 genes, 490 QTLs e 34 genes para PD, e 275 QTLs e 14 genes para PS. Os gráficos contendo as proporções de QTLs em cada classe definida no CattleQTLdb (exterior, saúde, carne e carcaça, leite, produção e reprodução), e as quantidades de QTLs relacionadas com as três características analisadas foram dispostos nas Figuras Suplementar 1, 2 e 3.

Genes identificados devem ser compreendidos como de importância para as características estudadas uma vez que foram identificados em sobreposição com as regiões genômicas identificadas nas análises de GWAS (Tabelas Suplementares 4, 5 e 6).

O gene *CSMD3* é um gene candidato novo que tem associação com características de conformação corporal (Toro-Ospina et al., 2022). Os genes

PCMTD1 e *STK3* foram descritos associados com características de crescimento e desenvolvimento corporal em bovinos taurinos (Smith et al., 2019; Duan et al., 2021), assim como o gene *PLAG1*, o qual foi confirmado desempenhar um papel importante em mudanças recentes na estatura de bovinos (Boitard et al., 2016), além de apresentar função pleiotrópica para características ponderais e reprodutivas (Mota et al., 2022).

O gene *ZC3H3* já foi identificado em outro estudo de GWAS no cromossomo 14, para características de carne e carcaça na raça Wagyu, esse gene foi associado com peso do bife para cortes parecidos com “fraldinha” e “patinho” (Zhang et al., 2019). Os genes *OXR1*, *TBC1D31*, *TG* e *DGAT1* são encontrados na literatura sendo associados com características de carcaça em raças taurinas (Gan et al., 2008; Yuan et al., 2013; Zhang et al., 2019; Srikanth et al., 2020; Khan et al., 2021), onde *TG* está associado à deposição de gordura e *DGTA1* é usado como gene candidato para melhoramento da qualidade da gordura na carcaça em bovinos chineses comerciais.

Os genes *ZFPM2*, *PCMTD1* e *NCOA2* foram associados com características de reprodução em bovinos zebuínos e taurinos (Camargo et al., 2015; Mota et al., 2017; Fonseca et al., 2018). Também foram identificados genes associados com características de produção de leite, como *TRAPPC9*, *DGTA1*, *SLC52A2*, *NCOA2*, *MROH1* (Wang et al., 2015; Ariyaratne et al., 2021). O gene *HSF1* é um gene candidato associado com tolerância de calor em bovinos, envolvidos na redução do dano celular causado pelo estresse térmico (Baena et al., 2018), animais mais adaptados a climas tropicais tendem a ter menor expressão desse gene, podendo ser utilizado como indicador de termo tolerância para bovinos taurinos (Rocha et al., 2019).

Apesar de até o presente não serem diretamente associados com características de crescimento, os genes identificados que estão associados com características de carne e carcaça, reprodução, produção de leite e tolerância de calor podem ser de importância para as características estudadas, já que as análises destacaram estas regiões e vários estudos indicam que as características estudadas apresentam correlação com estas outras mencionadas neste parágrafo (Fries e

Albuquerque, 1998; Mello, 2007; Lima Neto et al., 2009; Baena et al., 2018; Rocha et al., 2019; Ribeiro, 2020).

O enriquecimento funcional dos genes anotados foi realizado para identificar os termos de ontologia genética (GO) e agrupar os genes candidatos considerando a similaridade em termos associados com componente celular (CC), função molecular (FM) e processo biológico (PB). Para termos de PB, os genes *CPQ* e *TG* foram associados a geração do hormônio da tireoide (GO:0006590). Os genes *IL7* e *LY6E* foram associados com crescimento de órgãos (GO:0035265). Os genes *MMP16* e *SCX* foram associados ao processo de ossificação endocondral (GO:0001958). E os genes *SLA*, *SCX*, *TRAPPC9* e *ZFPM2* foram associados com processos de diferenciação celular (GO:0030154). Nos termos de CC, os genes *HSF1* e *RMDN1* foram associados com o polo do fuso mitótico (GO:0097431). Por fim, para termos de FM, os genes *HSF1* e *NCOA2* foram associados a enzima RNA polimerase II na região reguladora da transcrição intrônica (GO:0001162) e os genes *KLF10*, *PLAG1*, *HSF1* e *SCX* foram associados a atividade do fator de transcrição de ligação ao DNA que ativa ou aumenta a transcrição de conjuntos de genes específicos transcritos pela RNA polimerase II (GO:0001228).

Na figura 4, é possível observar a interação dos genes identificados para as três características em conjunto. Essa rede foi construída considerando os 89 genes candidatos encontrados no PN, PD e PS. As análises no STRING mapearam 83 genes, e após a remoção dos genes sem interação na rede, restaram 56 proteínas e 99 interações na análise da rede de genes. Pelo STRING era esperado que a rede encontrasse 32 interações ($p\text{-value} < 10^{-16}$), revelando que as proteínas tem mais interações entre elas do que esperado para um cenário aleatório.

O enriquecimento indica que as proteínas são ao menos parcialmente conectadas biologicamente (Szklarczyk et al., 2021). A rede de genes apresenta um *cluster* formado principalmente por genes identificados nas janelas do efeito materno do peso ao nascimento, onde destaca-se os genes *ADCK5*, *BOP1*, *DGAT1*, *EEF1D* e *FBXL6*. Observa-se também uma rede de genes menor, sem interação com a rede principal, contendo 3 genes (*SNAI2*, *MMP16* e *SCK*), oriundos do PN e PS.

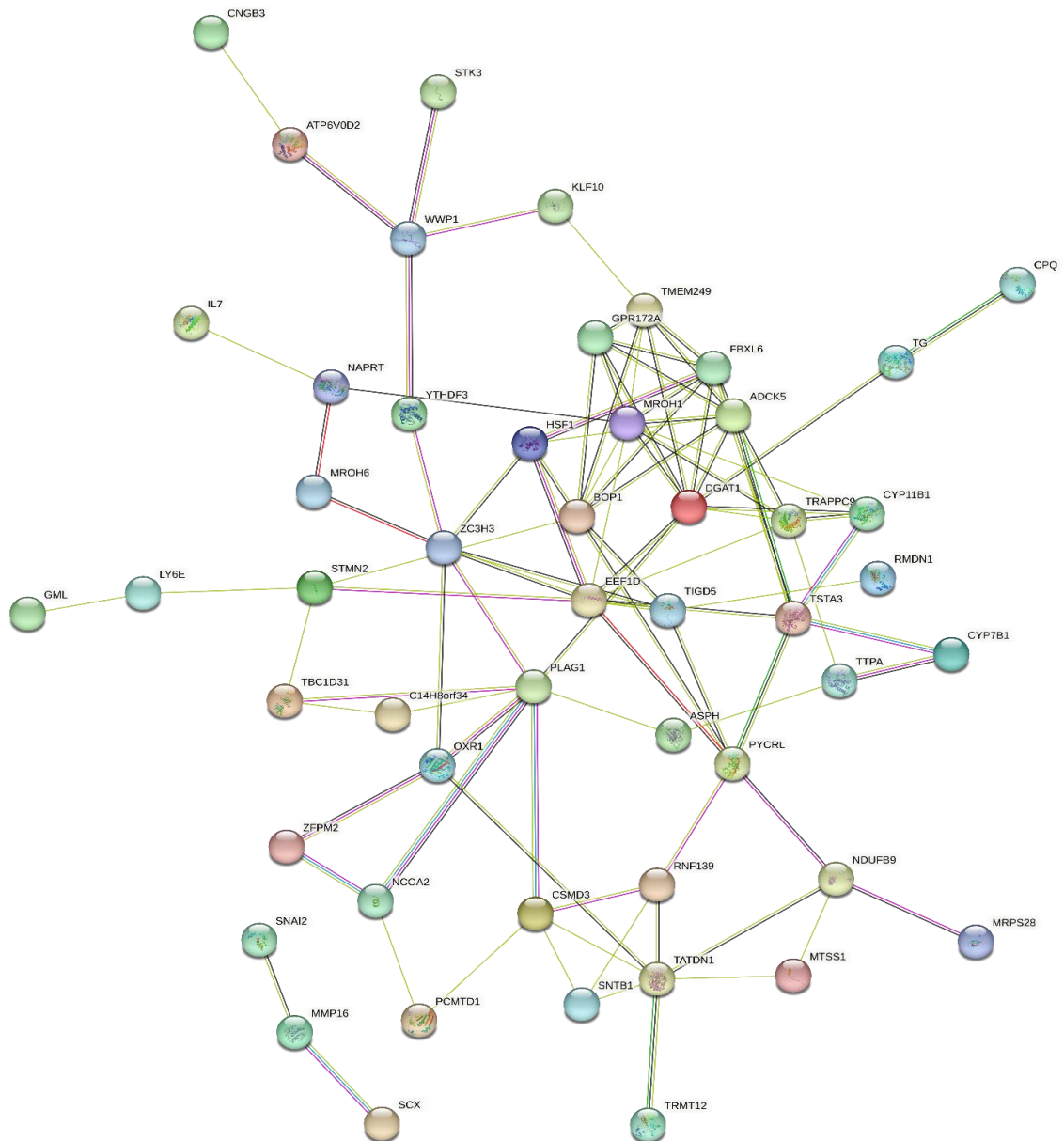


Figura 4. Interação da rede de genes candidatos identificados para as três características em conjunto.

