

UNIVERSIDADE ESTADUAL PAULISTA – UNESP
CÂMPUS DE JABOTICABAL

ESTUDO GENÔMICO DE CARACTERÍSTICAS DE DESEMPENHO
REPRODUTIVO EM BOVINOS DA RAÇA NELORE

Ana Paula Sbardella
Médica Veterinária

2020

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

ESTUDO GENÔMICO DE CARACTERÍSTICAS DE DESEMPENHO

REPRODUTIVO EM BOVINOS DA RAÇA NELORE

Discente: Ana Paula Sbardella

Orientador: Dr. Danísio Prado Munari

Coorientadoras: Dra. Priscila Arrigucci Bernardes

Dra. Nedenia Bonvino Stafuzza

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

2020

Sbardella, Ana Paula

Estudo genômico de características de desempenho reprodutivo em bovinos da raça Nelore / Ana Paula Sbardella. -- Jaboticabal, 2020
78 p. : il., tabs.

S276e

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

Orientador: Danísio Prado Munari

1. Bos taurus indicus. 2. Características reprodutivas. 3. Associação ampla do genoma. 4. Seleção genômica. 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.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: ESTUDO GENÔMICO DE CARACTERÍSTICAS DE DESEMPENHO REPRODUTIVO EM BOVINOS DA RAÇA NELORE

AUTORA: ANA PAULA SBARDELLA

ORIENTADOR: DANISIO PRADO MUNARI

COORIENTADORA: NEDENIA BONVINO STAFUZZA

COORIENTADORA: PRISCILA ARRIGUCCI BERNARDES

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

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

Pesquisador Dr. ROBERTO CARVALHEIRO
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal

Dr. RODRIGO PELICIONI SAVEGNAGO
Instituto de Zootecnia / Sertãozinho/SP

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

Pesquisadora Dra. LAIZA HELENA DE SOUZA IUNG
CRV Lagoa / Sertãozinho/SP

Jaboticabal, 19 de junho de 2020

DADOS CURRICULARES DO AUTOR

ANA PAULA SBARDELLA – nascida em 16 de janeiro de 1989, na cidade de Concórdia, Santa Catarina, filha de Vilmar Antonio Sbardella e Elisete Maria Dal’Pizzol Sbardella. Iniciou o curso de Medicina Veterinária em março de 2010 na Universidade Federal do Paraná, Setor Palotina, e obteve o título de bacharel em Medicina Veterinária em junho de 2014. Em agosto de 2014, ingressou no curso de mestrado pelo Programa de Pós-graduação em Genética e Melhoramento Animal na Faculdade de Ciências Agrárias e Veterinárias, UNESP, Campus Jaboticabal sob orientação do Prof. Dr. Danísio Prado Munari e coorientação do Dr. Marcos Vinícius Gualberto Barbosa da Silva, obteve o título de mestra em julho de 2016. Foi bolsista CAPES pelo período de 01 de agosto de 2014 a 29 de julho de 2016. Em agosto de 2016, ingressou no curso de doutorado pelo Programa de Pós-Graduação em Genética e Melhoramento Animal na Faculdade de Ciências Agrárias e Veterinárias, UNESP, Campus de Jaboticabal sob orientação do Prof. Dr. Danísio Prado Munari e coorientação da Dra. Priscila Arrigucci Bernardes e Dra. Nedenia Bonvino Stafuzza. Foi bolsista CAPES de agosto de 2016 a janeiro de 2019. Em abril de 2018, ingressou no curso de Master of Business Administration - MBA em Agronegócio na Universidade de São Paulo - ESALQ/USP, curso concluído em dezembro de 2019.

DEDICO

Aos meus pais, Elisete e Vilmar por serem simplesmente tudo em minha vida! O mérito é de vocês. Amo vocês!

Ao meu irmão Maicon, por ser minha inspiração como ser humano e como profissional. Te amo!

AGRADECIMENTOS

A Deus...

Aos meus pais Elisete e Vilmar, por todo esforço, amor e dedicação com nossa família; por serem meus exemplos de coragem, honestidade e simplicidade e por todos os valores e princípios passados.

Ao meu irmão Maicon, por todos os conselhos e cuidados comigo. Você é meu orgulho. Tua serenidade e tranquilidade me acalmam e me fazem ver o mundo na sua melhor versão.

Ao meu companheiro de vida Lucas, pela paciência, compreensão, parceria, lealdade e amor.

Ao meu afilhado Dudu, por permitir que eu volte no tempo e me veja criança.

Ao meu Orientador prof. Dr. Danisio Prado Munari, por todos os ensinamentos ao longo desses 6 anos de pós-graduação, por compreender minha inquietação e sempre ter uma palavra de apoio.

A minha coorientadora prof. Dr. Priscila Arrigucci Bernardes, por aceitar a coorientação mesmo sabendo das dificuldades que seria fazer isso a distância. Por toda paciência, atenção e ensinamentos, indiferente do dia e do horário (obs: prometo parar de mandar mensagens no final de semana).

Aos membros da banca do Exame Geral de Qualificação e de Defesa, Dr. Roberto Carneiro, Dr. Rodrigo Pelicioni Savegango, Dr. Henrique Nunes de Oliveira e Dra. Laiza Helena de Souza Lung, por todas as sugestões que permitiram o enriquecimento desse trabalho.

Aos amigos e colegas do Departamento de Ciências Exatas e do programa de Genética e Melhoramento Animal, pelo conhecimento compartilhado e pelos cafés intermináveis, que davam a força necessária para continuar.

Aos amigos e colegas da Terra Desenvolvimento Agropecuário, pela oportunidade e confiança.

A todos os demais familiares e amigos que contribuíram de alguma forma durante essa trajetória.

A UNESP – Jaboticabal, pela oportunidade de me desenvolver como profissional e como pessoa.

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

SUMÁRIO

RESUMO.....	iii
ABSTRACT	iv
CAPÍTULO 1 – Considerações Gerais	1
1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA	2
2.1 Aspectos Gerais	2
2.2 Características Reprodutivas	3
2.3 Estudos Genômicos	5
2.3.1 Associação Ampla do Genoma (GWAS).....	6
2.3.2 Seleção Genômica Ampla (GS)	8
3 REFERÊNCIAS BIBLIOGRÁFICAS	10
CAPÍTULO 2 - Genome-wide association study provides insights into important genes for reproductive traits in Nelore cattle	17
ABSTRACT	17
1 INTRODUCTION.....	17
2 MATERIAL AND METHODS	18
2.1 Data.....	18
2.2 Quality Control.....	19
2.2.1 Phenotype Data	19
2.2.2 Genotype Data	20
2.3 Genome-wide association analysis	20
2.4 Search for Associated Genes.....	22
3 RESULTS AND DISCUSSION	23
4 CONCLUSIONS.....	29
5 REFERENCES.....	30
CAPÍTULO 3 - Predição genômica de características reprodutivas em bovinos da raça Nelore.....	35
RESUMO.....	35
1 INTRODUÇÃO	35
2 MATERIAL E MÉTODOS.....	37
2.1 Conjunto de dados	37
2.2 Controle de Qualidade	38
2.3 Seleção Genômica Ampla (GS)	39
2.3.1 Análise de Agrupamento hierárquica e não-hierárquica	41
3 RESULTADOS E DISCUSSÃO.....	41
4 CONCLUSÃO.....	48

5 REFERÊNCIAS	48
CAPÍTULO 4 - CONSIDERAÇÕES GERAIS.....	52
APÊNDICES	53
Supplementary material A	57
Supplementary material B	62
Supplementary material C	63
Supplementary material D	64
Supplementary material E	65
Supplementary material F.....	66
Apêndice G.....	67

ESTUDO GENÔMICO DE CARACTERÍSTICAS DE DESEMPENHO REPRODUTIVO EM BOVINOS DA RAÇA NELORE

RESUMO – A informação genômica vem sendo aplicada como importante ferramenta em programas de melhoramento animal. Desta maneira, pode contribuir para melhorar o desempenho reprodutivo dos animais visando atender à crescente demanda por carne bovina. Este trabalho tem como objetivos: (1) Identificar a associação entre polimorfismos de nucleotídeo único (SNP) com as características de circunferência escrotal aos 365 dias de idade (CE365) e aos 450 dias de idade (CE450), período de gestação (PG), como característica do bezerro, idade ao primeiro parto (IPP), produtividade acumulada (PAC), parto precoce até 30 meses de idade (P30) e habilidade de permanência da vaca no rebanho (“Stayability” - STAY) a fim de detectar genes candidatos e vias biológicas associadas com características reprodutivas em bovinos de corte da raça Nelore; (2) Comparar a acurácia de predição de características reprodutivas usando os métodos Melhor preditor linear não-viesado (BLUP) e Melhor preditor linear não viesado genômico de único passo (ssGBLUP) e classificar cada animal segundo seu valor genético (EBV) e valor genético genômico (GEBV), estimados para as características estudadas, visando obter informações mais acuradas para a seleção de animais jovens; (3) Comparar grupos de animais para as características estudadas, resultantes de análises de agrupamento com as estimativas de EBV e GEBV, a fim de identificar aqueles que apresentam melhor potencial genético para o desempenho reprodutivo. Para isso, foram utilizados dados fenotípicos, de pedigree e genotípicos disponibilizados pela Associação Nacional de Criadores e Pesquisadores (ANCP). Os dados genotípicos são oriundos de 8.652 animais, genotipados ou imputados para painéis de marcadores SNP em alta densidade (Illumina Bovine HD BeadChip). Dados fenotípicos de 50.331 e 34.848 animais foram utilizados para estudo de associação genômica ampla (GWAS) e seleção genômica (GS), respectivamente. O GWAS em passo único ponderado (WssGWAS) foi utilizado para estimação dos efeitos dos SNP sobre as características estudadas. Para comparar a habilidade de predição foram utilizados os métodos de avaliação BLUP e ssGBLUP. Em seguida, as análises de agrupamento hierárquica e não hierárquica foram realizadas para agrupar os animais com base no EBV ou GEBV. Os resultados do presente estudo revelaram genes com importantes funções para características reprodutivas, como os influenciadores na fertilidade e precocidade. Os genes CAMK1D, ACOXL, MRPL33, PLCB1 e JAG1, entre outros, aparecem associados com mais de uma característica. Os resultados de habilidade de predição resultantes do método ssGBLUP (0,07 - 0,46), em sua maioria foram maiores do que os resultados obtidos pelo método BLUP (0,06 - 0,41). Ademais, uma diferença nos valores genéticos dos animais quando a informação genômica é incorporada nas análises foi observada, o que pode contribuir no direcionamento da seleção de animais jovens segundo o GEBV. O GWAS evidenciou que a identificação de regiões associadas com as características estudadas pode contribuir com futuros estudos de validação biológica e os genes identificados serem indicados como candidatos a seleção. Ainda, os resultados de GS indicaram que a inclusão da informação genômica melhora a habilidade de predição de características reprodutivas e que a seleção de animais superiores para as características de interesse pode ser diferente quando a informação genômica é considerada.

Palavras-chave: *Bos taurus indicus*, características reprodutivas, associação ampla do genoma, seleção genômica

GENOMIC STUDY OF REPRODUCTIVE PERFORMANCE TRAITS IN NELORE CATTLE

ABSTRACT - Genomic information has been applied as important tool in animal breeding programs. In this way, it can contribute to improving the reproductive performance of the animals in order to meet the growing demand for beef. The aim of this study was: (1) to identify the association between single nucleotide polymorphisms (SNP) with scrotal circumference at 365 days (SC365) and at 450 days (SC450), gestation length (GL) as a calf trait, age at first calving (AFC), accumulated productivity (ACP), heifer early calving until 30 months (HC30) and Stayability (STAY) traits, in order to detect candidate genes and biological pathways associated with reproductive traits in Nelore beef cattle; (2) To compare predictive accuracy of reproductive traits using the Best Linear Unbiased Prediction (BLUP) and the Single-step Genomic BLUP (ssGBLUP) methods and to order each animal according to their breeding value (EBV) and genomic breeding value (GEBV), estimated for the traits studied aiming to obtain more accurate information for young animals selection; (3) To compare clusters of the animals for the evaluated traits, resulting from cluster analyzes with EBV and GEBV, in order to identify those who have the best genetic potential for reproductive performance. The data set was composed of pedigree, phenotypes and genotypes of Nelore beef cattle, from the "Associação Nacional de Criadores e Pesquisadores" (ANCP). Genotyped data in HD or imputed to high density SNP marker panels are from 8,652 animals (males and females). The phenotype data from 50.331 e 34.848 animals were used for genome-wide association study (GWAS) and genome selection (GS), respectively. The weighted genomic single step (WssGWAS) was used to estimate the effects of SNP on the studied traits. To estimate the ability prediction, the BLUP and ssGBLUP evaluation methods were used. Then, the hierarchical and non-hierarchical cluster analyzes were performed to group the animals based on EBV or GEBV. The results obtained in this study revealed genes with important functions for reproductive traits, such as those influencing fertility and precocity. The CAMK1D, ACOXL, MRPL33, PLCB1 and JAG1 genes, among others, appear to be associated with more than one trait. The prediction ability results from ssGBLUP method (0.07 to 0.46) was higher than BLUP method (0.06 to 0.41). In addition, a difference in the genetic values of animals when genomic information is incorporated into the analyzes was observed, which can contribute to directing the selection of young animals according to the GEBV. The identification of regions associated with the studied traits can contribute to future studies of biological validation and the identified genes are indicated as candidates for selection. Also, the inclusion of genomic information improves the ability to predict reproductive traits. These results indicate that the selection of superior animals for the traits of interest may be different when the genomic information is considered.