2.3.4 Variâncias explicadas pela classe funcional

2.3.4.1 Similaridades entre as classes anotadas

Para a obtenção das variâncias genéticas explicadas pelas classes funcionais, somente o banco de dados de animais sequenciados foram utilizados. As GRM foram construídas para cada uma das 14 classes funcionais (Tabela 1), para cada característica, com uso do programa GCTA. Para determinar as similaridades entre

as GRM das classes, foi calculada a distância Euclidiana entre elas, conforme metodologia adotada por Koufariotis et al. (2018), onde foi gerado um mapa de calor (*heat map*) baseado na semelhança entre elas (Figura 5, Figura Suplementar 4). Quanto maior a similaridade entre as informações das GRM das classes, menor a distância Euclidiana e mais escura é a cor.

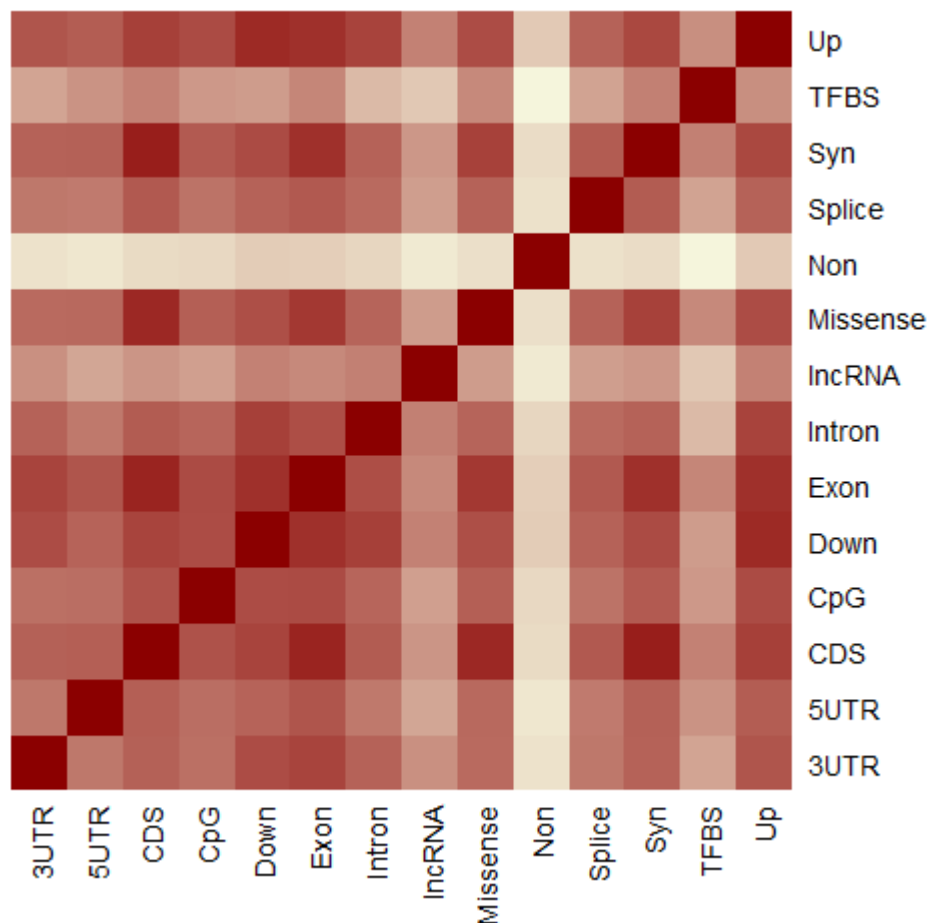


Figura 5. Mapa de calor da distância Euclidiana entre as GRM de todas as classes funcionais.

As GRM entre as classes CDS e Syn, CDS e *Exon*, CDS e *Missense*, Up e Down, Up e *Exon* e *Exon* e Syn foram as mais similares o que, segundo Koufariotis et al. (2018) indica haver alto LD entre as variantes de cada membro desses pares de classes (Koufariotis et al., 2018). As classes Non, TFBS e lncRNA foram as que menos possuem similaridades com outras classes.

2.3.4.2 Frequências alélicas

As frequências alélicas foram calculadas para cada classe funcional (Figura 6) e agrupadas segundo a frequência do alelo mais raro (MAF) nas seguintes classes: maf1 (menor que 0,05), maf2 (entre 0,05 e 0,1), maf3 (entre 0,1 e 0,2), maf4 (entre 0,2 e 0,3), maf5 (entre 0,3 e 0,4), e maf6 (entre 0,4 e 0,5). Para a maioria das classes, as variantes apresentaram frequência maior dentro da maf1, seguido pela maf3. As classes CDS, CpG e Exon apresentaram maior distribuição das variantes em maf1, podendo indicar que as variantes nessas classes podem estar mais sob seleção do que as demais classes. Já Splice, Intron e Non apresentaram mais variantes na maf3. No geral, maf4, maf5 e maf6 tem as menores porcentagens de variantes, em todas as classes.

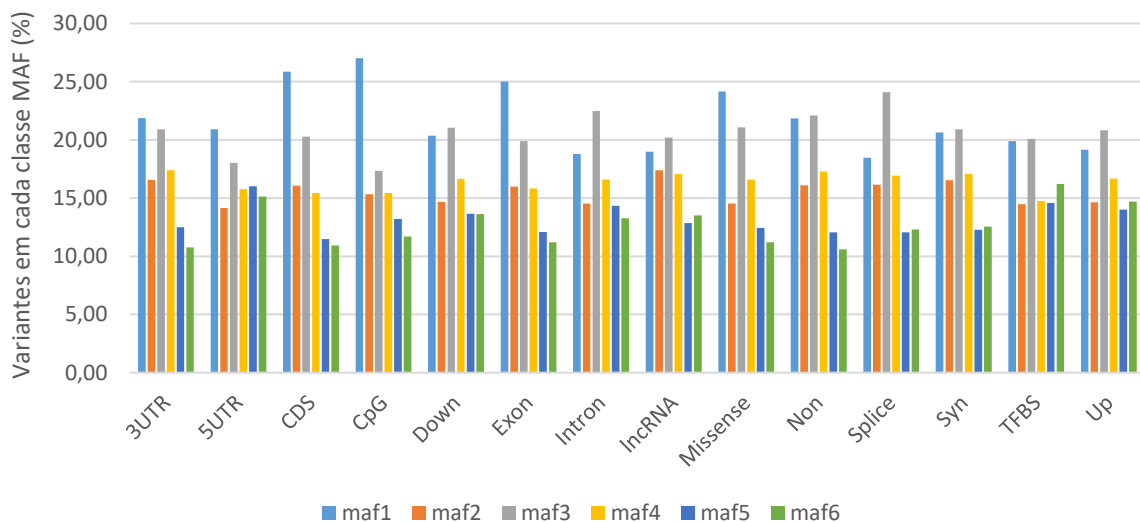


Figura 6. Distribuição das frequências alélicas em cada classe funcional, onde maf1 (<0,05), maf2 (0,05-0,1), maf3 (0,1-0,2), maf4 (0,2-0,3), maf5 (0,3-0,4), e maf6 (0,4-0,5).

A variância explicada por classe e por SNP de cada classe MAF foram apresentadas em gráficos de barras (Figura 7). As classes maf3 e maf5 foram as que mais contribuíram para a variância do PN, onde a classe maf5 capturou maior proporção da variância. Para PS, as classes entre maf3 e maf6 explicaram proporções semelhantes das variâncias, enquanto a maf2 explicou menor proporção. Para PD a

variância também foi capturada para as classes entre maf2 e maf6, porém com distribuição diferente da PD, em que a classe de MAF intermediária capturou maior proporção da variância capturada. Bhuiyan et al. (2018) encontraram maior contribuição para a variância de variantes classificadas em MAF 0,1 a 0,2 (maf3) e 0,3 a 0,4 (maf5) em características de carcaça, condizendo com o observado para PN e PS, apesar das características e raças diferirem entre os estudos.

Para PN, cada SNP da classe maf5 explicou uma parte maior da variância em relação aos SNPs das demais classes. Embora a variância capturada pelos SNPs da maf4 tenha sido maior para PD, a variância média explicada por cada SNP da classe maf3 teve maior valor, isso ocorre devido ao menor número de SNPs classificados na maf3. Como mencionado anteriormente, a variância explicada para o PS foi bem distribuída entre as classes acima maf1, porém com maior variância explicada para cada SNP da maf5, devido a ocorrência de menor número de variantes nessa classe. Apesar da presença de variantes na classe maf1, não houve variância capturada nesta classe.

Esses resultados estão de acordo com o esperado, uma vez que a variância explicada por cada variante depende de seu efeito, mas também das frequências dos alelos (Hill et al., 2008). Também pode-se observar que variantes com MAF acima de 0,05 devem ser enfatizadas nos estudos para PN, PD e PS. Os resultados das análises com a GRM para as diferentes classes genômicas, juntamente com os resultados das análises anteriores de WssGWAS e prospecção de genes candidatos, indicam arquiteturas genômicas complexas para características ponderais estudadas na raça Nelore.

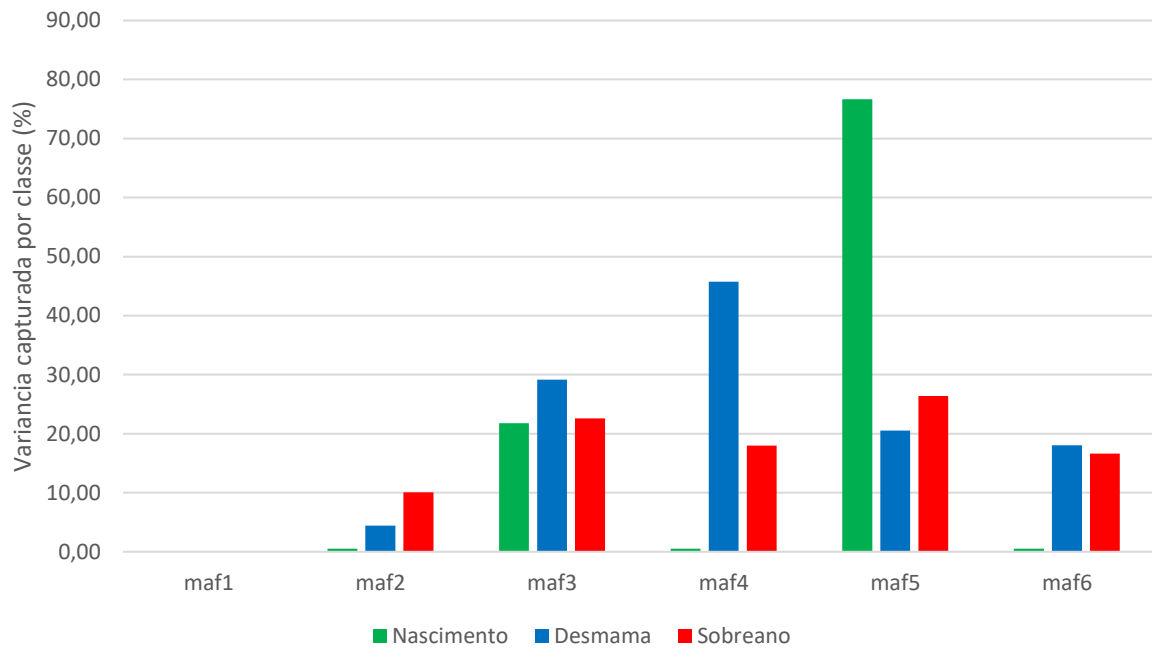
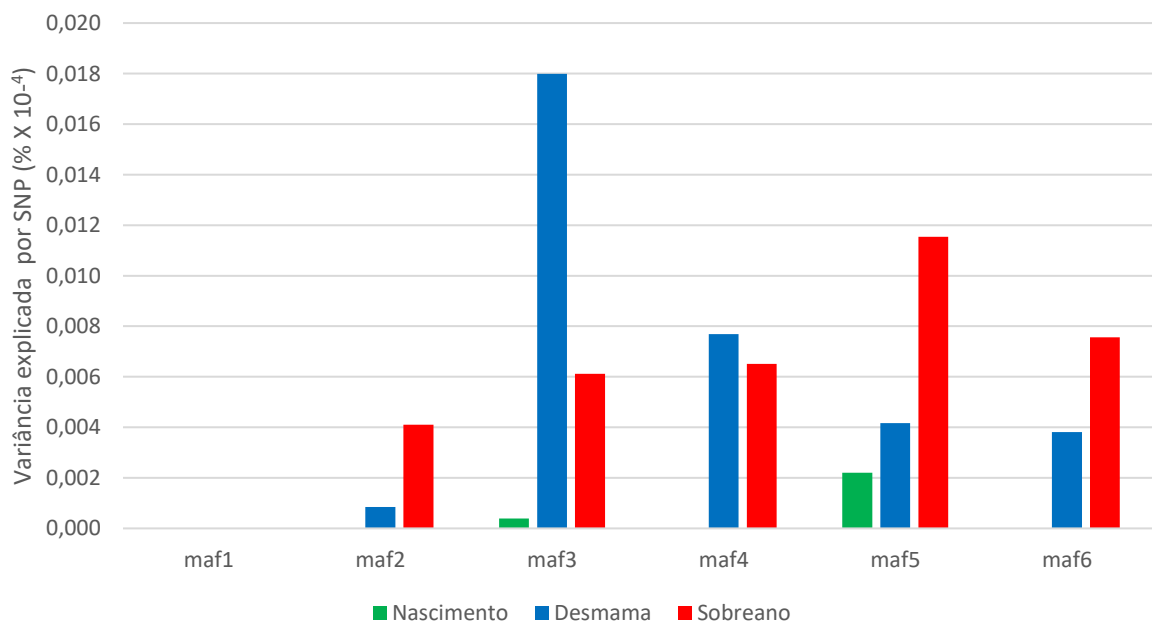
A**B**

Figura 7. Proporção da variância capturada e variância explicada para as classes MAF, para peso ao nascimento, peso a desmama e peso ao sobreano. **A)** variância capturada por cada classe **B)** variância explicada por SNP em cada classe, segundo metodologia descrita no material e métodos.

2.3.4.3 Classes funcionais

As classes funcionais Non, 5UTR e lncRNA capturaram maior proporção da variância para PN, PD e PS, respectivamente (Figura 8). O PN apresentou variância apenas para a classe Non. Para PD foi capturada variâncias para as classes 5UTR, Intron e TFBS. O PS foi a característica com maior quantidade de classe funcionais com variância capturada, sendo elas CpG, Down, Intron, lncRNA, Non, Splice e Up. Observa-se que as classes Splice e 5UTR tem maior destaque na variância explicada por SNPs do que na variância capturada, isso acontece por serem as duas classes com menor número de variantes identificadas, sendo 390 e 799 variantes, respectivamente. Estes resultados estão de acordo com Koufariotis et al. (2018), que estudando classes funcionais para características de leite em bovino da raça Holandês, observou variância explicada significativa para a classe Splice, indicando grande importância para priorização na seleção genômica de variantes.

O inverso pode ser observado para Intron, que apesar de ter aproximadamente 290 mil variantes, tem variância explicada menor do que Splice e 5UTR, indicando que essas variantes tem efeitos menores nas características, contribuindo pouco para a variação. Desta forma, é importante ressaltar que as classes 5UTR e Splice contem variantes que contribuem mais para a variação das características ponderais.

Diferente do resultado obtido, era esperado que as variantes da classe Missense apresentassem efeitos significativos nas características, devido essas variantes poderem alterar sequencias genéticas e sequencias de aminoácidos (Koufariotis et al.,2014). O mesmo foi observado para classes funcionais de regiões codificadoras, onde não obtivemos variância capturada nessas classes.