Keywords: *Bos taurus indicus*, reproduction traits, genome-wide association, genomic selection

CAPÍTULO 1 – Considerações Gerais

1 INTRODUÇÃO

A crescente demanda de proteína animal na alimentação humana, aumentou a demanda na produção de carne bovina. Para atender essa maior demanda, produtores e pesquisadores vêm trabalhando ao longo do tempo, utilizando-se dos benefícios da genética e nutrição animal, na busca por um desempenho reprodutivo que seja compatível com o crescimento da pecuária bovina e que resulte em maior produtividade (FAO, 2020; Nowacka-Woszuik, 2020)

O rebanho brasileiro é composto em sua maioria por raças zebuínas, das quais se destaca a raça Nelore, que é bastante explorada comercialmente (Alves, 2019). A raça Nelore tem uma importante participação na economia brasileira por sua adaptação ao sistema de produção do país, no entanto quando comparada a raças europeias, o Nelore é menos precoce (Sartori et al., 2010). Desta maneira, as características reprodutivas são de elevada importância econômica para pecuária do país, pois tem relações com eficiência produtiva e reprodutiva do rebanho (Brumatti et al., 2011). Assim, a utilização de características reprodutivas em programas de seleção vem sendo aplicada objetivando a redução da idade ao primeiro parto e o aumento do número de progênes deixadas pela vaca ao longo da vida produtiva (Goddard e Hayes, 2007). No entanto, a incorporação de algumas características reprodutivas como critérios de seleção em programas de melhoramento genético ainda ocorre lentamente, tendo como principais motivos o fato de que essas características são de baixa herdabilidade, expressas em idades mais tardias e limitadas pelo sexo (Dias et al., 2004; Cammack et al., 2009).

A utilização e aplicação mais eficiente de métodos de melhoramento genético, tornam as perspectivas do avanço na produção de bovinos de corte favoráveis. Com a chegada das 'ÔMICAS', termo utilizado para designar genômica, transcriptômica, proteômica, metabolômica, entre outras, tem-se otimizado as predições genéticas e viabilizado a inclusão de novas características em programas de melhoramento com diferenças esperadas na Progênie (DEP's) mais confiáveis. A incorporação de informações genômicas em programas de melhoramento permite predizer o mérito genético com maior confiabilidade e a redução do intervalo de geração (Goddard e Hayes, 2002; Boichard et al., 2016).

Com a finalidade de contribuir para a melhoria dos processos de seleção de características reprodutivas em bovinos de corte, os objetivos deste estudo foram: (1) Identificar a associação entre polimorfismos de nucleotídeo único (SNP) com as características de circunferência escrotal aos 365 dias de idade (CE365) e aos 450 dias de idade (CE450), período de gestação (PG), como característica do bezerro, idade ao primeiro parto (IPP), produtividade acumulada (PAC), parto precoce até 30 meses de idade (P30) e habilidade de permanência da vaca no rebanho (“Stayability” - STAY), a fim de detectar genes candidatos e vias biológicas associadas com características reprodutivas em bovinos de corte da raça Nelore; (2) Comparar a acurácia de predição de características reprodutivas usando os métodos Melhor preditor linear não viesado (BLUP) e Melhor preditor linear não viesado genômico de único passo (ssGBLUP) e classificar cada animal segundo seu valor genético (EBV) e valor genético genômico (GEBV), estimados para as características estudadas, visando obter informações mais acuradas para a seleção de animais jovens; (3) Comparar grupos de animais para as características estudadas, resultantes de análises de agrupamento com as estimativas de EBV e GEBV, a fim de identificar aqueles que apresentam melhor potencial genético para o desempenho reprodutivo.

2 REVISÃO DE LITERATURA

2.1 Aspectos Gerais

A bovinocultura de corte é uma atividade econômica de importância para o agronegócio brasileiro, sendo responsável por expressiva parcela dos lucros gerados pela indústria da carne no Brasil. No ano de 2019, a produção mundial de carne bovina foi de aproximadamente 61 milhões de toneladas (USDA, 2020). Dessas, dez milhões de toneladas foram produzidas no Brasil, que é o segundo país com maior produção de carne bovina no mundo, e o maior exportador mundial de carne bovina (USDA, 2020). Esses dados associados ao baixo custo de produção consequente do sistema de produção a pasto, que é predominante no país e as tecnologias disponíveis evidenciam o potencial e competitividade da pecuária de corte brasileira.

O melhoramento genético de bovinos Nelore no Brasil, realizado pelos programas de melhoramento genético, tem função importante para aumentar a produtividade de carne bovina e como suporte para que esta cadeia atinja seu potencial. Basicamente, os programas de melhoramento fornecem estimativas de

diferença esperada na progênie (DEP) reunidas em índices de seleção de acordo com o interesse econômico. Os programas também disponibilizam assistência técnica para a realização de acasalamentos dirigidos que visam o aumento da produção de carne.

Os métodos de seleção tradicionalmente recomendados pelos programas de melhoramento genético do Brasil permitem o aumento do ganho genético para muitas características de importância econômica. No entanto, estratégias adicionais devem ser implementadas para que a aplicação de métodos mais eficientes que permitam a mensuração de novas características e a utilização de animais jovens geneticamente superiores, possam ser implementados para disseminar ganho genético também para rebanhos comerciais (Carvalho, 2014).

2.2 Características Reprodutivas

As características reprodutivas são economicamente relevantes, com grande potencial na melhoria da rentabilidade do sistema de produção de gado de corte (Brumatti et al., 2011). A precocidade sexual é especialmente importante para raças *Bos taurus indicus* que são comumente menos precoces que as raças *Bos taurus taurus* (Sartori et al., 2010). Dentre as características implementadas em programas de melhoramento genético, circunferência escrotal (CE) se destaca para machos. Em fêmeas, idade ao primeiro parto (IPP), parto precoce até 30 meses de idade (P30) e habilidade de permanência da vaca no rebanho ("Stayability" - STAY) são bastante utilizadas e, em menor escala, período de gestação (PG) e produtividade acumulada (PAC).

A CE é a que compõe os índices de seleção há mais tempo, por sua facilidade de mensuração e por apresentar associação genética favorável com características de fertilidade e precocidade sexual, o que justifica sua contínua exploração e inclusão em processos de seleção (Bolognini et al., 2007). Essa característica é expressa em centímetros e apresenta estimativas de herdabilidade variando entre 0,31 e 0,65, conforme a idade em que é mensurada (Grossi et al., 2009; Moreira et al., 2015; Terakado et al., 2015; Buzanskas et al., 2017).

As características reprodutivas de fêmeas têm adquirido maior importância econômica nos programas de melhoramento genético. Dentre as características associadas à precocidade sexual, IPP é a mais utilizada, seguida de P30. A IPP é influenciada pelo peso a puberdade e por isso é afetada, além do potencial genético, que também interfere na taxa de crescimento da fêmea, pela variação de quantidade

e qualidade de forragens durante o ano e suplementação disponibilizada para novilha (Bhatti et al., 2007). A IPP indica a maturidade sexual das novilhas, interferindo no intervalo de geração e conseqüente aumento do ganho genético anual. Essa característica é facilmente mensurada e apresenta estimativas de herdabilidade variando de 0,08 a 0,3 (Dias et al 2004; Buzanskas et al., 2017; Schimidt et al., 2019; Costa et al., 2020).

A característica P30 é medida pela ocorrência de fêmeas, desafiadas precocemente, que pariram um bezerro vivo até 30 meses de idade. A P30 possui estimativas de herdabilidade que variam de 0,37 a 0,53 (Vozzi et al., 2008; Kluska et al., 2018). Por contemplar a prenhez precoce e parição de bezerro vivo, a P30 tem importância econômica na pecuária de corte e por isso é indicada como critério de seleção para precocidade sexual (Vozzi et al., 2008; Lôbo et al., 2016).

A característica STAY expressa a habilidade da vaca em permanecer até os 76 meses de idade no rebanho, produzindo no mínimo três bezerros. Isso a torna uma característica economicamente importante, o que justifica sua inclusão nos índices de seleção (Formigoni et al., 2005; Bittencourt et al., 2006), pois possibilita a seleção de touros que produzam filhas que permaneçam no rebanho produzindo progênes por mais tempo (Silva et al., 2003). A STAY possui estimativas de herdabilidade que variam de 0,09 a 0,25 (Van Melis et al., 2007; Kluska et al., 2018; Costa et al., 2020).

O PG pode ser avaliado como característica da mãe ou do bezerro. Essa característica está relacionada ao período reprodutivo da vaca, embora não seja indicador de precocidade (Moreira et al., 2015). O PG é determinado principalmente pela idade da vaca, sendo que vacas mais jovens apresentam PG menores (Crews, 2006; Jamrozik e Miller et al., 2014), sendo também influenciado pela raça e sexo do bezerro. As estimativas de herdabilidade para PG variam entre 0,27 e 0,49 (Nogalski e Piwczyński, 2012; Chud et al., 2014; Moreira et al., 2015).

A PAC indica a produtividade da vaca em kg de bezerros desmamados por vaca por ano, expressando assim a capacidade da vaca iniciar precocemente a produção de bezerros. Essa característica é obtida por um índice que depende da IPP, da permanência da vaca no rebanho e da sua capacidade de desmamar bezerros mais pesados (Lôbo et al., 2000). A PAC é associada com características de precocidade e fertilidade e possui estimativas de herdabilidade que variam de 0,11 a 0,25 (Schwengber et al., 2001; Faria et al., 2007; Grossi et al., 2008; Chud et al., 2014).

Algumas características indicativas de precocidade sexual, tais como CE, IPP e ocorrência de parto precoce, são consideradas nos programas de melhoramento de bovinos zebuínos no Brasil (Costa et al., 2019). As DEP são estimadas para essas características, mas a maioria dos programas de melhoramento incluem apenas CE em índices de seleção, pois é de fácil mensuração e possui herdabilidade de magnitude média a alta (Van Melis et al., 2010; Laureano et al., 2011). Quando se observa as correlações genéticas entre CE e as demais características reprodutivas, as mesmas são de baixa magnitude (Boligon et al., 2007), evidenciando que a inclusão exclusiva de CE nos índices de seleção pode não ser uma estratégia eficiente para melhorar o desempenho reprodutivo (Costa et al., 2019). Isso também indica a necessidade de inclusão das demais características reprodutivas em índices de seleção para atingir melhor desempenho reprodutivo.

2.3 Estudos Genômicos

O genoma bovino tem aproximadamente 3.000 MB, distribuídos em 29 pares de cromossomos autossomos acrocêntricos e dois cromossomos sexuais (NCBI, 2016). A primeira montagem e anotação do genoma bovino (UMD3.1) foi realizada em dezembro de 2009 e segundo o banco de dados do Ensembl (Cunningham et al., 2018) é composta por 43.984 genes transcritos.

O sequenciamento da molécula de ácido desoxirribonucleico (DNA) permite identificar as bases nitrogenadas pelas quais o DNA é composto. Mutações de transição e transversão que ocorrem nessas bases são sequenciadas e identificadas como marcadores moleculares do tipo polimorfismos de nucleotídeo único (SNP - "Single Nucleotide Polymorphism"), que são variações no genoma que podem estar associadas a diferentes fenótipos nos indivíduos de uma população. As informações provenientes dos marcadores tipo SNP são disponibilizadas em painéis de diversas densidades.

A disponibilização de painéis de SNP viabilizou estudos como os de associação, predição (O'Brien et al., 2014) e estudos populacionais (Ferencakovic et al., 2011), os quais vem sendo desenvolvidos em diferentes espécies. Os marcadores associados a características de interesse econômico podem auxiliar no processo de melhoramento genético, quando incluída sua informação em modelos para estimação de valores genéticos genômicos para essas características.

Os benefícios da utilização da informação gerada a partir de marcadores moleculares no processo de melhoramento genético ocorrem principalmente em características de difícil mensuração, de expressão tardia ou ainda, características que requerem o abate dos animais para mensuração. As vantagens incluem ganhos na acurácia de predição de valor genético genômico (GEBV); aumento na proporção da variância genética que é explicada pelos marcadores; estimação do efeito dos *loci* de características quantitativas (QTL – “Quantitative Trait Loci”) com maior exatidão e redução no intervalo de geração (Goddard e Hayes, 2002; Boichard et al., 2016). Isso é possível, pois os animais podem ser genotipados para obtenção de suas informações moleculares, quando nascem ou até mesmo durante sua gestação. Em adição, quando essas informações são utilizadas juntamente com tecnologias reprodutivas, é possível reduzir a idade ao primeiro parto e aumentar o número de progênies ao longo da vida da vaca (Goddard e Hayes, 2007).

2.3.1 Associação Ampla do Genoma (GWAS)

Estudos de associação ampla do genoma (GWAS - “Genome-Wide Association Study”) são utilizados principalmente para identificar associações entre regiões cromossômicas e características complexas, ou seja, controladas por múltiplos genes. Os marcadores moleculares do tipo SNP são os mais utilizados para identificar essas regiões ao longo do genoma e então, associá-las as características de interesse. O GWAS testa a associação entre cada SNP com a característica, sendo que associações significativas ocorrem quando o marcador molecular está em desequilíbrio de ligação (DL) com a mutação causal que afeta a característica (Hayes e Goddard, 2010).