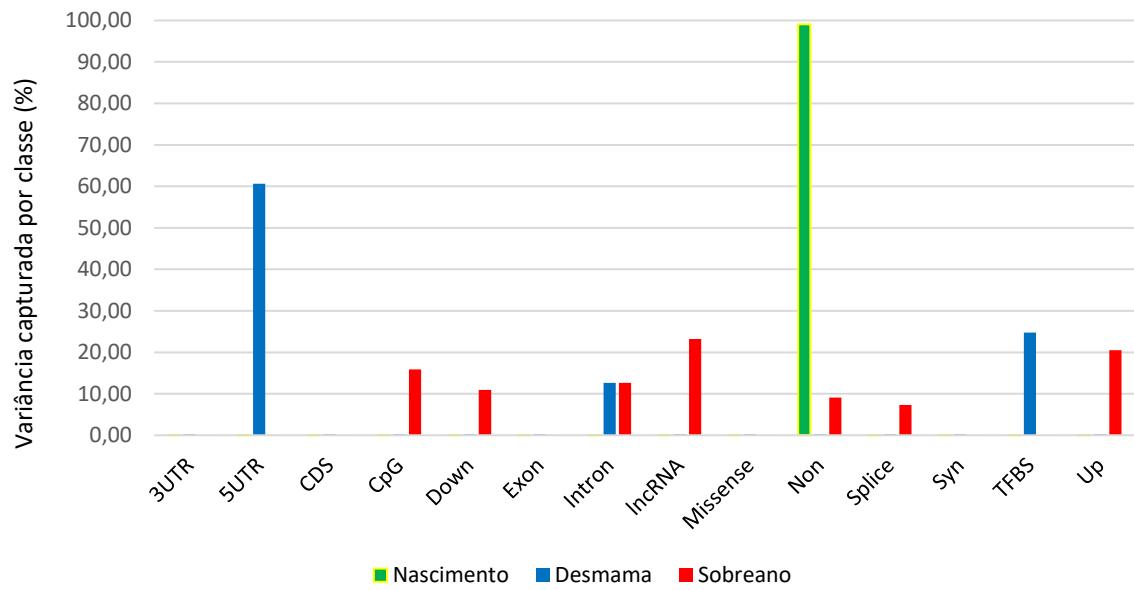
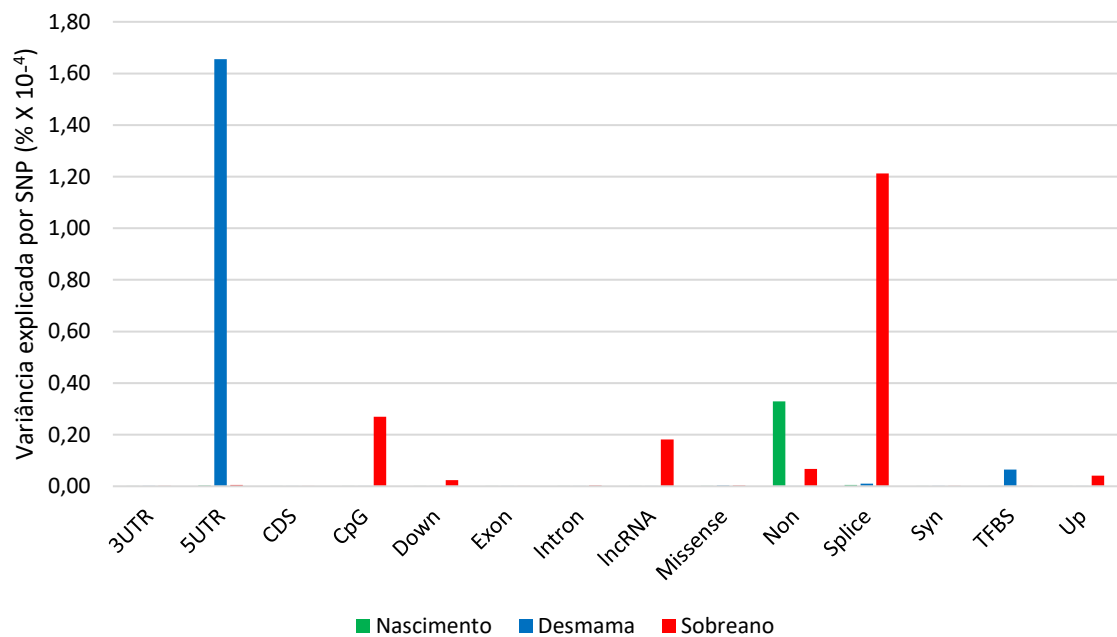
A**B**

Figura 8. Proporção da variância capturada e variância explicada para as classes funcionais, para peso ao nascimento, peso a desmama e peso ao sobreano. **A)** variância capturada por cada classe **B)** variância explicada por SNP em cada classe, segundo metodologia descrita no material e métodos.

Em geral, as classes funcionais de regiões regulatórias dos genes (5UTR, CpG, Down, lncRNA, TFBS e Up) têm maior significância no presente estudo e capturam grande proporção da variância para as características analisadas, estando de acordo com Koufariotis et al. (2018), que encontrou de média a alta variância para as classes a região regulatória, podendo potencialmente ser associada com o peso ao nascimento, peso à desmama e peso ao sobreano, através das funções regulatórias nos genes envolvidos com essas características. A classe não codificante Non foi a classe que obteve maior variância capturada por SNP dentre as características (Figura 8).

A fim de completar os resultados obtidos para as classes de variantes codificantes e regulatórias do genoma bovino, análises foram realizadas para calcular as variâncias capturadas e explicadas pelos SNP por um modelo conjunto das classes, no qual consiste em realizar a análise inserindo as classes funcionais simultaneamente no modelo. Porém, devido a estruturação do banco de dados, essa análise foi limitada pela quantidade de animais nos bancos de dados, pela quantidade de variantes nas classes e pela similaridade entre elas, impossibilitando condução de uma análise conjunta. Bhuiyan et al. (2018) documentou a mesma dificuldade para este tipo de análise conjunta quando o número de variantes na classe foi pequeno.

As variantes encontradas em regiões regulatórias e não codificantes tem importância nas análises de variantes associadas com características ponderais de interesse na seleção de bovinos Nelore. Estes resultados podem beneficiar a implementação da informação de sequência pelos programas de melhoramento genético que selecionam características ponderais, com foco em redução do custo de genotipagem pela priorização de variantes de regiões regulatórias na seleção genômica. As regiões enfatizadas como importantes no presente estudo devem ser investigadas em novos estudos de validação e de expressão gênica na raça Nelore para melhor conhecimento de suas funções biológicas e influências nos fenótipos de interesse. Também são necessários trabalhos futuros para verificar se o uso das classes funcionais pode contribuir para aumentar a acurácia dos GEBV, além da importância de validar os resultados obtidos, quando analisado o genoma completo.

2.4 Conclusão

Os resultados do GWAS sugerem que variantes associadas com pesos avaliados em diferentes idades estão distribuídas em diversas regiões do cromossomo 14, exibindo o padrão poligênico das características estudadas. Algumas regiões do cromossomo 14 podem apresentar efeito pleiotrópico em PN, PD e PS. Alguns dos genes candidatos identificados, como *CSMD3*, *PCMTD1*, *STK3* e *PLAG1* estão relacionados com conformação corporal e desenvolvimento muscular e os genes *HSF1*, *NCOA2*, *KLF10* e *PLAG1* tem funções moleculares associadas a regiões regulatórias. A anotação de variantes baseadas na função biológica e em categorias de MAF foram complementares aos resultados do GWAS, contribuindo para o entendimento de como as variantes atuam na variação fenotípica de características ponderais na raça Nelore, apresentando diferentes arquiteturas genéticas entre as três características investigadas. Variantes em classes de regiões regulatórias e *splice sites* empenham papéis importantes nas características estudadas e podem ser incluídas em avaliações genômicas para características ponderais na raça Nelore.

2.5 Referências bibliográficas

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ (2010) Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. **Journal of Dairy Science**, 93(2), 743-752.

Ariyaratne HBPC, Correa-Luna M, Blair HT, Garrick DJ, Lopez-Villalobos N (2021) Identification of genomic regions associated with concentrations of milk fat, protein, urea and efficiency of crude protein utilization in grazing dairy cows. **Genes**, 12(3), 456.

Baena MM, Tizioto PC, Meirelles SLC, Regitano LCDA (2018) HSF1 and HSPA6 as functional candidate genes associated with heat tolerance in Angus cattle. **Revista Brasileira de Zootecnia**, 47.

Barbosa ACB, Carneiro PLS, Rezende MPG, Ramos IO, Martins Filho R, Malhado CHM (2017) Parâmetros genéticos para características de crescimento e reprodutivas em bovinos Nelore no Brasil. **Archivos de Zootecnia**, v. 66, n. 255, p. 449-452.

Bhuiyan MAS, Lim D, Park M, Lee S, Kim Y, Gondro C, Park B, Lee S (2018) Functional partitioning of genomic variance and genome-wide association study for carcass traits in Korean Hanwoo cattle using imputed sequence level SNP data. **Frontiers in Genetics**, v.9, p.217.

Boitard S, Boussaha M, Capitan A, Rocha D, Servin B (2016) Uncovering adaptation from sequence data: lessons from genome resequencing of four cattle breeds. **Genetics**, 203(1), 433-450.

Bourdon C, Bardou P, Le Guillou S, Tosser-Klopp G, Le Provost F (2019) RumimiR: a detailed microRNA database focused on ruminant species. **Database**, Vol. 2019.

Camargo GM, Costa RB, Lucia G, Regitano LC, Baldi F, Tonhati H (2015) Polymorphisms in TOX and NCOA2 genes and their associations with reproductive traits in cattle. **Reproduction, Fertility and Development**, 27(3), 523-528.

Carlson M (2020) org.Bt.eg.db: Genome wide annotation for Bovine. R Packag. version 3.8.2., <https://doi.org/10.18129/B9.bioc.org.Bt.eg.db>.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. **Gigascience** 4:7-23.

Duan X et al (2021) Genome-wide association analysis of growth curve parameters in Chinese Simmental beef cattle. **Animals**, 11(1), 192.

Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W (2005) BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. **Bioinformatics**, 21(16), 3439-3440.

Euclides Filho K (2009) Evolução do melhoramento genético de bovinos de corte no Brasil. **Revista Ceres**, v. 56, n. 5, p. 620-626.

Evangelista AF, Cavalcante DH, Malhado CHM, Campelo JEG, Carvalho G, Barros Júnior CP, Sousa Junior SC (2020) Estimação de parâmetros genéticos para características de crescimento em bovinos Nelore Mocho da Região Norte do Brasil. **Livestock Research For Rural Development**, v. 32, n. 10.

Felipe EF, Silva NAM, Moraes GF (2021) Meta-análise para estimativas de herdabilidade de características ponderais em bovinos da raça nelore. **Brazilian Journal of Development**, 7(6), 59365-59379.

Fernandes Júnior GA et al (2021) Imputation accuracy to whole-genome sequence in Nellore cattle. **Genetics Selection Evolution**, 53(1), 1-10.