Muitos modelos têm sido propostos para realização das análises de GWAS. Inicialmente, o modelo utilizando regressão de marcadores pontuais foi proposto, no entanto, apenas uma pequena fração da variação genética pode ser explicada utilizando SNP significativos (Hayes, 2013; Wu et al., 2014). Então, metodologias para modelos mais complexos para GWAS foram propostos. O GWAS baseado em blocos de haplótipos em janelas no genoma associados aos fenótipos pode ser mais adequado, uma vez que os haplótipos podem estar em maior DL com os QTL quando comparado com a metodologia de regressão pontual utilizando painéis de menor densidade (Hayes, 2013). No entanto ainda existem fatores limitantes na sua utilização, como a estrutura e a fase dos blocos de haplótipos que raramente são

observados nos dados de genotipagem; diferentes métodos para construção dos blocos geram diferentes resultados e o aumento de grau de liberdades dentro do bloco (Niu, 2004; Shim et al., 2009; Wu et al., 2014).

Modelos que ajustam todos os SNP de forma simultânea e os consideram como efeitos aleatórios (Meuwissen et al., 2001), diminuem a ocorrência de estimativas superestimadas e a ocorrência de estimação de efeitos viesados, que são algumas desvantagens dos modelos citados anteriormente (Goddard et al., 2009; Goddard e Hayes 2012; Lee et al., 2013). Diferentes modelos de estimação dos efeitos dos SNP sobre as características de importância econômica foram sugeridos por Meuwissen et al. (2001), incluindo os modelos bayesianos (BayesA e BayesB) e BLUP.

O método Melhor preditor linear não viesado genômico de único passo (ssGBLUP) proposto por Misztal et al. (2009) e Christensen e Lund (2010) é uma adaptação do método Melhor preditor linear não viesado (BLUP) e permite a integração simultânea das informações de fenótipo, genótipo e de pedigree em análises de predição e associação genômica. A metodologia permite a incorporação de animais que tenham apenas informação de fenótipo, pois o método não é restrito a animais genotipados. Nesse procedimento, as informações de parentesco contidas nos registros de pedigree são acrescidas das informações derivadas dos marcadores moleculares e os GEBV são utilizados para estimar os efeitos e pesos (para ssGBLUP ponderado) dos SNP (Zhang et al., 2010; Sun et al., 2011; Wang et al., 2012).

Em bovinos de corte, estudos de GWAS foram conduzidos para diversas raças (Bolormaa et al., 2011; Kim et al., 2011; Nishimura et al., 2012), associando a variação genética principalmente com características de carcaça e qualidade de carne. Além disso, outros estudos com características reprodutivas também foram realizados. Regiões cromossômicas associadas com CE (Irano et al., 2016) e com IPP (Mota et al., 2017) foram identificadas permitindo a indicação de genes candidatos que apresentam funções associadas ao desempenho reprodutivo de bovinos da raça Nelore. Para contribuir com a seleção de características reprodutivas de fêmeas em bovinos de corte Brahman, novos SNP identificados podem ser incorporados em futuras versões de painéis de SNP (Fortes et al., 2018).

Em animais da raça Nelore, Melo et al. (2017) identificaram associação entre reconcepção de novilhas e ocorrência de parto até os 53 meses de idade com as vias biológicas de metabolismo de lipídios e processos imunes, respectivamente. Ainda, em animais Nelore, regiões associadas à precocidade sexual foram identificadas

(Melo et al., 2018) e regiões genômicas que eram conhecidas por afetar características de precocidade sexual em gado Brahman australiano foram validadas em população de animais da raça Nelore (Melo et al., 2019). Nascimento et al. (2018) identificaram regiões cromossômicas associadas a IPP, as quais contêm genes e QTL relacionados ao desempenho reprodutivo.

2.3.2 Seleção Genômica Ampla (GS)

Os programas de seleção têm a finalidade de acelerar a taxa de mudança genética por unidade de tempo, direcionada por um objetivo de interesse econômico. Muitas tecnologias, tais como a dos marcadores moleculares e as biotecnologias de reprodução, que possam atuar no aumento da variação genética capturada ou na diminuição do intervalo de geração, aumentam a velocidade do ganho genético (Van Eenennaam et al., 2014).

A incorporação da informação de marcadores SNP em modelos estatísticos que são tradicionalmente utilizados para predição do mérito genético foi proposta inicialmente por Nejati-Javaremi et al. (1997) e Meuwissen et al. (2001). Os autores observaram que a inclusão desta informação, associada aos registros fenotípicos, poderia aumentar a acurácia de predição do mérito genético dos indivíduos. Nessa época, as plataformas de genotipagem ainda não haviam sido desenvolvidas, mas com a disponibilidade das mesmas, a predição genômica tornou-se possível e vem sendo aplicada com a denominação de seleção genômica (GS – “Genomic Selection”).

A GS utiliza indivíduos genotipados com painéis de média (50k) a alta (800k) densidade de marcadores, ou até mesmo com painéis de baixa densidade imputados para média ou alta densidade, permitindo que a GS ocorra em maior escala em termos de volume de animais, para obter os resultados de predição. Com base em grande número de indivíduos em uma população de referência, uma equação de predição é derivada com o objetivo de prever os valores genéticos genômicos (GEBV) como a soma de efeitos de todos os marcadores SNP. Nessa estratégia de seleção, o efeito do marcador não precisa exceder um limiar de significância para ser incluído na equação de predição, por isso é capaz de capturar toda a variância genética associada aos marcadores (Van Eenennaam et al., 2014).

Em bovinos de corte, a utilização da GS aumenta a precisão das DEP geradas para animais jovens e permite a identificação de animais superiores em idades

precoces. Essas informações associadas a tecnologias reprodutivas permitem intensificação no uso de animais jovens superiores geneticamente aumentando o ganho genético (Carvalho, 2014). A GS na indústria de bovinos de corte está sendo utilizada de maneira lenta, quando comparada a indústria leiteira (Hayes et al., 2013).

Métodos de seleção genômica determinam equações de predição para estimar os efeitos dos SNP com base em uma população de treinamento e para que essa equação seja validada, a mesma é aplicada em uma população de animais candidatos a seleção, com informação genotípica (Goddard e Hayes, 2009). Alguns fatores impactam na habilidade preditiva, como a herdabilidade, o número de animais e de SNP, o método utilizado e o relacionamento entre a população de referência e validação (Makowsky, 2011; Neves et al., 2014; Oldenbroek e Waaij, 2015; Silva et al., 2016).

Os métodos bayesianos, dentre eles Bayes A, Bayes B e Lasso, e o método GBLUP, são muito utilizados para predizer o mérito genético, sendo a principal diferença entre eles a distribuição dos efeitos dos SNP assumida (Hayes e Goddard, 2010). Alguns estudos reportaram maior habilidade de predição utilizando os métodos bayesianos quando comparados ao método GBLUP (Jimenez-Montero et al., 2013; Meuwissen et al., 2013; Costa et al., 2019). No entanto, quando todos os indivíduos da população não são genotipados, um dos principais métodos de escolha é o ssGBLUP, descrito por Aguilar et al. (2010). Esse método pode apresentar ganhos na acurácia quando comparado com outros métodos, possivelmente por permitir a combinação das informações de pedigree, fenótipo e marcadores SNP em uma única análise (Legarra et al., 2014).

A viabilidade técnica da aplicação da predição genômica em bovinos da raça Nelore foi demonstrada por Neves et al., (2014). Os autores relataram ainda a necessidade de constante atualização na população de referência para que a predição em animais jovens fosse acurada. Estudo com características de carcaça em bovinos Nelore reportou acurácia de predição de moderada a alta, sendo que diferentes modelos bayesianos foram testados e todos obtiveram resultados similares (Fernandes Júnior et al., 2016).

Acurácias de baixa a moderada magnitude para perfil de ácidos graxos na carne em bovinos Nelore foram reportadas por Chiaia et al. (2018) utilizando o método ssGBLUP, que se mostrou uma alternativa apropriada para obtenção de GEBV confiáveis. Costa et al. (2019) estudaram a acurácia de predição dos modelos GBLUP

e Bayesianos para características reprodutivas de reconcepção de novilhas, IPP e ocorrência de parto precoce em fêmeas da raça Nelore, concluindo que a precisão da predição é mais dependente da própria característica do que do modelo utilizado.

A inclusão da prática da seleção genômica em uma população deve ser sempre monitorada. Além disso, é necessária a constante atualização da população de treinamento com novos animais genotipados, permitindo aumentar a acurácia de predição em animais jovens, uma vez que a acurácia depende da relação entre a população de animais nos grupos de treinamento e validação (Neves et al., 2014).

3 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** 93:743-752.

Alves, AAC (2019) Applying machine learning methods for genomic analysis of reproductive traits in nelore cattle. 114 f. Tese (Doutorado em Genética e Melhoramento animal – Unesp, Jaboticabal.

Bhatti SA, Sarwar M, Khan MS, Hussain SMI (2007) Reducing the age at first calving through nutritional manipulations in dairy buffaloes and cows: a review. **Pakistan Veterinary Journal** 27:42-47.

Bittencourt TCC, Lôbo RB, Bezerra LAF, (2006) Objetivos de seleção para sistemas de produção de gado de corte em pasto: ponderadores econômicos. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia** 58:196-204.

Boichard D, Ducrocq V, Croiseau P, Fritz S, (2016) Genomic selection in domestic animals: Principles, applications and perspectives. **Comptes Rendus Biologies** 339:274-277.

Boligon AA, Rorato PRN, Albuquerque LG (2007) Correlações genéticas entre medidas de perímetro escrotal e características produtivas e reprodutivas de fêmeas da raça Nelore. **Revista Brasileira de Zootecnia** 36:565-571.

Bolorma S, Porto Neto LR, Zhang YD, Bunch RJ, Harrison BE, Goddard ME, Barendse WA, (2011) genome-wide association study of meat and carcass traits in Australian cattle. **Journal of Animal Science** 89:2297-2309.

Brumatti RC, Ferraz JBS, Eler JP, Formigoni IB, (2011) Desenvolvimento de índice de seleção em gado corte sob o enfoque de um modelo bioeconômico. **Archivos Zootecnia** 60:205-213.

Buzanskas ME, Pires PS, Chud TCS, Bernardes PA, Rola LD, Savegnago RP, Lôbo RB, Munari DP (2017) Parameter estimates for reproductive and carcass traits in Nelore beef cattle. **Theriogenology** 09-57.

Cammack KM, Thomas MG, Enns RM (2009) REVIEW: Reproductive Traits and Their Heritabilities in Beef Cattle. **Professional Animal Scientists** 25:517-528.

Carvalho R, (2014) Genomic Selection in Nelore Cattle in Brazil. In: 10th World Congress of Genetics Applied to Livestock Production. **Presentations...** Vancouver.

Chiaia HLJ, Peripolli E, Silva RMO, et al. Genomic prediction ability for beef fatty acid profile in Nelore cattle using different pseudo-phenotypes. **Journal of Applied Genetics** 59:493-501.

Christensen OF, Lund MS, (2010) Genomic prediction when some animals are not genotyped. **Genetics Selection Evolution** 42:2.

Chud TCS, Caetano SL, Buzanskas ME, et al. (2014) Genetic analysis for gestation length, birth weight, weaning weight, and accumulated productivity in Nellore beef cattle. **Livestock Science** 170:16-21.

Costa ED, Ventura HT, Veronezea R, Silva FF, Pereira MA, Lopes PS (2020) Estimated genetic associations among reproductive traits in Nellore cattle using Bayesian analysis. **Animal Reproduction Science** 1-21.

Costa RB, Irano N, Diaza IDPS, Takada L, et al, (2019) Prediction of genomic breeding values for reproductive traits in Nellore heifers. **Theriogenology** 1-27.

Crews JRDH (2006) Age of dam and sex of calf adjustments and genetic parameters for gestation length in Charolais cattle. **Journal Animal of Science** 84:25-31.

Cunningham F, Achuthan P, et al, (2018) Ensembl **Nucleic Acids Research**. doi:10.1093/nar/gky1113

Dias LT, El Faro L, Albuquerque LG, (2004) Estimativas de herdabilidade para idade ao primeiro parto de novilhas da raça Nelore. **Revista Brasileira de Zootecnia** 33:97-102.

FAO – Food and Agriculture Organization. Animal production and health. 2020. Disponível em: http://www.fao.org/ag/againfo/themes/en/animal_production.html. Acesso em: 17/03/2020.

Faria CU, Magnabosco CU, Reyes AI, Lôbo RB, et al. (2007) Bayesian inference on field data for genetic parameters for some reproductive and related traits of Nellore cattle (*Bos indicus*). **Genetics and Molecular Biology** 30:343-348.

Ferencakovic M, Hamzic E, Gredler B, Curik I, Sölkner J, (2011) Runs of Homozygosity Reveal Genome-wide Autozygosity in the Austrian Fleckvieh Cattle. **Agriculturae Conspectus Scientificus** 76:325-328.

Fernandes Júnior G A, Rosa G J M, Valente BD, Carvalho R, et al. (2016) Genomic prediction of breeding values for carcass traits in Nellore cattle. **Genetics Selection Evolution** 48:7.

Formigoni IB, Ferraz JBS, Silva JAIV, Eler JP, Brumatti RC, (2005) Valores econômicos para habilidade de permanência e probabilidade de prenhez aos 14 meses em bovinos de corte. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia** 57:220-226.