Fernandes LS (2022) **Estimativas de parâmetros genéticos e estudo de associação genômica ampla para comportamento de proteção materna em bovinos da raça nelore**. 80 f. Tese (Doutorado em Genética e Melhoramento Animal) – Unesp, Jaboticabal.

Fonseca PADS, Santos FC, Lam S, Suárez-Vega A, Miglior F, Schenkel FS, Diniz LFA, Id-Lahoucine S, Carvalho MRS, Cánovas A (2018) Genetic mechanisms underlying spermatid and testicular traits within and among cattle breeds: systematic review and prioritization of GWAS results. **Journal of Animal Science**, 96(12), 4978-4999.

Fragomeni BO, Lourenco DAL, Legarra A, VanRaden PM, Misztal I (2019) Alternative SNP weighting for single-step genomic best linear unbiased predictor evaluation of stature in US Holsteins in the presence of selected sequence variants. **Journal of Dairy Science**, 102(11), 10012-10019.

Fries LA, Albuquerque LG (1998) Pressuposições e restrições dos modelos animais com efeitos maternos em gado de corte. In: Costa MJRP, Cromberg VU Ed: **Comportamento Materno em Mamíferos: bases teóricas e aplicações aos ruminantes domésticos**. Jaboticabal: SBET-ETICO p.179-214.

Gan QF, Zhang LP, Li JY, Hou GY, Li HD, Gao X, Ren HY, Chen JB, Xu SZ (2008) Association analysis of thyroglobulin gene variants with carcass and meat quality traits in beef cattle. **Journal of Applied Genetics**, 49, 251-255.

Hay EH, Roberts A (2018) Genome-wide association study for carcass traits in a composite beef cattle breed. **Livestock Science**. 213, 35–43.

Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. **PLoS genetics**, v. 4, n. 2, p. e1000008.

Hu ZL, Park CA, Reecy JM (2022) Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. **Nucleic Acids Research**, Volume 50, Issue D1, Pages D956–D961. DOI: 10.1093/nar/gkab1116.

Hubbard TJP et al (2009) Ensembl 2009. **Nucleic Acids Research**.;37:D690–D697.

Kamei LM, Azambuja Ribeiro EL, Fonseca NAN, Muniz CADSD, Camiloti TV, Koritiaki NA, Souza Fortaleza AP (2017) Genetic parameters of growth traits in Nellore cattle. **Semina: Ciências Agrárias** 38(3), 1503-1509.

Khan MZ, Ma Y, Ma J, Xiao J, Liu Y, Liu S, Khan A, Khan IM, Cao Z (2021) Association of DGAT1 with cattle, buffalo, goat, and sheep milk and meat production traits. **Frontiers in Veterinary Science**, 8, 712470.

Klees S, Heinrich F, Schmitt AO, Gultas M (2021) agReg-SNPdb: A Database of Regulatory SNPs for Agricultural Animal Species. **Biology**, v. 10, n. 8, p. 790.

Koufariotis L, Chen YPP, Bolormaa S, Hayes BJ (2014) Regulatory and coding genome regions are enriched for trait associated variants in dairy and beef cattle. **BMC Genomics**, 15(1), 1-16.

Koufariotis LT, Chen YPP, Stothard P, Hayes BJ (2018) Variance explained by whole genome sequence variants in coding and regulatory genome annotations for six dairy traits. **BMC Genomics** 19, 237.

Lima Neto HR, Bergmann JAG, Gonçalves TDM, Araújo FRDC, Bezerra LAF, Saiz RD, Lobo RB, Silva MA (2009). Parâmetros genéticos para características de carcaça avaliadas por ultrassonografia em bovinos da raça Guzerá. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, 61, 251-258.

Maiorano AM, Cardoso DF, Carneiro R, Junior GAF, Albuquerque LG, Oliveira HN (2022) Signatures of selection in Nelore cattle revealed by whole-genome sequencing data. **Genomics**, 114, 110304.

Marín-Garzón NA et al (2021) Genome-wide scan reveals genomic regions and candidate genes underlying direct and maternal effects of preweaning calf mortality in Nellore cattle. **Genomics**, v. 113, n. 3, p. 1386-1395.

Mello RO (2007). **Eficiência produtiva e econômica, características da carcaça e qualidade da carne de bovinos mestiços confinados e abatidos com diferentes pesos corporais**. 160 f. Tese (Doutorado em Zootecnia) – UFV, Viçosa.

Misztal I, Tsuruta S, Strabel T, Auvray B, Bruet T, Lee DH (2002) BLUPF90 and related programs (BGF90). In: 7th WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION. **Anais...** Montpellier: WCGALP, p. 19-23.

Mota LF, Carvajal AB, Bernardes PA, Buzanskas ME, Baldi F, Lôbo RB, Munari DP (2022) Integrating genome-wide association study and pathway analysis reveals physiological aspects affecting heifer early calving defined at different ages in Nelore cattle. **Genomics**, 114(4), 110395.

Mota RR, et al (2017) Genome-wide association study and annotating candidate gene networks affecting age at first calving in Nellore cattle. **Journal of Animal Breeding and Genetics**, 134(6), 484-492.

Nepomuceno LL, Andrade RJ, Lopes FB, Lira TS, Vieira LF, Santos GCJ, Ferreira JL (2017) Genetic associations between scrotal circumference and productive and reproductive traits in herd Nellore raised in northern Tocantins State, Brazil. **Revista Acadêmica: Ciência Animal**, v. 10, no. 3, pp. 253-261.

Pedrosa VB, Eler JP, Ferraz JBS, Pinto LFB (2014) Utilização de modelos unicaracterística e multicaracterística na estimação de parâmetros genéticos na raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 66, no. 6, pp. 1802-1812.

Plummer M, Best N, Cowles K, Vines K (2006) CODA: Convergence Diagnosis and Output Analysis for MCMC. **R news** 6:7-11.

Ribeiro VMP (2020) **Genes candidatos para características de crescimento, reprodução e resistência em bovinos de corte**. 112 f. Tese (Doutorado em Zootecnia) – UFMG, Belo Horizonte.

Rocha RDFB, Baena MM, Estopa AC, Gervásio IC, Ibelli AMG, Gionbelli TRS, Gionbello MP, Freitas RTF, Meirelles SLC (2019) Differential expression of HSF1 and HSPA6 genes and physiological responses in Angus and Simmental cattle breeds. **Journal of Thermal Biology**, 84, 92-98.

Rodrigues MR (2020) **Parâmetros genéticos e fenotípicos de características de crescimento em bovinos da raça Nelore criados no norte do Brasil**. 61 f. Dissertação (Mestrado em Sanidade Animal e Saúde Pública nos Trópicos) – UFT, Araguaína.

Smith JL, Wilson ML, Nilson SM, Rowan TN, Oldeschulte DL, Schnabel RD, Decker JE, Seabury CM (2019) Genome-wide association and genotype by environment interactions for growth traits in US Gelbvieh cattle. **BMC Genomics**, 20(1), 1-13.

Srikanth K et al (2020) A Gene-Set Enrichment and Protein–Protein Interaction Network-Based GWAS with Regulatory SNPs Identifies Candidate Genes and Pathways Associated with Carcass Traits in Hanwoo Cattle. **Genes**, 11(3), 316.

Szklarczyk D et al (2021) The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. **Nucleic Acids Research**, Volume 49, Database issue. D605-D612.

Toro-Ospina AM, Rios ACH, Santos WB, Schettini GP, Aristizabal VHV, Claros GT, Morea EGO (2022) Genetic Architecture and Signatures of Selection in the Caqueteño Creole (Colombian Native Cattle). **Diversity**, 14(10), 828.

VanRaden PM (2008) Efficient methods to compute genomic Predictions. **Journal of Dairy Science** 91:4414-4423.

Veneroni GB, Meirelles SL, Grossi DA, Gasparin G, Ibelli AMG, Tizioto PC, Oliveira HN, Alencar MM, REGITANO LDA (2010) Prospecting candidate SNPs for backfat in Canchim beef cattle. **Genetics and Molecular Research** 9:1997-2003.

Wang H, Misztal I, Aguilar I, Legarra A, Muir WM (2012) Genome-wide association mapping including phenotypes from relatives without genotypes. **Genetic Research** 94:73-83.

Wang X et al (2015) Genome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. **BMC Genetics**, 16(1), 1-9.

Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for Genome-wide Complex Trait Analysis. **American Journal Human Genetics**, v. 88, n. 1, p. 76-82.

Yu G, Wang L-G, Han Y, He Q-Y. 2012 clusterProfiler:an R package for comparing biological themes among gene clusters. **OMICS: a journal of integrative biology**. 16, 284–287.

Yuan Z, Li J, Li J, Gao X, Gao H, Xu S (2013) Effects of DGAT1 gene on meat and carcass fatness quality in Chinese commercial cattle. **Molecular Biology Reports**, 40, 1947-1954.

Zhang R et al (2019) Genome-wide association study identifies the PLAG1-OXR1 region on BTA14 for carcass meat yield in cattle. **Physiological Genomics**, 51(5), 137-144.