Fortes MRS, Enculescu C, Neto LRP, Lehnert SA, McCulloch R, Hayes B, (2018) Candidate mutations used to aid the prediction of genetic merit for female reproductive traits in tropical beef cattle. **Revista Brasileira de Zootecnia** 47:e20170226, <https://doi.org/10.1590/rbz4720170226>

Goddard ME and Hayes BJ, (2002) Optimisation of response using molecular data. In: 7th World Congress on Genetics Applied to Livestock Production, **Communication...** Montpellier, France.

Goddard ME and Hayes BJ, (2007) Genomic selection. **Journal of Animal Breeding and Genetic** 124:323–330.

Goddard ME, Hayes B (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews | Genetics** 10:381-391.

Goddard ME, Hayes BJ, (2012) **Genome-Wide Association Studies and Linkage Disequilibrium in Cattle**. Bovine genomics. Edited by James E. Womack. Oxford, United Kingdom: John Wiley e Sons 192-210. <https://doi.org/10.1002/9781118301739.ch13>.

Goddard ME, Wray NR, Verbyla K, Visscher PM (2009) Estimating Effects and Making Predictions from Genome-Wide Marker Data. **Statistical Science** 24:517-529.

Grossi DA, Frizzas OG, Paz CCP, Bezerra LAF, Lôbo RB, Oliveira JA, Munari DP, (2008) Genetic associations between accumulated productivity, and reproductive and growth traits in Nelore cattle. **Livestock Science** 117:139-146.

Grossi DA, Venturini GC, Paz CCP, Bezerra LAF, Lôbo RB, Oliveira JA, Munari DP (2009) Genetic associations between age at first calving and heifer body weight and scrotal circumference in Nelore cattle. **Journal of Animal Breeding and Genetics** 126:387-393.

Hayes B (2013) Overview of Statistical Methods for Genome-Wide Association Studies (GWAS). **Methods in Molecular Biology** 1019:149-69. doi: 10.1007/978-1-62703-447-0_6

Hayes B, Goddard M (2010) Genome-wide association and genomic selection in animal breeding. **NRC Research Press** 53:876-883.

Hayes BJ, Lewin HA, Goddard ME, (2013) The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. **Trends in Genetics** 29:4.

Irano N, Camargo GMF, Costa RB, Terakado APN, Magalhães AFB, Silva RMO, Dias MM, Bignardi AB, Baldi F, Carneiro R, Oliveira HN, Albuquerque LG, (2016) Genome-Wide Association Study for Indicator Traits of Sexual Precocity in Nelore Cattle. **PLoS ONE**, 11:8. doi:10.1371/journal.pone.0159502

Jamrozik J, Miller SP, (2014) Genetic evaluation of calving ease in Canadian Simmentals using birth weight and gestation length as correlated traits. **Livestock Science** 162:42–49.

Jiménez-Montero JA, González-Recio O, Alenda R (2013) Comparison of methods for the implementation of genome-assisted evaluation of Spanish dairy cattle. **Journal of Dairy Science** 96 :625–634.

Kim Y, Ryu J, Woo J, Kim JB, Kim CY, Lee C, (2011) Genome-wide association study reveals five nucleotide sequence variants for carcass traits in beef cattle. **Animal Genetics** 42:361-365.

Kluska S, Olivieri BF, Bonamy M, Chiaia H LJ, Feitosa FLB, Berton MP, Peripolli E, Lemos MVA, Tonussi RL, Lôbo RB, Magnabosco CU, DI Croce F, Osterstock J, Pereira ASC, Munari DP, Bezerra LA, Lopes F B, Baldi F, (2018) Estimates of genetic parameters for growth, reproductive, and carcass traits in Nelore cattle using the single step genomic BLUP procedure. **Livestock Science** 216:203-209.

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

Lee SH, Choi BH, Lim D, et al. (2013) Genome-Wide Association Study Identifies Major Loci for Carcass Weight on BTA14 in Hanwoo (Korean Cattle). **Plos One** 8:1-9.

Legarra A, Christensen OF, Aguilar I, Misztal I, (2014) Single Step, a general approach for genomic selection. **Livestock Science** 166:54–65.

Lôbo R, Bezerra LAF, Oliveira HN, Garnero AV, Schwengber EB, Marcondes CR, (2000) Avaliação genética de animais jovens, touros e matrizes: **Sumário 2000**. SP, Ribeirão Preto. Brasil.

Lôbo RB, Bezerra LAF, Figueiredo LGG, Baldi F, Faria CU, Vozzi PA, Magnabosco CU, Bergmann JAG, Oliveira HN, (2016) **Sumário de Touros das Raças Nelore, Guzerá, Brahman e Tabapuã**: Edição Agosto de 2016. Ribeirão Preto, ANCP 116 p. ISSN 1981-1705

Makowsky R, Pajewski NM, Klimentidis YC, Vazquez AI, Duarte CW, Allison DB, Campos GDL (2011), Beyond Missing Heritability: Prediction of Complex Traits. **PLOS Genetics** 7:1-9.

Melo TP, Fortes MRS, Bresolin T, Mota LFM, Albuquerque LG, Carneiro R (2018) Multitrait meta-analysis identified genomic regions associated with sexual precocity in tropical beef cattle. **Journal of Animal Science** 96:4087-4099.

Melo TP, Fortes MRS, Hayes B, Albuquerque LG, Carneiro R (2019) Across-breed validation study confirms and identifies new loci associated with sexual precocity in Brahman and Nelore cattle. **Journal of Animal Breeding and Genetics** 00:1-16.

Melo TPD, DE Camargo GMF, DE Albuquerque LG, Carneiro R, (2017). Genome-wide association study provides strong evidence of genes affecting the reproductive performance of Nelore beef cows. **PLoS ONE** 12(5): e0178551. <https://doi.org/10.1371/journal.pone.0178551>

Meuwissen T, Hayes B, Goddard M (2013) Accelerating Improvement of Livestock with Genomic Selection. **Annual Reviews Animal Biosciences** 1:221-237.

Meuwissen TH, Hayes BJ, Goddard ME, (2001) Prediction of total genetic value using genome-wide dense marker maps. **Genetics** 157:1819-1829.

Misztal I, Legarra A, Aguilar I, (2009) Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. **Journal of Dairy Science** 92:4648-4655.

Moreira HL, Buzanskas ME, Munari DP, Canova ÉB, Lôbo RB, De Paz CCP, (2015) Reproductive traits selection in nelore beef cattle. **Ciências e Agrotecnologia** 39:355-362.

Mota RR, Guimarães SEF et al, 2017. Genome-wide association study and annotating candidate gene networks affecting age at first calving in nelore cattle. **J Anim Breed Genet** 134:484-492. <https://doi.org/10.1111/jbg.12299>

Nascimento AV, Romero ARS, Utsunomiya YT, et al. (2018) Genome-wide association study using haplotype alleles for the evaluation of reproductive traits in Nelore cattle. **Plos One** 13:1-14.

NCBI - NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. Disponível em: <https://www.ncbi.nlm.nih.gov/genome?term=bos%20taurus>. Acesso em: 27/11/2016.

Nejati-Javaremi A, Smith C, Gibson JP, (1997). Effect of total allelic relationship on accuracy of evaluation and response to selection. *Journal Animal Science* 75(7):1738-1745. doi:10.2527/1997.7571738x

Neves HHR, Carneiro R, et al. (2014) Accuracy of genomic predictions in *Bos indicus* (Nelore) Cattle. **Genetics Selection Evolution** 46:17.

Nishimura S, Watanabe T, Mizoshita K, Tatsuda K, Fujita T, Watanabe N, Sugimoto Y, Takasuga A, (2012) Genome-wide association study identified three major QTL for carcass weight including the PLAG1-CHCHD7 QTN for stature in Japanese Black Cattle. **BMC Genetics** 13:40.

Niu T, (2004) Algorithms for Inferring Haplotypes. **Genetic Epidemiology** 27:334-347.

Nogalski Z, Piwczyński D, (2012) Association of Length of Pregnancy with Other Reproductive Traits in Dairy Cattle. **Asian-Australasian Journal of Animal Science** 25:22-27

Nowacka-Woszek J, (2020) Nutrigenomics in livestock-recent advances. **Journal of Applied Genetics** 61:93-103.

O'Brien AMP, Mészáros G, Utsunomiya YT, Sonstegard TS, Garcia J F, Van Tassell CP, Carvalheiro SR, Sölkner MVB, (2014) J. Linkage disequilibrium levels in *Bos indicus* and *Bos taurus* cattle using medium and high density SNP chip data and different minor allele frequency distributions. **Livestock Science** 166:121-132.

Oldenbroek K and Waaij LVD, (2015). Textbook Animal Breeding and Genetics for BSc students. **Centre for Genetic Resources The Netherlands and Animal Breeding and Genomics Centre**. Groen Kennisnet: <https://wiki.groenkennisnet.nl/display/TAB/>

Sartori R, Bastos MR, Baruselli PS, Gimenes LU, Ereno RL, Barros CM, (2010) Physiological differences and implications to reproductive management of *Bos taurus* and *Bos indicus* cattle in a tropical environment. **Soc Reprod Fertil Suppl.** 67: 357–375.

Schmidt PI, Ferreira IA, Silveira DD, Campos GS, Souza FRP, R.Carvalheiro, Boligon AA (2019) Reproductive performance of cows and genetic correlation with weight gains and principal components of traits used in selection of Nelore cattle. **Livestock Science** 229:77-84.

Schwengber EB, Bezerra LAF, Lôbo RB, (2001) Produtividade acumulada como critério de seleção em fêmeas da raça nelore. **Ciência Rural** 31:483-486.

Shim H, Chun H, Engelman C, Payseur BA (2009) Genome-wide association studies using single-nucleotide polymorphisms versus haplotypes: an empirical comparison with data from the North American Rheumatoid Arthritis Consortium. **BMC Proceedings** 35:1-6.

Silva JAIV, Eler JP, Ferraz JBS, Oliveira HN, (2003) Análise Genética da Habilidade de Permanência em Fêmeas da Raça Nelore1. **Revista Brasileira de Zootecnia** 32:598-604.

Silva RMO, Fragomeni BO, et al. (2016) Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population. **Journal of Animal Science** 94:3613-3623.

Sun X, Fernando RL, Garrick DJ, and Dekkers JCM, (2011) An iterative approach for efficient calculation of breeding values and genome-wide association analysis using weighted genomic BLUP. **Journal of Animal Science** 89.

Terakado APN, Boligon AA, Baldi F, Silva JAIV, Albuquerque LG (2015) Genetic associations between scrotal circumference and female reproductive traits in Nelore cattle. **Journal of Animal Science** 93:2706-2713.

USDA – UNITED STATES DEPARTMENT OF AGRICULTURE. Livestock and Poultry: World Markets and Trade. 2020. Disponível em: https://apps.fas.usda.gov/psdonline/circulars/livestock_poultry.pdf. Acesso em: 09/03/2020.

Van Eenennaam AL, Weigel KA, Young AE, Cleveland MA, Dekkers JC M, (2014) Applied Animal Genomics: Results from the Field. **Annual Review Animal Biosciences** 2:105-139.

Van Melis MH, Eler JP, Oliveira HN, Rosa GJM, Silva JAV, Ferraz JBS, Pereira E (2007) Study of stayability in Nelore cows using a threshold model. **Journal of Animal Science** 1-23.

Van Melis MH, Oliveira HN, Eler JP, Ferraz JBS, Casellas J, Varona L, (2010) Additive genetic relationship of longevity with fertility and production traits in Nelore cattle based on bivariate models. **Genetics and Molecular Research** 9 (1):176-187.

Vozzi PA, Oliveira HN, Siqueira RG, Paz CC, Assagra WLO, Bezerra L AF, Lôbo RB, (2008) Probabilidade de Parto Precoce como critério de seleção para precocidade sexual em bovinos da raça Nelore. In: **SBMA - VII Simpósio Brasileiro de Melhoramento Animal**. São Carlos - SP.

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.

Wu Y, Fan H, Wang Y, Zhang L, Gao X, Chen Y, Li J, Ren H, Gao H, (2014) Genome-Wide Association Studies Using Haplotypes and Individual SNPs in Simmental Cattle. **PLOS One** 9:10.

Zhang Z, Liu J, Ding X, Bijma P, DE Koning DJ, and Zhang Q, (2010) Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. **PLOS One** 5:9.

CAPÍTULO 2 - Genome-wide association study provides insights into important genes for reproductive traits in Nelore cattle

ABSTRACT – The identification of genomic regions associated with reproductive traits as well as their biological processes allows a better understanding of the phenotypic variability of these traits. This information could be applied in animal breeding programs to accelerate genetic gain. The aim of this study was to evaluate the association between single nucleotide polymorphisms (SNP) with scrotal circumference at 365 days of age (SC365) and at 450 days of age (SC450), gestation length (GL) as a calf trait, age at first calving (AFC), accumulated productivity (ACP), heifer early calving until 30 months (HC30) and stayability (STAY) traits, in order to identify candidate genes and biological pathways associated with reproductive traits in Nelore beef cattle. The data set consisted of pedigree, phenotypes and genotypes of Nelore beef cattle, from the "Associação Nacional de Criadores e Pesquisadores" (ANCP). The association analyses were performed using the Weighted Single-Step Genome-Wide Association method (WssGWAS) and the regions, consisted of 10 consecutive SNPs, that explained more than 0.5% of additive genetic variance were considered as significant association. A total of 3, 6, 7, 5, 10, 25 and 12 windows were associated with SC365, SC450, GL, AFC, ACP, HC30 and STAY respectively. The results revealed genes with important functions for reproductive traits, as fertility and precocity. Some genes were associated with more than one trait, among them, CAMK1D, TASP1, ACOXL, RAB11FIP5 and SFXN5. Moreover, the genes were enriched in functional terms, like negative regulation of fat cell differentiation, fatty acid alpha-oxidation and sphingolipids signaling pathway. The identification of the genes associated with the traits as well as genes enriched in the terms and pathway mentioned above, should contribute to future biological validation studies and may be used as candidate genes in Nelore breeding programs.

Key words: Beef cattle, GWAS, reproduction, candidate genes

1 INTRODUCTION

The greater inclusion of animal protein in human food has increased the demand for beef production. Producers and researchers have been working, using the benefits of genetics, nutrition and animal reproduction, in the search for a performance that is compatible with the growth of beef cattle production and that results in an increase in the productivity (FAO, 2020; Nowacka-Woszuik, 2020).

Reproductive traits are economically relevant for beef cattle production system (Brumatti et al., 2011). However, these traits result in slow genetic gain, as they are controlled by many genes and have relatively low to moderate heritability, that is, they are strongly influenced by environmental factors (Cardoso et al., 2015). Furthermore, most reproductive trait is measured late in the life of the animals and are limited by

sex, making selection and genetic gain difficult (Dias et al., 2004; Cammack et al., 2009; Boligon and Albuquerque 2011).

Genomic studies have been widely used for beef cattle. Single nucleotide polymorphism (SNP) applied in genome-wide association studies (GWAS) can be used to identify possible associations between chromosomal regions and complex traits, such as reproductive traits. GWAS have been developed aiming to identify candidate genes and to understand the genetic mechanisms involved with the reproductive traits (Irano et al., 2016; Melo et al., 2018). Thus, the use of more informative markers and the disposal of those that generate noise in the predictions, could be improve the accuracy of genomic predictions (Nascimento et al., 2018).

The SNP associated with genes that influence reproductive traits, once identified, can be incorporated into SNP panels to increase the accuracy of genomic predictions for fertility and contribute to the selection process (Cochran et al., 2013). Therefore, the aim of this study is to evaluate the association between SNP and scrotal circumference at 365 days of age (SC365) and at 450 days of age (SC450), gestation length (GL) as a calf trait, age at first calving (AFC), accumulated productivity (ACP), heifer early calving until 30 months (HC30) and stayability (STAY) traits, in order to identify candidate genes and biological pathways associated with reproductive traits in Nelore beef cattle.

2 MATERIAL AND METHODS

2.1 Data

The data set was composed of pedigree, phenotypes and genotypes of Nelore beef cattle from the Associação Nacional de Criadores e Pesquisadores (ANCP). The ANCP animal breeding program aims to identify, select and offer to the beef market, genetically superior reproducers for productive and reproductive traits. The pedigree used was composed of 140,199 Nelore animals born from 1934 to 2016, with the maximum number 23 generations and with an effective population size equal to 403.84 (CFC and ENGOG software's). The phenotypic data set consisted a total of 50,331 animals with observations for least one of the SC365, SC450, GL, AFC, ACP, HC30 and STAY traits in Nelore beef cattle.

The ACP index was calculated under the expression described in the studies by Lôbo et al. (2000):

$$ACP = \frac{W_{210} \times n_c \times 365}{ADC_n - 550}$$

where, W_{210} is the average body weight of weaned calves corrected for 210 days of age, n_c is the total number of calves produced and ADC_n is the dam's age at last calving. The 365 and 550 constants enable fertility expression on an annual basis.

For HC30, the challenged females that had confirmed pregnancy and calving for the first time until 30 months of age, their phenotype was categorized as indicative of success, for females that did not calving until this age, as failure, and for females that did not reach this age and did not confirm pregnancy, as missing information. For Stayability, cows that have attained at least three calving at 76 months of age had the phenotype categorized as "success", otherwise, cows that did not had three calving until this age, had the phenotype described "failure".

The animals were raised on pasture, weaned between 6 and 8 months of age and the reproductive management consisted of the mating season lasting from 60 to 120 days using artificial insemination or controlled natural mating. The contemporary groups (CG) were constituted by sex (when the trait was measured in both sexes), animals born on the same farm, year and season and belonging to the same management group.

The genotypic data set was composed of 8.652 animals (males and females). From these animals, a total of 960 bulls were genotyped with HD SNP panel (Illumina Bovine HD BeadChip), 1,000 animals were genotyped with a medium-density SNP panel (50k-Illumina BovineSNP50 BeadChip) and 6,692 animals were genotyped with a low-density SNP panel (12k-Clarifide Nelore 2.0). The animals genotyped with lower density were imputed to medium-density (50k) and subsequently to the HD panel using FImpute v.2.2 (Sargolzaei et al., 2014). In the imputation the bulls genotyped with HD SNP panel were used as reference population and the accuracy estimated of imputation was higher than 0.98.

2.2 Quality Control

2.2.1 Phenotype Data

The CG with frequency less than five animals were excluded. The connectivity analysis between CG for each trait was performed using the AMC software (Roso and

Schenkel, 2006), and the disconnected CG were disregarded in the analyzes. The residual normality and variance homogeneity assumptions were performed and observations with ± 3.0 standard deviations were excluded for each trait. The number of observations after phenotype data editing is summarized in Table 1.

Table 1. Number of observations (N), mean and respective standard deviation (SD), minimum, maximum and number of contemporary groups (CG) for scrotal circumference at 365 days of age (SC365), scrotal circumference at 450 days of age (SC450), gestation length (GL) as a calf trait, age at first calving (AFC), accumulated productivity (ACP), heifer early calving until 30 months (HC30) and stayability (STAY)

Trait	N	Mean*	SD	Minimum	Maximum	CG
SC365(cm)	15,054	21.56	2.40	14.40	29.10	426
SC450(cm)	13,694	25.01	3.23	16.00	35.40	381
GL (days)	16,252	296.80	6.07	278.00	314.00	202
AFC (months)	15,384	35.24	5.91	21.00	49.00	237
ACP (kg)	6,205	150.70	31.12	63.00	249.00	150
HC30	3,236	0.30	-	0	1.00	97
STAY	12,981	0.42	-	0	1.00	162

* For HC30 and STAY the mean represent the “success” proportion for each trait

2.2.2 Genotype Data

The SNP with call rate less than 90% and minor allele frequency (MAF) less than 5% were excluded. Also, samples with a call rate less than 90% were not considered in analyses. Only autosomal SNPs were considered. After genomic data quality control, 460,838 SNP and 8,545 animals with observations for least one of trait were available.

2.3 Genome-wide association analysis

The methodology used was the weighted single-step genome-wide association (WssGWAS), that is a modification of BLUP method, with the numerator relationship matrix A^{-1} replaced by the genomic matrix H^{-1} (Aguilar et al., 2010):

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

where A is the numerator relationship matrix for all animals; A_{22} is the numerator relationship matrix for genotyped animals and G is the genomic relationship matrix. The G matrix was obtained according to VanRaden (2008):

$$G = ZDZ'q$$

where Z is the coefficient matrix adjusted for allele frequency; D is the weight matrix for each SNP and q is the weighting factor. The weighting factor was calculated by ensuring that the mean diagonal in G will be close to A_{22} .

For derivation of the SNP effects and weights, the animal effect was decomposed in genotyped animals (a_g) and not genotyped (a_n), as described by Wang et al. (2012). The animal effect of the genotyped animals is a function of the SNP effects:

$$a_g = Z_u$$

where Z is a matrix relating genotypes of each loci and u is a vector of SNP effects. The animal effect variance is:

$$var(a_g) = var(Z_u) = ZDZ'\sigma_u^2 = G^*\sigma_a^2$$

where D is the diagonal matrix of weights for variances of SNP; σ_u^2 is the genetic additive variance captured by each SNP and G^* is the weighted of relationship genomic matrix.

Thus, the SNP effect was obtained following the equation of Wang et al. (2012):

$$\hat{u} = \lambda DZ'G^{*-1}\hat{a}_g = DZ'[ZDZ']^{-1}\hat{a}_g$$

where λ is a normalizing constant, defined by VanRaden et al. (2009):

$$\lambda = \frac{\sigma_u^2}{\sigma_a^2} = \frac{1}{\sum_{i=1}^M 2p_i(1-p_i)}$$

where, M is the number of SNP and p_i is the allele frequency of the second allele of the i th SNP. The iterative process described by Wang et al. (2012) in six steps was used: Step one, $D = I$; step two, to calculate \hat{a}_g with WssGWAS; step three, to calculate the SNP effect; step four, to calculate the variance of each SNP, $d_i = \hat{u}_i^2 2p_i(1-p_i)$, where i is the i th SNP; step five, the normalized the value of the SNP to keep constant the additive genetic variance and step six, to calculate the G matrix. A weighted process that recalculated the weights of SNP, consisted of repeating once steps 2 to 6 (Wang et al., 2012).

This process increases SNP weights explaining larger genetic variance and decrease SNP weights explaining small genetic variance. The percentage of genetic variance explained by the i th region was calculated as follows:

$$\frac{Var(a_i)}{\sigma_i^2} \times 100\% = \frac{Var(\sum_{j=1}^{10} Z_j \hat{u}_j)}{\sigma_a^2} \times 100\%$$

where a_i is the genetic value of the i th region that consist of 10 consecutives SNP, σ_a^2 is the total genetic variance, Z_j is the vector of SNP content of the j th SNP for all individuals and \hat{u}_j is the marker effect of the j th SNP within the i th region.

The analyzes for estimation of the genomic association were performed using software from the BLUPF90 program family (Misztal et al., 2002; Aguilar et al., 2011). The SC365, SC450, GL, AFC and ACP traits were evaluated under linear model, while for HC30 and STAY traits, a threshold model was applied.

The general animal models used for SC365, SC450, AFC, ACP, STAY and HC30 (1) and for GL (2) were:

$$(1) y_t = Xb + Z_1a + e$$

$$(2) y_t = Xb + Z_1a + Z_2m + e$$

where y_t is the vector of phenotypic observations for each trait, except for STAY and HC30, that is the threshold vector; b is the vector of fixed effects, that included CG; a is the vector of effects of the animals; m is the vector of maternal effect e is the vector of residual effects; X , Z_1 and Z_2 are the incidence matrix for b , a and m , respectively. The variances of a and e are:

$$\begin{bmatrix} a \\ m \\ e \end{bmatrix} = \begin{bmatrix} H\sigma_a^2 & H\sigma_{am} & 0 \\ H\sigma_{am} & H\sigma_m^2 & 0 \\ 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

where σ_a^2 is the additive genetic variance; σ_m^2 is the maternal additive genetic variance; σ_e^2 is the residual variance; σ_{am} is the additive-maternal covariance and H is the matrix that combines the relationship and genomic information matrix and I is the identity matrix.

2.4 Search for Associated Genes

The chromosome regions consisting of windows with 10 consecutive SNPs, that explained more than 0.5% of additive genetic variance were considered as associated with the studied trait and then the investigation of the genes that this region contained was carried out with Ensembl Genome Browser (<http://www.ensembl.org/index.html>) using the Ensembl gene 94 version of the gene model and UMD3.1 bovine genome assembly as reference. The enrichment of genes as well as its classification according to its biological function and identification of metabolic pathways was performed using the "Database for Annotation, Visualization and Integrated Discovery (DAVID) v. 6.8"

(<http://david.abcc.ncifcrf.gov/>) for each trait, with a P-Value significance criterion of 0.05.

3 RESULTS AND DISCUSSION

The associated regions with the traits studied are showed in Supplementary material A and presented in Table 2. A total of 3, 6, 2, 5, 10, 25 and 12 windows with 10 SNP were associated with SC355, SC450, GL, AFC, ACP, HC30 and STAY, respectively.