Zhao L et al (2021) NONCODEV6: an updated database dedicated to long Non-coding RNA annotation in both animals and plants. **Nucleic Acids Research**, v.49, n. D1, p. D165-D171.

Zhou Y et al (2020) Comparative whole genome DNA methylation profiling across cattle tissues reveals global and tissue-specific methylation patterns. **BMC Biology**, v. 18, n. 1, p. 1-17.

APÊNDICE

Tabela Suplementar 1. Posição das 20 janelas com maior variância genética no cromossomo 14 para o peso ao nascimento.

Genético direto			Genético materno		
Janela		Variância	Janela		Variância
Início	Fim		Início	Fim	
66819965	66860322	1,27	11454332	1180817	1,52
17997177	18018110	0,99	58096523	85129331	0,76
71146583	71175238	0,78	75780114	75851385	0,73
74540410	74569760	0,72	58102959	58134245	0,71
76496799	76533620	0,70	48456599	48521600	0,69
3999352	4021404	0,68	70809198	70842864	0,68
52672328	52754547	0,64	1096344	1142313	0,67
12088985	12141193	0,58	561250	638162	0,67
65511307	65552983	0,56	67776868	67888128	0,66
3546946	3589056	0,55	588160	676575	0,64
17431022	17492989	0,55	58144518	58192822	0,62
59113223	59182529	0,53	13403132	13432383	0,56
3977117	4005326	0,53	79055874	79097926	0,55
59083232	59156531	0,50	59533518	59564637	0,55
79782432	79808853	0,49	43158811	43216895	0,54
20980495	21034025	0,48	70780318	70826500	0,52
49447249	49487643	0,48	63084103	63110924	0,51
39599354	39710072	0,47	41943842	41965713	0,51
12083341	12136189	0,46	59518050	59555054	0,50
71962685	72003733	0,46	29390659	29426711	0,50

Tabela Suplementar 2. Posição das 20 janelas com maior variância genética no cromossomo 14 para o peso a desmama.

Genético direto			Genético materno		
Janela		Variância	Janela		Variância
Início	Fim		Início	Fim	
42808565	42846993	1.13	11453318	11489850	1.04
79478326	79505107	0.92	71962685	72003733	1.00
16664647	16695012	0.88	35568640	35627074	0.98
17453361	17523884	0.78	11444087	11479294	0.89
67307041	67337501	0.73	35511046	35591127	0.79
4221099	4243733	0.70	51249961	51283737	0.63
67305322	67336207	0.70	52965513	53077021	0.59
76024875	76059594	0.66	11927034	11972236	0.58
32378045	32394139	0.62	15565785	15645092	0.58
76009262	76057522	0.60	55700820	55739359	0.54
33261290	33297963	0.57	11934153	11984541	0.52
37317717	37342138	0.56	15556723	15631968	0.51
82012774	82032086	0.56	11223815	11270109	0.50
28074344	28175797	0.55	52810818	52920347	0.49
43369028	43434359	0.55	67305322	67336207	0.48
82001294	82015406	0.52	76207102	76245704	0.48
61736112	61780005	0.49	55641755	55672854	0.46
79845053	79898057	0.48	67307041	67337501	0.45
80701769	80734004	0.47	23342093	23398807	0.45
44188176	44268544	0.47	53671344	53740799	0.44

Tabela Suplementar 3. Posição das 20 janelas com maior variância genética direta no cromossomo 14 para o peso ao sobreano.

Janela		Variância
Início	Fim	
42822301	42853303	0.78
15367788	15407903	0.68
18040234	18077555	0.65
71977644	72029041	0.60
27128043	27166761	0.59
57643589	57686713	0.55
8228215	8259410	0.52
19921096	19962709	0.51
5817336	5840240	0.48
24114804	24158384	0.47
49412518	49447249	0.45
51200087	51249961	0.45
52302869	52352362	0.44
33899470	33954905	0.42
35902169	35932857	0.42
1421501	1546349	0.42
16664647	16695012	0.40
67307041	67337501	0.39
24516834	24573854	0.39
80649315	80678212	0.38

Tabela Suplementar 4. Genes na região das 20 janelas com maior variância genética no cromossomo 14 para o peso ao nascimento.

Janela		Identificação do gene	Nome do Gene	Gene		Descrição do gene
Início	Fim			Início	Fim	
1096344	1142313	ENSBTAG00000014642	NAPRT	1127518	1131606	<i>nicotinate phosphoribosyltransferase</i>
1096344	1142313	ENSBTAG00000014643	EEF1D	1113551	1126356	<i>eukaryotic translation elongation factor 1 delta</i>
1096344	1142313	ENSBTAG00000016810	PYCR3	1101602	1108734	<i>pyrroline-5-carboxylate reductase 3</i>
1096344	1142313	ENSBTAG00000021474	GSDMD	1139890	1146602	<i>gasdermin D [Source:VGNC</i>
1096344	1142313	ENSBTAG00000034613	MROH6	1132386	1136717	<i>maestro heat like repeat family member 6</i>
1096344	1142313	ENSBTAG00000034691	GFUS	1094043	1099107	<i>GDP-L-fucose synthase</i>
1096344	1142313	ENSBTAG00000048160	TIGD5	1111425	1113356	<i>tigger transposable element derived 5</i>
1145433	1180817	ENSBTAG00000021472	ZC3H3	1153788	1215120	<i>zinc finger CCCH-type containing 3</i>
1145433	1180817	ENSBTAG00000021474	GSDMD	1139890	1146602	<i>gasdermin D</i>
20980495	21034025	ENSBTAG00000017492	PCMTD1	21013842	21062048	<i>protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing 1</i>
20980495	21034025	ENSBTAG00000043686	7SK	20998521	20998797	<i>7SK RNA</i>
20980495	21034025	ENSBTAG00000050040	U6	20991476	20991582	<i>U6 spliceosomal RNA</i>
29390659	29426711	ENSBTAG00000001299	CYP7B1	29199622	29460150	<i>cytochrome P450 family 7 subfamily B member 1</i>
3546946	3589056	ENSBTAG00000013955	TRAPPC9	3200372	3587738	<i>trafficking protein particle complex subunit 9</i>
41943842	41965713	ENSBTAG00000016284	IL7	41937221	41996546	<i>interleukin 7</i>
43158811	43216895	ENSBTAG00000001302	MRPS28	43120078	43222183	<i>mitochondrial ribosomal protein S28</i>
561250	638162	ENSBTAG00000000856	FBXL6	580951	583957	<i>F-box and leucine rich repeat protein 6</i>
561250	638162	ENSBTAG00000000857	SLC52A2	578057	580805	<i>solute carrier family 52 member 2</i>
561250	638162	ENSBTAG000000009811	BOP1	634453	652447	<i>BOP1 ribosomal biosynthesis factor</i>
561250	638162	ENSBTAG00000011064	ADCK5	556020	570498	<i>aarF domain containing kinase 5</i>
561250	638162	ENSBTAG00000020751	HSF1	613328	634349	<i>heat shock transcription factor 1</i>
561250	638162	ENSBTAG00000026356	DGAT1	603813	612791	<i>diacylglycerol O-acyltransferase 1</i>
561250	638162	ENSBTAG00000035158	TMEM249	584956	586597	<i>transmembrane protein 249</i>
561250	638162	ENSBTAG00000050112	SCRT1	594540	600190	<i>scratch family transcriptional repressor 1</i>
58096523	58129331	ENSBTAG00000008285	OXR1	57844325	58141613	<i>oxidation resistance 1</i>
58102959	58134245	ENSBTAG00000008285	OXR1	57844325	58141613	<i>oxidation resistance 1</i>
588160	676575	ENSBTAG00000009811	BOP1	634453	652447	<i>BOP1 ribosomal biosynthesis factor</i>
588160	676575	ENSBTAG00000009816	SCX	646600	648020	<i>scleraxis bHLH transcription factor</i>
588160	676575	ENSBTAG00000014458	MROH1	652785	702430	<i>maestro heat like repeat family member 1</i>
588160	676575	ENSBTAG00000020751	HSF1	613328	634349	<i>heat shock transcription factor 1</i>
588160	676575	ENSBTAG00000026356	DGAT1	603813	612791	<i>diacylglycerol O-acyltransferase 1</i>
588160	676575	ENSBTAG00000050112	SCRT1	594540	600190	<i>scratch family transcriptional repressor 1</i>

59083232	59156531	ENSBTAG00000001649	ZFPM2	58819937	59239114	<i>zinc finger protein, FOG family member 2</i>
59113223	59182529	ENSBTAG00000001649	ZFPM2	58819937	59239114	<i>zinc finger protein, FOG family member 2</i>
65511307	65552983	ENSBTAG000000026234	STK3	65475373	65785445	<i>serine/threonine kinase 3</i>
66819965	66860322	ENSBTAG000000011572	TSPYL5	66852944	66854263	<i>TSPY like 5</i>
66819965	66860322	ENSBTAG000000049287	-	66854284	66857370	-
67776868	67888128	ENSBTAG000000045108	bta-mir-2285dc	67820604	67820664	<i>bta-mir-2285dc</i>
74540410	74569760	ENSBTAG000000006877	MMP16	74381958	74776968	<i>matrix metalloproteinase 16</i>
75780114	75851385	ENSBTAG000000051390	CNBD1	75334895	75781689	<i>cyclic nucleotide binding domain containing 1</i>
76496799	76533620	ENSBTAG000000021092	ATP6V0D2	76530194	76586635	<i>ATPase H⁺ transporting V0 subunit d2</i>

Tabela Suplementar 5. Genes na região das 20 janelas com maior variância genética no cromossomo 14 para o peso a desmama.