Table 2. Genomics regions associated with scrotal circumference at 365 days of age (SC365) and 450 of age (SC450), gestation length (GL) as calf trait, age at first calving (AFC), accumulated productivity (ACP), heifer early calving until 30 months (HC30) and stayability (STAY) traits in Nelore cattle and percentage of additive genetic variance

Chromosome	Position(bp)	% additive genetic variance
Scrotal circumference at 365 days of age		
BTA7	99797785-99807157	0.7598
BTA27	3459974-3476332	0.5827
BTA27	3763511-3783637	0.8425
Scrotal circumference at 450 days of age		
BTA4	93491117-93502503	1.7190
BTA5	54114064-5415865	2.3197
BTA17	30807597-30828891	0.5404
BTA19	16748888-16777564	0.6842
BTA23	30929417-30952562	1.2410
BTA26	489143-4913438	2.7796
Gestation length		
BTA1	89275294-89301259	0.6100
BTA13	75605755-75651541	0.6721
Age at first calving		
BTA3	71141852-71202433	0.5065
BTA11	21246-21281912	0.7776
BTA11	21287816-21311113	0.7324
BTA12	19870944-19879683	0.5148
BTA25	40542599-40557377	1.0960
Accumulated productivity		
BTA1	2054452-20570066	0.7828
BTA3	7752656-7785494	1.5813
BTA6	1442287-14434633	0.5264
BTA6	94344402-94375956	0.5681
BTA6	94388672-94426934	0.9580
BTA9	67670626-67696355	0.6988
BTA16	9824597-9861772	0.5885

Continuation...		
BTA16	1706924-17085278	0.8929
BTA20	19183661-19199006	0.5869
BTA25	31194971-31217552	0.5478
<hr/>		
Heifer early calving until 30 months		
BTA9	56459368-56477584	0.5289
BTA10	63914118-63931413	0.5813
BTA11	11320434-11348287	0.5008
BTA11	11355768-11368368	0.6445
BTA11	11374317-11381414	0.9560
BTA11	11410105-11424583	0.5888
BTA11	11439886-11463529	0.5572
BTA11	11496633-11536209	0.6600
BTA11	16578809-16627592	0.6191
BTA11	67791729-67805289	0.5965
BTA13	64147929-64202735	0.5643
BTA13	64212843-64261331	1.5723
BTA13	6426382-64340236	2.7497
BTA13	64358933-64432139	1.1148
BTA13	64450162-64500844	0.5546
BTA14	51334426-51347339	0.5336
BTA14	51399505-5142635	0.5006
BTA14	51586769-51622257	0.8759
BTA25	7204736-722268	0.8833
BTA25	17664012-17694691	1.3587
BTA25	18619922-18639119	0.7122
BTA25	18672747-18714923	1.5879
BTA25	18946238-18966876	1.3475
BTA25	40714999-40732777	0.5351
BTA27	35208391-35220792	0.6705
<hr/>		
Stayability		
BTA1	33285504-33305018	0.8321
BTA3	43009131-43028743	0.6630
BTA4	55984897-56013042	0.6909
BTA4	79187125-79206524	0.9377
BTA4	79221965-7930004	0.6522
BTA9	4370684-439556	2.4011
BTA9	4448707-4477071	1.2592
BTA11	1424088-1433201	0.5943
BTA13	1339928-13434427	0.5580
BTA15	15283071-15318647	0.5176
BTA26	25583443-25591332	0.6311
BTA27	22208624-22221777	0.9445
<hr/>		

The genes found in the significant regions for each trait are presented in the Supplementary material B. In the regions associated with SC365 (Table 2) no gene was found. A total of 5 known genes and 3 unknown genes are associated with SC450, and in the enrichment analysis no term was significantly enriched (≤ 0.05). A single known gene (CDH22) was found in GL-associated regions (Supplementary material B) and in the enrichment analysis no term was significantly enriched (≤ 0.05).

A total of 231 genes were found in the AFC-associated regions (Supplementary material B), with 178 known and 53 unknown genes. The enrichment analysis (≤ 0.05) for AFC associated genes in 6 terms (Supplementary material C) including mitochondrial translational initiation and elongation. In these pathways, the mitochondrial ribosomal protein S5 (MRPS5) and mitochondrial ribosomal protein S9 (MRPS9) genes have been enriched and have already been associated with feed intake (Olivieri et al., 2016). Basarab et al. (2003) showed that animals with positive feed intake consumed more feed and retained more energy. The AFC has a negative genetic correlation with weaning and yearling weight gain (-0.20 and -0.24, respectively) indicating that the selection for greater weight gain results in AFC decrease (Boligon et al., 2010).

A total of 244 genes were found in the ACP-associated regions (Supplementary material B), of which 156 are known genes and 88 are unknown. The total of 22 terms (Supplementary material D) were significantly enriched (≤ 0.05), being related to immune system processes and body growth.

The neuropeptide Y receptor Y1 (NPY1R) gene was enriched in the regulation of multicellular organism growth term. The NPY1R has been found in association with Angus heifers' fertility by Neupane (2017) and the neuropeptide Y associated with maternal behavior dependent on nutritional status in mouse (Muroi and Ishii, 2015). Added to this, and considering that ACP indicates the abilities of the female to calve at a young age, to maintain the regularity of calving, and to wean heavy calves (Grossi et al., 2008), the maternal behavior can influence the weaning weight of animals and cow fertility can influence not only in AFC, but also to maintain a shorter calving interval, which reflects the occurs of regular calving, evidencing the importance of the NPY1R gene for cow's ACP throughout life.

In the regions associated with HC30, 580 genes were found (Supplementary material B), of which 434 are known and 146 are unknown. Analyzing the total of genes, 22 terms (Supplementary material E) were significantly enriched (≤ 0.05),

among them fatty acid alpha-oxidation, negative regulation of fat cell differentiation and innate immune response.

The hydroxyacid oxidase (glycolate oxidase) 1 (HAO1), phytanoyl-CoA 2-hydroxylase (PHYH) and phytanoyl-CoA dioxygenase, peroxisomal-like (LOC100300115) genes was enriched into the fatty acid alpha-oxidation term, that is the pathway that degrades 3-methyl branched fatty acids, which are not degraded by the normal degradation pathway of fatty acids (peroxisomal beta-oxidation pathway) (Binns et al., 2009). The fatty acids metabolism is associated with success pregnancy (Melo et al., 2017).

Into the negative regulation of fat cell differentiation term, the E2F transcription factor 1 (E2F1), GATA binding protein 3 (GATA3), additional sex combs like 1, transcriptional regulator (ASXL1) and tribbles pseudokinase 3 (TRIB3) genes was enriched. This term involves processes that inhibit the differentiation of adipocytes (Binns et al., 2009). Considering the association of the two terms mentioned above with HC30, the results in the present study suggest that the favorable selection of alleles that inhibit degradation of 3-methyl branched fatty acids and which favor differentiation of adipocytes can result in higher fat deposition in these animals, which would be sexually precocious, i.e., with more success for HC30.

The biological process of innate immune response was significantly enriched (≤ 0.05) associated with HC30 trait. Melo et al. (2017) identified association between the number of calves at 53 months of age trait and the immune processes pathway. The phenotypic correlations were reported for Banos et al. (2013) between immune traits with reproductive performance traits, in special, the authors report that a higher concentration of CD8⁺ cells are associated with longer calving intervals. Thus, the immune system can interfere not only in the HC30 trait, but also in cow longevity, once that the longer calving interval result in lower stay in the herd.

A total of 52 genes were found in the STAY-associated regions (Supplementary material B), of which, 37 are known genes and 15 are unknown genes. The analysis of the biological functions of the genes enriched 6 terms (≤ 0.05) (Supplementary material F), among them, sphingolipid signaling pathway.

The lipid pathways are associated with the reproductive traits, such as the Nelore heifer reconception trait, indicating that cows with greater capacity to accumulate fat have better rebreeding performance (Melo et al 2017) and consequently greater STAY.

The sphingolipid signalin and its metabolic products have a signaling function in several other pathways, as the metabolic sphingosine-1-phosphate (S1P) that acts as a growth and survival factor and the ceramide (Cer) that activates the apoptotic pathways (Kanehisa and Goto, 2000). The related of this pathway to the activation of growth factor interfere in sexual precocity of the heifers, that can be take less time to reach the need weight for the reproductive maturity, thus decreasing AFC and increasing the cow's STAY. Also, by activating survival factor, this pathway can be associated with the longevity of cows in the herd, and thus, in STAY.

The phospholipase C beta 1 (PLCB1), enriched in the sphingolipid signaling pathway, and phospholipase C beta 4 (PLCB4) genes, both located in BTA 13, were associated with STAY. Early embryonic development in mammals depends on maternal miRNA stocks prior to the initiation of zygote regulation, this miRNA involvement is described in rats (Tang et al., 2007) and bovine (Abd El Naby et al., 2011). The PLCB1 is a target gene of the miR-205, and this miRNA was differentially expressed during pre-implantation embryonic development and both, PLCB1 and miR-205, have a reciprocal expression pattern (Abd El Naby et al., 2011). The PLCB1 and PLCB2 are gene target in KEGG pathways the miR-301b, which have association with ovarian follicle development in cattle (Zielak-Steciwko et al., 2014).

The jagged 1 (JAG1) and mitochondrial ribosomal protein L33 (MRPL33) genes, both located in BTA 13, were found associated with STAY. Although these genes have not been enriched in biological pathways, they have important reproductive functions. The JAG1 gene has a signaling function during the embryogenesis (Tripurani et al., 2011) and the MRPL33 gene was cited by Melo (2015) for being in association with heifer rebreeding in Nelore cattle. This fact reinforces the association between MRPL33 and STAY found in our work, because the stayability cow in the herd is influenced by its ability to reconception even when it's a heifer.

Many uncharacterized genes were found associated with the traits. These genes do not yet have an identified ortholog, which means that they do not have a homologous gene that has the same function in different organisms, therefore, in the enrichment analyzes these genes are not found. All the biological process, cell components, molecular functions and pathways that were significantly enriched (≤ 0.05) in the gene ontology analysis presented in Supplementary material C-F.

None of the genes found in the associated regions with each trait studied were associated with all the traits. The associated genes with SC450, GL and ACP do not

coincide with any gene associated with the other traits. According Kluska et al. (2018), the SC450 presented genetic correlation with AFC, HC30 and STAY equal to -0.29 ± 0.06 , $0.52 [0.31-0.74]$ and $0.35 [0.20-0.50]$, respectively. This indicates that selection for SC can decrease AFC (Moreira et al., 2015) and increase HC30 and STAY, therefore, it's an indicator of sexual precocity of Nelore heifers. Grossi et al. (2008) found a negative genetic correlation between ACP and AFC (-0.33 ± 0.04) in Nelore dams, i.e., the gains in ACP are related to a lower AFC. The genetic correlation existence between the studied traits may indicate that they are controlled by common genes.

The common genes associated with the reproductive traits are presented in Table 3. A single common and known gene is associated with AFC and STAY; 3 common genes are in association with AFC and HC30 (2 known genes); 31 genes common to HC30 and STAY (22 known genes).

Table 3. Common gene list associated with de with reproductive traits in Nelore cattle

Traits: Total Common Genes
AFC and STAY: 1 ACOXL
AFC and HC30: 3 RAB11FIP5, SFXN5, EMX1
HC30 and STAY: 31 SPTLC3, RF00273, ISM1, TASP1, ESF1, NDUFAF5, SEL1L2, FLRT3, MACROD2, KIF16B, ENSBTAG00000032609, SNRPB2, OTOR, NANP, ENSBTAG00000023169, ENSBTAG00000004354, ENSBTAG00000022922, ENSBTAG00000040046, CCDC3, CAMK1D, RF00619, CDC123, NUDT5, SEC61A2, DHTKD1, UPF2, RF00012, PROSER2, ECHDC3, USP6NL, CELF2

Different studies reported that AFC and STAY present negative genetic correlation ($-0.64 [-0.83 - -0.44]$ and -0.69) (Kluska et al., 2018; Rizzo et al., 2015), thus, the decrease in AFC would lead to an increase in STAY. The acyl-CoA oxidase like gene (ACOXL) located in the BTA 11 was found in the regions associated with AFC and STAY, so the selection targeted to this gene would allow a possible gain for both traits. ACOXL is a primary rate-limiting enzyme in peroxisomal fatty acids β -oxidation (Reubsæet et al., 1988; Oaxaca-Castillo et al., 2007) and has been found in association with pregnancy outcome (pregnant or not pregnant) using fixed-time artificial insemination in Brahman heifers (Porto-Neto et al., 2015).

The RAB11 Family Interacting Protein 5 (RAB11FIP5) and Sideroflexin 5 (SFXN5) genes, both located in BTA 11, was associated with AFC and HC30. The RAB11FIP5 and SFXN5 genes have already been described associated with shear force (Tizioto et al., 2013), a method used to evaluate the tenderness of meat. There is a high degree of interdependence of puberty traits and growth and body composition (Pardo et al., 2018). Pacheco et al. (2005) report that subcutaneous fat thickness is favored when younger animals are slaughtered. Still, part of a chromosomal region associated with the occurrence of early calving is also described when meat tenderness in Nelore cattle is investigated (Magalhães, 2015; Irano et al., 2016). In view of this, the RAB11FIP5 and SFXN5 genes seem to influence the sexual precocity in cattle, justifying the association of this genes with AFC and HC30 traits identified in the present study.

The gene calcium/calmodulin dependent protein kinase ID (CAMK1D) was found associated with HC30 and STAY. This gene is located in BTA 13 and its association with heifer rebreeding and age at first calving was found by Costa et al. (2015) in a GWAS study in Nelore dams. In addition, Melo et al., 2018 report that this gene may contribute to sexual precocity in Nelore cattle.

Threonine aspartase 1 (TASP1) gene, also associated with HC30 and STAY, generate alpha and beta subunits, forming the active alpha2-beta2, necessary to cleave the protein used for the maintenance of HOX gene expression. When the HOX gene has its expression altered, it interferes with uterine development, preventing the implantation of the embryo and causing infertility in women (Du and Taylor, 2016). In bovine, the HOX expression profile suggests a role during early developmental events including oocyte maturation, the maternal to embryonic transition, and the first differentiation events (Paul et al., 2011). In the same study, the comparison between two bovine and mouse species indicated that HOX expression at initial stages is partly conserved among mammals. Thus, the idea of the influence of the TASP1 gene on cow fertility is reinforced and justifies its association with STAY and HC30 traits.

4 CONCLUSIONS

The results in the present study allowed a better understanding of the associated genes with reproductive traits studied in Nelore cattle. Many genes found in the windows associated appear to be related with more than one trait, indicating that these

genes have a pleiotropic effect and are important for reproductive efficiency. Among these common genes, the CAMK1D and TASP1 genes associated with HC30 and STAY; the ACOXL gene that appears associated with AFC and STAY and RAB11FIP5 and SFXN5 genes associated with AFC and HC30 should be highlight due to their relationship with reproductive efficiency.