Janela		Identificação do gene	Nome do Gene	Gene		Descrição do gene
Início	Fim			Início	Fim	
15556723	15631968	ENSBTAG00000020398	<i>TRMT12</i>	15616492	15618384	<i>tRNA methyltransferase 12 homolog</i>
15556723	15631968	ENSBTAG00000020399	<i>RNF139</i>	15593322	15607732	<i>ring finger protein 139</i>
15556723	15631968	ENSBTAG00000020405	<i>NDUFB9</i>	15553533	15560939	<i>NADH:ubiquinone oxidoreductase subunit B9</i>
15556723	15631968	ENSBTAG00000033446	<i>TATDN1</i>	15560985	15592470	<i>TatD DNase domain containing 1</i>
15565785	15645092	ENSBTAG00000020399	<i>RNF139</i>	15593322	15607732	<i>ring finger protein 139</i>
15565785	15645092	ENSBTAG00000033446	<i>TATDN1</i>	15560985	15592470	<i>TatD DNase domain containing 1</i>
16664647	16695012	ENSBTAG00000021156	<i>TBC1D31</i>	16631171	16698112	<i>TBC1 domain family member 31</i>
23342093	23398807	ENSBTAG00000004022	<i>PLAG1</i>	23330541	23375751	<i>PLAG1 zinc finger</i>
23342093	23398807	ENSBTAG000000049910	<i>CHCHD7</i>	23376061	23383133	<i>coiled-coil-helix-coiled-coil-helix domain containing 7</i>
23342093	23398807	ENSBTAG000000054153	<i>A0A3Q1M6D7_BOVIN</i>	23395823	23397050	-
28074344	28175797	ENSBTAG00000018557	<i>YTHDF3</i>	28141155	28204989	<i>YTH N6-methyladenosine RNA binding protein 3</i>
28074344	28175797	ENSBTAG00000022471	<i>TTPA</i>	28074682	28104548	<i>alpha tocopherol transfer protein</i>
32378045	32394139	ENSBTAG00000022588	<i>C14H8orf34</i>	32378189	32741893	<i>chromosome 14 C8orf34 homolog</i>
35511046	35591127	ENSBTAG000000048963	<i>A0A3Q1MM58_BOVIN</i>	35573943	35575232	-
35568640	35627074	ENSBTAG00000002062	<i>TRPA1</i>	35611321	35666918	<i>transient receptor potential cation channel subfamily A member 1</i>
35568640	35627074	ENSBTAG000000048963	<i>A0A3Q1MM58_BOVIN</i>	35573943	35575232	-
37317717	37342138	ENSBTAG000000048533	<i>A0A3Q1MIT2_BOVIN</i>	37324498	37325736	-
37317717	37342138	ENSBTAG000000054912	-	37314661	37320969	-
42808565	42846993	ENSBTAG00000020099	<i>STMN2</i>	42811857	42870107	<i>stathmin 2</i>
44188176	44268544	ENSBTAG00000014026	<i>PAG1</i>	44178038	44329953	<i>phosphoprotein membrane anchor with glycosphingolipid microdomains 1</i>
51249961	51283737	ENSBTAG000000038281	<i>CSMD3</i>	50969629	52273953	<i>CUB and Sushi multiple domains 3</i>
61736112	61780005	ENSBTAG00000014396	<i>KLF10</i>	61770129	61776507	<i>Kruppel like factor 10</i>
67305322	67336207	ENSBTAG00000011908	<i>CPQ</i>	66989157	67551536	<i>carboxypeptidase Q</i>
67307041	67337501	ENSBTAG00000011908	<i>CPQ</i>	66989157	67551536	<i>carboxypeptidase Q</i>
76009262	76057522	ENSBTAG000000036304	<i>CNGB3</i>	75914395	76103436	<i>cyclic nucleotide gated channel subunit beta 3</i>
76024875	76059594	ENSBTAG000000036304	<i>CNGB3</i>	75914395	76103436	<i>cyclic nucleotide gated channel subunit beta 3</i>
76207102	76245704	ENSBTAG00000015720	<i>WWP1</i>	76218264	76322858	<i>WW domain containing E3 ubiquitin protein ligase 1</i>

76207102	76245704	ENSBTAG00000015734	<i>RMDN1</i>	76178740	76212523	<i>regulator of microtubule dynamics 1</i>
82001294	82015406	ENSBTAG00000044126	<i>SNTB1</i>	81871605	82123649	<i>syntrophin beta 1</i>
82012774	82032086	ENSBTAG00000044126	<i>SNTB1</i>	81871605	82123649	<i>syntrophin beta 1</i>

Tabela Suplementar 6. Genes na região das 20 janelas com maior variância genética no cromossomo 14 para o peso ao sobreano.

Janela		Identificação do gene	Nome do Gene	Gene		Descrição do gene
Início	Fim			Início	Fim	
1421501	1546349	ENSBTAG00000004595	<i>GML</i>	1513837	1542041	<i>glycosylphosphatidylinositol anchored molecule like</i>
1421501	1546349	ENSBTAG00000004596	<i>TBC1D31</i>	1503538	1508569	-
1421501	1546349	ENSBTAG00000017040	<i>LY6E</i>	1422652	1426356	<i>lymphocyte antigen 6 family member E</i>
15367788	15407903	ENSBTAG00000020407	<i>MTSS1</i>	15383762	15553664	<i>MTSS I-BAR domain containing 1</i>
16664647	16695012	ENSBTAG00000021156	<i>TBC1D31</i>	16631171	16698112	<i>TBC1 domain family member 31</i>
19921096	19962709	ENSBTAG00000013227	<i>SNAI2</i>	19941415	19945016	<i>snail family transcriptional repressor 2</i>
27128043	27166761	ENSBTAG00000026283	<i>ASPH</i>	27016060	27203495	<i>aspartate beta-hydroxylase</i>
33899470	33954905	ENSBTAG00000020312	<i>NCOA2</i>	33837784	34119578	<i>nuclear receptor coactivator 2</i>
35511046	35591127	ENSBTAG00000048963	<i>A0A3Q1MM58_BOVIN</i>	35573943	35575232	-
42822301	42853303	ENSBTAG00000020099	<i>STMN2</i>	42811857	42870107	<i>stathmin 2</i>
51200087	51249961	ENSBTAG00000038281	<i>CSMD3</i>	50969629	52273953	<i>CUB and Sushi multiple domains 3</i>
67307041	67337501	ENSBTAG00000011908	<i>CPQ</i>	66989157	67551536	<i>carboxypeptidase Q</i>
8228215	8259410	ENSBTAG00000007823	<i>TG</i>	8217490	8453614	<i>thyroglobulin</i>
8228215	8259410	ENSBTAG00000007828	<i>SLA</i>	8250306	8315788	<i>Src like adaptor</i>

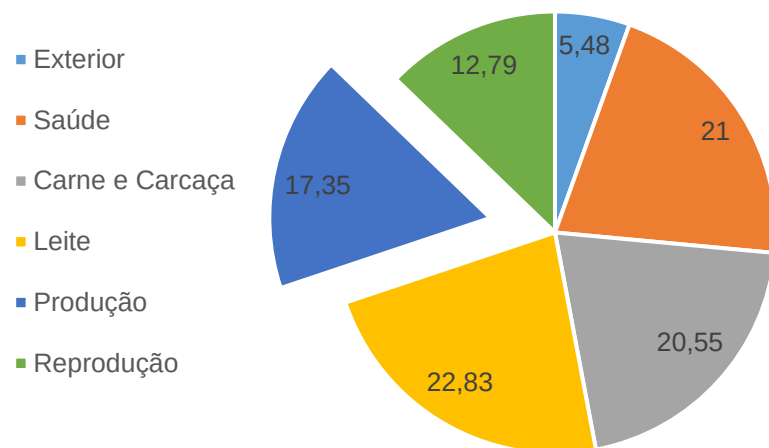
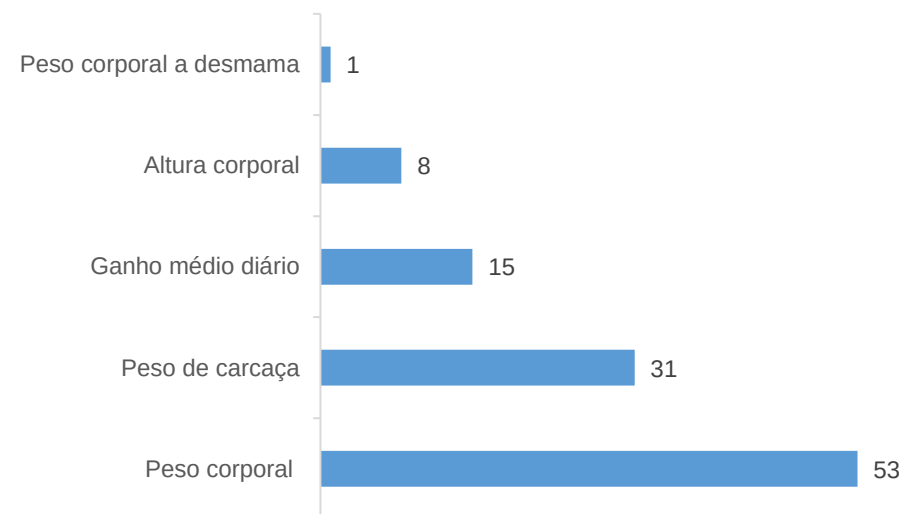
A**B**

Figura Suplementar 1. QTLs presentes nas 20 janelas adjacentes para peso ao nascimento. **A)** porcentagem das classes de QTLs. **B)** QTLs associados com o fenótipo.

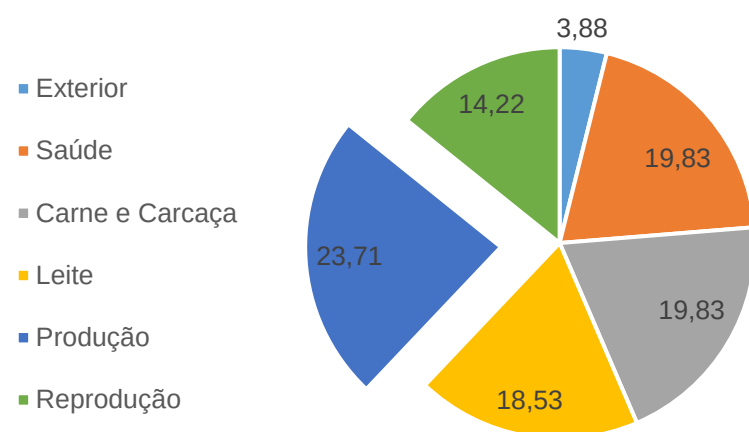
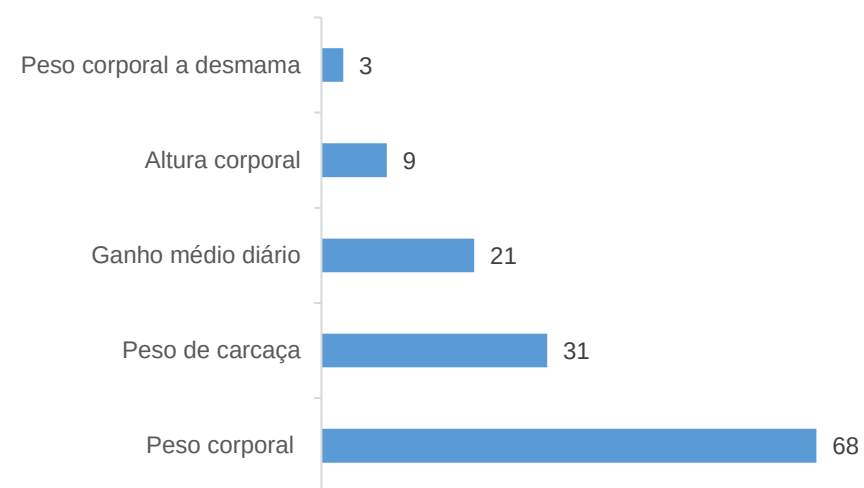
A**B**

Figura Suplementar 2. QTLs presentes nas 20 janelas adjacentes para peso. **A)** porcentagem das classes de QTLs **B)** QTLs associados com o fenótipo.

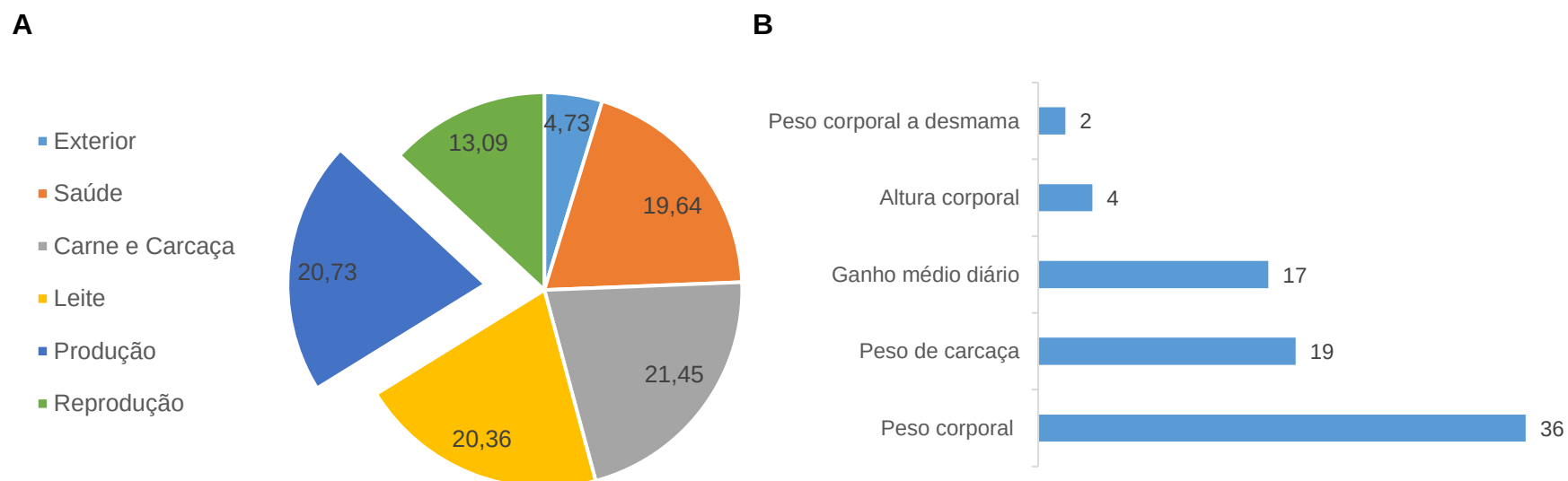


Figura Suplementar 3. QTLs presentes nas 20 janelas adjacentes para peso ao sobreano. **A)** porcentagem das classes de QTLs **B)** QTLs associados com o fenótipo.

Figura Suplementar 4. Matriz de similaridade entre as GRM das classes funcionais no cromossomo 14 baseada na distância Euclidiana.

	3UTR	5UTR	CDS	CpG	Down	Exon	Intron	lncRNA	Missense	Non	Splice	Syn	TFBS	Up
3UTR	0,00	1118,17	903,83	1048,09	719,61	641,14	915,80	1340,21	995,04	2107,86	1121,51	927,65	1528,95	798,49
5UTR	1118,17	0,00	885,35	1034,76	933,60	800,92	1124,61	1545,87	983,53	2158,21	1143,81	908,98	1370,78	872,33
CDS	903,83	885,35	0,00	773,19	637,83	345,08	863,31	1396,55	380,41	2037,23	836,29	294,16	1209,96	603,17
CpG	1048,09	1034,76	773,19	0,00	718,15	711,08	947,62	1488,07	890,75	2012,30	1071,09	845,17	1422,70	707,19
Down	719,61	933,60	637,83	718,15	0,00	452,46	602,82	1213,13	747,74	1904,76	911,77	706,64	1455,36	399,71
Exon	641,14	800,92	345,08	711,08	452,46	0,00	738,33	1281,44	530,48	1924,94	833,70	452,30	1253,20	450,73
Intron	915,80	1124,61	863,31	947,62	602,82	738,33	0,00	1197,55	945,41	1998,06	995,80	918,88	1733,38	633,11
lncRNA	1340,21	1545,87	1396,55	1488,07	1213,13	1281,44	1197,55	0,00	1456,39	2185,98	1478,58	1416,34	1873,16	1208,61
Missense	995,04	983,53	380,41	890,75	747,74	530,48	945,41	1456,39	0,00	2088,43	920,34	611,35	1279,72	718,17
Non	2107,86	2158,21	2037,23	2012,30	1904,76	1924,94	1998,06	2185,98	2088,43	0,00	2100,01	2057,92	2284,89	1881,69
Splice	1121,51	1143,81	836,29	1071,09	911,77	833,70	995,80	1478,58	920,34	2100,01	0,00	859,87	1521,55	911,25
Syn	927,65	908,98	294,16	845,17	706,64	452,30	918,88	1416,34	611,35	2057,92	859,87	0,00	1200,92	673,23
TFBS	1528,95	1370,78	1209,96	1422,70	1455,36	1253,20	1733,38	1873,16	1279,72	2284,89	1521,55	1200,92	0,00	1331,02
Up	798,49	872,33	603,17	707,19	399,71	450,73	633,11	1208,61	718,17	1881,69	911,25	673,23	1331,02	0,00