Many pathways and terms were associated with the reproductive traits in Nelore cattle, but the fatty acid alpha-oxidation and negative regulation of fat cell differentiation terms highlight for being related with HC30, while the sphingolipid signaling pathway highlight for being related with STAY. The identification of the genes associated with the traits as well as genes enriched in the terms and pathway mentioned above, should contribute to future biological validation studies and may be used as candidate genes in Nelore breeding programs.

5 REFERENCES

Abd EL Naby WS, Hagos TH, Hossain MM, Salilew-Wondim D, Gad AY, Rings F, Cinar MU, Tholen E, Looft C, Schellander K, Hoelker M, Tesfaye D, (2011) Expression analysis of regulatory microRNAs in bovine cumulus oocyte complex and preimplantation embryos. **Zygote** 21:31-51, doi:10.1017/S0967199411000566.

Aguilar I, Misztal I, Legarra A, Tsuruta S, Davis R, (2011) Efficient computations of genomic relationship matrix and other matrices used in the single-step evaluation. **Journal of Animal Breeding and Genetics** 128:422-8.

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 Dairy Science** 93:743-752.

Banos G, Wall E, Coffey MP, Bagnall A, Gillespie S, Russell GC, McNeilly TN (2013) Identification of Immune Traits Correlated with Dairy Cow Health, Reproduction and Productivity. **Plos One** 8:1-14.

Basarab JA, Price MA, Aalhus JL, Okine EK, Snelling WM, Lyle KL (2003) Residual feed intake and body composition in young growing cattle. **Canadian journal of animal science** 83:189-204.

Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R, (2009) QuickGO: a web-based tool for Gene Ontology searching. **Bioinformatics** 25:3045-3046. DOI:10.1093/bioinformatics/btp536.

Boligon AA, Albuquerque LG, (2011) Genetic parameters and relationships of heifer pregnancy and age at first calving with weigth gain, yearling and mature weigth in Nelore cattle. **Livestock Science** 141:12-16.

Boligon AA, Albuquerque LG, Mercadante MEz, Lôbo RB, (2010) Study of relations among age at first calving, average weight gains and weights from weaning to maturity in Nelore cattle. **Revista Brasileira de Zootecnia** 39:746-751.

Brumatti RC, Ferraz JBS, Eler JP, Formigoni IB, (2011) Desenvolvimento de índice de seleção em gado corte sob o enfoque de um modelo bioeconômico. **Archivos Zootecnia** 60:205-213.

Cammack KM, Thomas MG, Enns RM (2009) REVIEW: Reproductive Traits and Their Heritabilities in Beef Cattle. **Professional Animal Scientists** 25:517-528.

Cardoso RC, Alves BRC, Sharpton SM, Williams GL, Amstalden M (2015) Nutritional Programming of Accelerated Puberty in Heifers: Involvement of POMC Neurons in the Arcuate Nucleus. **Journal of Neuroendocrinology** 8:647-57.

Cochran SD, Cole JB, Null DJ, and Hansen PJ, (2013). Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. **BMC Genetics** 14:49.

Costa RB, Camargo GMF, Diaz IDPS, Irano N, Dias MM, Carvalheiro R, Boligon AA, Baldi F, Oliveira HN, Tonhati H, Albuquerque LG, (2015). Genome-wide association study of reproductive traits in Nelore heifers using Bayesian inference. **Genetics Selection Evolution** 47:67. DOI: 10.1186/s12711-015-0146-0.

Dias LT, El Faro L, Albuquerque LG, (2004) Estimativas de herdabilidade para idade ao primeiro parto de novilhas da raça Nelore. **Revista Brasileira de Zootecnia** 33:97-102.

Du H, Taylor HS, (2016) The Role of Hox Genes in Female Reproductive Tract Development, Adult Function, and Fertility. **Cold Spring Harb Perspect Med** 6:a023002

FAO – Food and Agriculture Organization. Animal production and health. 2020. Available in: http://www.fao.org/ag/againfo/themes/en/animal_production.html. Access in: 17/03/2020.

Grossi DA, Frizzas OG, Paz CCP, Bezerra LAF, Lôbo RB, Oliveira JA, Munari DP, (2008) Genetic associations between accumulated productivity, and reproductive and growth traits in Nelore cattle. **Livestock Science** 117:139-146.

Irano N, Camargo GMF, Costa RB, Terakado APN, Magalhães AFB, Silva RMO, Dias MM, Bignardi AB, Baldi F, Carvalheiro R, Oliveira HN, Albuquerque LG, (2016) Genome-Wide Association Study for Indicator Traits of Sexual Precocity in Nelore Cattle. **PLoS ONE**, 11:8. doi:10.1371/journal.pone.0159502

Irano N, Camargo GMF, et al. (2016) Genome-Wide Association Study for Indicator Traits of Sexual Precocity in Nelore Cattle. **PLoS ONE** 8:e0159502.

Kanehisa M, Goto SK, (2000) Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research** 28:27-30.

Kluska S, Olivieri BF, Bonamy M, Chiaia HLJ, Feitosa FLB, Berton MP, et al. (2018) Estimates of genetic parameters for growth, reproductive, and carcass traits in Nelore cattle using the single step genomic BLUP procedure. **Livestock Science** 216:203-209.

Lôbo R, Bezzera LAF, Oliveira HN, Garnero AV, Schwengber EB, Marcondes CR, (2000) Avaliação genética de animais jovens, touros e matrizes: **Sumário 2000**. SP, Ribeirão Preto. Brasil.

Magalhães AFB (2015) **Utilização de informações genômicas para o melhoramento genético de características da carne em bovinos da raça nelore**. 60 f. Tese (Doutorado em Zootecnia) – Unesp, Jaboticabal.

Melo TP, (2015) **Genome-wide association study of reproduction traits in nelore cattle, including additional phenotypic information from non-genotyped animals**. 59 f. Dissertação (Mestrado em Genética e Melhoramento Animal) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Jaboticabal.

Melo TP, Camargo GMF, Albuquerque LG, Carneiro R (2017) Genome-wide association study provides strong evidence of genes affecting the reproductive performance of Nelore beef cows. **Plos One** 12:1-14.

Melo TP, Bresolin T, Mota LFM, Fortes MRS, Albuquerque LG, Carneiro R, (2018) Multi-trait meta-analysis for sexual precocity traits in Nelore cattle. **Proceedings of the World Congress on Genetics Applied to Livestock Production** 11.178.

Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH, (2002) “**BLUPF90 and related programs (BGF90)**”, in Proceedings of the 7th World Congress on Genetics Applied to Livestock Production 28:21-22.

Moreira HL, Buzanskas ME, Munari DP, Canova ÉB, Lôbo RB, DE Paz CCP, (2015) Reproductive traits selection in nelore beef cattle. **Ciências e Agrotecnologia** 39:355-362.

Muroi Y, Ishii T, (2015) Neuropeptide Y is crucial for nutritional state-dependent regulation of maternal behavior. **Psychoneuroendocrinology** 51:392-402. Doi: <http://dx.doi.org/10.1016/j.psyneuen.2014.09.022>.

Nascimento AV, Romero ARS, Utsunomiya YT, et al. (2018) Genome-wide association study using haplotype alleles for the evaluation of reproductive traits in Nelore cattle. **Plos One** 13:1-14.

Neupane M, Geary TW, Kiser JN, Burns GW, Hansen PJ, Spencer TE, Neibergs HL, (2017) Loci and pathways associated with uterine capacity for pregnancy and fertility in beef Cattle. **PLoS ONE**, 12 (12): e0188997. DOI: <https://doi.org/10.1371/journal.pone.0188997>.

Nowacka-Woszuik J, (2020) Nutrigenomics in livestock-recent advances. **Journal of Applied Genetics** 61:93-103.

Oaxaca-Castillo D, Andreoletti P, Vluggens A, Yu S, Veldhoven PPV, Reddy JK, Cherkaoui-Malki M, (2007) Biochemical characterization of two functional human liver acyl-CoA oxidase isoforms 1a and 1b encoded by a single gene. **Biochemical and Biophysical Research Communications** 360:314–319.

Olivieri BF, Mercadante MEZ, Cyrillo JNSG, Branco RH, Bonilha SFM, Albuquerque LG, Silva RMO, Baldi F, (2016) Genomic Regions Associated with Feed Efficiency Indicator Traits in the Experimental Nellore Cattle Population. **PLoS ONE**. DOI:10.1371/journal.pone.0164390.

Pacheco PS, Silva JHS, Restle, Arboitte MZ, Brondani IL, Alves Filho DC, Freitas AK, (2005) Características Quantitativas da Carcaça de Novilhos Jovens e Superjovens de Diferentes Grupos Genéticos. **Revista Brasileira Zootecnia** 5:1666-1677.

Pardo AM, Villarreal EL, Papaleo Mazzucco J, Melucci OG, Santamaría S, Ferrario J, Melucci LM, (2018) Sexual precocity and productivity of beef cattle female under grazing conditions. **Animal Production Science** doi: 10.1071/AN17804.

Paul D, Bridoux L, Rezsóhazy R, Donnay I, (2011) HOX Genes Are Expressed in Bovine and Mouse Oocytes and Early Embryos. **Molecular Reproduction & Development** 78:436–449.

Porto-Neto LR, Edwards S, Fortes MRS, Lehnert SA, Reverter A, McGowan M, (2015) Genome-wide association for the outcome of fixed-time artificial insemination of Brahman heifers in northern Australia. **Journal of Animal Science** 93:5119-5127. doi:10.2527/jas2015-9401.

Reubsæet FAG, Veerkamp JH, B&kens SGF, Trijbels JMF, LAH Monnens, (1988) Acyl-CoA oxidase activity and peroxisomal fatty acid oxidation in rat tissues. **Biochimica et Biophysica Acta** 958:434-442.

Rizzo ECA, Neto FRA, Diaz IDPS, Dias MM, Costa RB, Ventura HT, Oliveira HN, Falcão AJ, (2015) Genetic association of productive and reproductive traits with stayability in Nellore cattle: analysis using Bayesian models. **Genetics and Molecular Research**,14:14956-14966.

Roso VM, Schenkel FS, (2006) AMC – A computer program to assess the degree of connectedness among contemporary groups. In: 8th World Congress on Genetics Applied to Livestock Production. **Resumos...** Belo Horizonte, MG. V. 34.

Sargolzaei M, Chesnais JP and Schenkel FS (2014) A new approach for efficient genotype imputation using information from relatives. **BMC Genomics**, 15:471-478.

Tang F, Kaneda M, O'carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovskiy A, Lao K, Surani MA, (2007) Maternal microRNAs are essential for mouse zygotic development. **Genes & Development** 21:644-648.

Tizioto PC, Decker JE, et al. (2013) Genome scan for meat quality traits in Nelore beef cattle. **Physiol Genomics** 45:1012–1020.

Tripurani SK, Lee K-B, Wang L, Wee G, Smith GW, Lee YS, Latham KE, Yao JA, (2011) Novel Functional Role for the Oocyte-Specific Transcription Factor Newborn

Ovary Homeobox (NOBOX) during Early Embryonic Development in Cattle. **Endocrinology** 152:1013-1023.

Vanraden PM, (2008) Efficient methods to compute genomic predictions. **Journal Dairy Science** 91:4414-4423

Vanraden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS, (2009) Invited review: Reliability of genomic predictions for North American Holstein bulls. **Journal of Dairy Science** 92:16-24.

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.

Zielak-Steciwko AE, Browne JA, Mcgettigan PA, Gajewska M, DZI Eciół M, Szulc T, EVans ACO, (2014) Expression of microRNAs and their target genes and pathways associated with ovarian follicle development in cattle. **Physiol Genomics**, 46:735-745.

CAPÍTULO 3 - Predição genômica de características reprodutivas em bovinos da raça Nelore

RESUMO – A seleção genômica vem sendo aplicada em programas de melhoramento genético animal para obter valores genéticos mais acurados que auxiliam na tomada de decisão ao selecionar animais superiores geneticamente, os quais são selecionados precocemente. Para contribuir com a aplicação de seleção genômica em bovinos Nelore, os objetivos do presente trabalho foram: (1) Comparar a habilidade de predição dos métodos Melhor preditor linear não-viesado (BLUP) e Melhor preditor linear não viesado genômico de passo único (ssGBLUP) para as características de circunferência escrotal aos 365 dias de idade (CE365) e 450 dias de idade (CE450) de idade, período de gestação (PG), como característica do bezerro, idade ao primeiro parto (IPP) e produtividade acumulada (PAC) e classificar cada animal segundo seu valor genético (EBV) e valor genético genômico (GEBV), estimados para as características estudadas, visando obter informações mais acuradas para a seleção de animais jovens; (2) Comparar grupos de animais para as características reprodutivas avaliadas, resultantes de análises de agrupamento com as estimativas de EBV e GEBV, a fim de identificar aqueles que apresentam melhor potencial genético para o desempenho reprodutivo. Foram utilizadas informações de pedigree, fenótipos de 34.848 animais e genótipos de 8.652 animais provenientes da Associação Nacional de Criadores e Pesquisadores (ANCP). Os grupos de treinamento e validação foram determinados de forma aleatória para validação cruzada (método “K-fold Cross Validation”) e a obtenção do EBV e GEBV foram obtidos para o grupo de validação utilizando os métodos BLUP e ssGBLUP, respectivamente. Posteriormente, a habilidade de predição dos métodos foi mensurada por meio da correlação entre o fenótipo ajustado e o GEBV ou EBV. As acurácias de predição variaram entre 0,06 a 0,41 e 0,07 a 0,46 para os métodos BLUP e ssGBLUP, respectivamente. A inclusão da informação genômica melhora a habilidade de predição de características reprodutivas. A utilização do método ssGBLUP para predição de valores genéticos para as características CE365, CE450, PAC e PG no programa de melhoramento genético da raça Nelore é indicada. No entanto, o método não apresenta a mesma eficiência para IPP. As análises de agrupamento contribuem para o direcionamento da seleção de animais jovens segundo o GEBV. Os resultados indicam que a seleção de animais superiores para as características de interesse pode ser diferente quando a informação genômica é considerada.

Palavras chave: Bovinos de corte, reprodução, habilidade de predição, agrupamento

1 INTRODUÇÃO

A eficiência reprodutiva é fundamental e economicamente relevante para produção de carne bovina (Cammack et al., 2009). O constante aumento da demanda por carne bovina exige esforços para suprir esse mercado e o melhoramento genético animal, por meio da identificação de animais com superioridade genética e de

combinação de genótipos superiores (Resende et al., 2012), permite selecionar animais mais precoces e com maior permanência produtiva no rebanho. No entanto, a seleção para características reprodutivas possui resposta mais lenta quando comparado à seleção para características produtivas, isso ocorre, pois, características reprodutivas são de mais baixa herdabilidade (Muir, 2007). Assim, alternativas que melhorem a resposta à seleção e conseqüentemente aceleram o ganho genético são necessárias para melhorar o desempenho reprodutivo dos rebanhos bovinos de corte.

Com a finalidade de acelerar o ganho genético em direção a determinado objetivo de interesse econômico, tecnologias como a dos marcadores moleculares e biotecnologias de reprodução vêm sendo usadas. O foco do uso de marcadores tipo polimorfismos de nucleotídeo único (SNP) é aumentar a velocidade da mudança genética por meio do aumento da precisão das predições e da variação genética capturada ou pela diminuição do intervalo de gerações (Van Eenennaam et al., 2014).

A seleção genômica (GS) incorpora a matriz de informação genômica na seleção tradicional que é baseada em registros de pedigree (Legarra et al., 2014). Como resultado, a melhora na acurácia de predição dos valores genéticos é observada (Lourenço et al., 2013; Boichard et al., 2016), permitindo que a tomada de decisão de seleção para características reprodutivas seja mais assertiva (Boddhireddy et al., 2014; Guarini et al., 2018).

As informações genômicas, quando incorporadas em programas de avaliação genética, não só aumentam a acurácia dos valores genéticos, mas principalmente permitem a seleção de animais em idades precoces e conseqüente utilização intensificada de touros jovens com superioridade genética comprovada (Carvalho, 2014). Isso acelera o ganho genético, sendo grande avanço na seleção de bovinos, em especial para características de baixa herdabilidade.

Para contribuir com a aplicação da GS em bovinos da raça Nelore, os objetivos do presente trabalho foram: (1) Comparar a habilidade de predição dos métodos Melhor preditor linear não viesado (BLUP) e Melhor preditor linear não viesado genômico de passo único (ssGBLUP) para as características de circunferência escrotal aos 365 dias de idade (CE365) e 450 dias de idade (CE450) de idade, período de gestação (PG), como característica do bezerro, idade ao primeiro parto (IPP) e produtividade acumulada (PAC) e classificar cada animal segundo seu valor genético (EBV) e valor genético genômico (GEBV), estimados para as características estudadas, visando obter informações mais acuradas para a seleção de animais

jovens; (2) Comparar grupos de animais para as características estudadas, resultantes de análises de agrupamento com as estimativas de EBV e GEBV, a fim de identificar aqueles que apresentam melhor potencial genético para o desempenho reprodutivo.

2 MATERIAL E MÉTODOS

2.1 Conjunto de dados

O conjunto de dados utilizado é constituído por pedigree, fenótipos e genótipos de animais da raça Nelore, provenientes da Associação Nacional de Criadores e Pesquisadores (ANCP). O programa de melhoramento genético animal da ANCP tem como objetivo identificar, selecionar e oferecer reprodutores e matrizes geneticamente superiores para características produtivas e reprodutivas ao mercado. O pedigree utilizado é composto por registros de 140.199 animais da raça Nelore nascidos entre os anos de 1.934 e 2.016, com 23 gerações máximas e tamanho efetivo da população de 403,84 (Programas CFC e ENGOG). O conjunto total de dados fenotípicos inicial era composto por observações de 34.848 animais para as características de CE365, CE450, PG, IPP e PAC em bovinos da raça Nelore.

O índice PAC foi obtido utilizando a seguinte equação (Grossi et al., 2008):

$$PAC = \frac{P_{210} \times n_c \times 365}{IVP_n - 550}$$

onde, P_{210} é o peso corporal médio dos bezerros desmamados corrigido para 210 dias de idade, n_c é o número total de bezerros produzidos e IVP_n é a idade da vaca ao último parto. As constantes 365 e 550 permitem a expressão da característica anualmente.

Os animais foram criados em regime de pastagem, desmamados entre 6 e 8 meses de idade e o manejo reprodutivo consistia no uso de estação de monta de 60 a 120 dias utilizando inseminação artificial ou monta natural controlada. Os efeitos considerados para definição dos grupos de contemporâneos (GC) foram o sexo (quando a característica é medida em ambos os sexos), ano e trimestre de nascimento, fazenda de origem e grupo de manejo.

O conjunto de dados genotípicos é composto por genótipos de 8.652 animais (machos e fêmeas). Desses animais, um total de 960 touros foram genotipados com painel de alta densidade (HD) (Illumina Bovine HD BeadChip), 1.000 animais foram

genotipados com o painel de média densidade (50k Illumina BovineSNP50 BeadChip) e 6.692 animais foram genotipados com painel de baixa densidade (LD) (2k Clarifide Nelore 2.0). Os animais genotipados com o painel LD e 50K foram imputados para o painel HD utilizando o software FImpute v.2.2 (Sargolzaei et al., 2014). Na imputação os touros genotipados em HD foram usados como população de referência, sendo que a acurácia de imputação foi maior que 0,98.

2.2 Controle de Qualidade

Na edição dos dados fenotípicos, os grupos de contemporâneos que apresentaram frequência menor que cinco animais foram excluídos. A análise de conectabilidade entre os grupos de contemporâneos para cada característica foi realizada utilizando o software AMC (Roso e Schenkel, 2006), utilizando o padrão do software de 10 laços, para identificar grupos desconexos, os quais foram desconsiderados das análises. As pressuposições de normalidade dos resíduos e homogeneidade da variância foram realizadas e as observações de resíduo de todo o conjunto de dados que estavam acima ou abaixo de 3,0 desvios-padrão foram excluídas. O número de observações para cada característica após a edição dos dados está apresentado na Tabela 1.

Tabela 1. Número de observações (n), média e respectivo desvio-padrão (SD), mínimo, máximo e número de grupo de contemporâneos (GC) para circunferência escrotal aos 365 dias de idade (CE365) e 450 dias de idade (CE450), período de gestação (PG), como característica do bezerro, idade ao primeiro parto (IPP) e produtividade acumulada (PAC)

Característica	n	Média	SD	Mínimo	Máximo	GC
CE365(cm)	15.054	21,56	2,40	14,40	29,10	426
CE450(cm)	13.694	25,01	3,23	16,00	35,40	381
PG (dias)	16.252	296,80	6,07	278,00	314,00	202
IPP (meses)	15.384	35,24	5,91	21,00	49,00	237
PAC (kg/ano)	6.205	150,70	31,12	63,00	249,00	150

Os genótipos passaram por controle de qualidade excluindo os SNP que apresentaram taxa de leitura menor que 90% e com frequência do alelo raro ("Minor Allele Frequency" - MAF) abaixo de 5%. Informações em cromossomos sexuais não estavam disponíveis. No controle de qualidade das amostras foram excluídas as que

tiveram taxa de leitura menor que 90%. Após o controle de qualidade dos dados genômicos, restaram dados de 460.838 SNP e de 8.545 animais.

2.3 Seleção Genômica Ampla (GS)

Os métodos BLUP e ssGBLUP foram usados para verificar a habilidade de predição e calcular o valor genético dos animais. Os grupos de treinamento e validação para ambas análises foram determinados para validação cruzada, mascarando os fenótipos de todos os animais do grupo de validação e predizendo essa informação com o modelo derivado com os indivíduos do grupo de treinamento. Para isso, os animais foram divididos em três grupos definidos aleatoriamente, sendo uma primeira análise realizada considerando um grupo na validação e os demais no treinamento e, seguindo desta forma até que todos os grupos tivessem sido considerados como população de validação (método “K-fold Cross Validation”). O número de animais para a composição das populações de treinamento e validação e a média de parentesco de cada animal presente no grupo com os 86 animais mais aparentados presentes no mesmo grupo está descrito na Tabela 2.

Tabela 2. Número de animais compondo as populações de treinamento e validação

Característica	Treinamento		Validação		Parentesco
	N	G	N	G	Treinamento x Validação
C365	13.881	7.479	1.173	1.173	0,152
C455	12.651	7.609	1.043	1.043	0,155
PG	14.881	7.281	1.371	1.371	0,144
IPP	12.986	6.254	2.398	2.398	0,150
PAC	5.482	7.929	723	723	0,147

CE365: Circunferência escrotal aos 365 dias de idade; CE450: Circunferência escrotal aos 450 dias de idade; IPP: Idade ao primeiro parto; PAC: Produtividade acumulada; PG: Período de gestação, como característica do bezerro; Parentesco: Média de parentesco de cada animal com os 86 animais mais aparentados, considerando animais treinamento e validação; N: Número total de animais com fenótipo; G: Número total de animais com genótipo

O primeiro método utilizado foi o BLUP, o qual utiliza informações de pedigree e fenotípicas e não inclui informações genômicas. Os modelos animais utilizados para CE365, CE450, IPP e PAC (1) e para PG (2) foram:

$$(1) y_t = Xb + Z_1u + e$$

$$(2) y_t = Xb + Z_1u + Z_2m + e$$

em que, y_t é o vetor de fenótipos para cada característica, b é o vetor de efeitos fixos (GC), u é o vetor de efeitos genéticos aditivos; m é o vetor de efeito genético materno

e é o vetor de efeitos aleatórios residuais e X , Z_1 e Z_2 são as matrizes de incidência para b , u e m , respectivamente. Considerando um modelo infinitesimal, as variâncias de u , m e e são:

$$\begin{bmatrix} u \\ m \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_u^2 & A\sigma_{um} & 0 \\ A\sigma_{um} & A\sigma_m^2 & 0 \\ 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

em que σ_u^2 é a variância genética aditiva; σ_m^2 é a variância genética aditiva materna; σ_e^2 é a variância residual; σ_{um} é a covariância aditiva maternal e A é o numerador da matriz de parentesco obtida a partir das informações de pedigree e I é a matriz identidade.

O segundo método utilizado foi o ssGBLUP, em que todas as informações de pedigree, fenotípicas e genotípicas são combinadas em uma única análise (Aguilar et al., 2010; Christensen e Lund, 2010). Os modelos animais foram os mesmos utilizados para BLUP tradicional. O método ssGBLUP foi implementado substituindo a matriz A , utilizada no BLUP, pela matriz H . A inversa da matriz H foi utilizada de acordo com Legarra et al. (2009) e Aguilar et al. (2011):

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

em que, H^{-1} é a inversa da matriz de parentesco incorporada pela inversa da matriz de parentesco genômica (G^{-1}) e pela inversa do numerador da matriz de parentesco dos animais genotipados (A_{22}^{-1}). A matriz G foi construída de acordo com VanRaden (2008):

$$G = \frac{(M - P)(M - P)'}{2 \sum_{j=1}^m p_j (1 - p_j)}$$

em que, M é a matriz de alelos marcadores com m colunas (m = número total de marcadores) e n linhas (n = número total de indivíduos genotipados), e P é a matriz contendo a frequência do segundo alelo (p_j), expresso como $2p_j$. As análises foram realizadas com os softwares da família BLUPF90 (Misztal et al., 2002; Aguilar et al., 2011).

A habilidade de predição dos modelos para cada conjunto de dados foi mensurada por meio da correlação de Pearson entre o valor genético estimado (EBV ou GEBV) e o fenótipo corrigido para efeitos fixos, utilizando a população de validação. Para verificar a qualidade do ajuste do modelo aplicado foi utilizada a medida do quadrado médio do erro de predição (MSE - "Mean Squared Error"), sendo que o

melhor ajuste do modelo é identificado pelo menor valor de MSE. O viés das predições genômicas foi avaliado utilizando o seguinte modelo de regressão:

$$y = b_0 + b_1x + e$$

em que y é o fenótipo ajustado, b_0 é o intercepto, b_1 é o coeficiente de regressão linear, x é o valor genético estimado (EBV ou GEBV) e e o resíduo. As estimativas de herdabilidade, utilizando a matriz genômica, foram 0,33, 0,42, 0,42, 0,14, e 0,18 para CE365, CE450, PG, IPP e PAC, respectivamente.

Além disso, os EBVs e GEBVs a partir do pedigree foram ordenados de acordo com os valores mais favoráveis para cada característica. Os 20 primeiros animais para EBV e GEBV desse ranking, foram comparados, primeiro, observando a porcentagem de animais coincidentes em ambos os grupos e posteriormente, pela correlação de Spearman, calculada entre a identificação dos 20 primeiros animais do pedigree que tinham os maiores EBVs e GEBVs a fim de identificar se as duas metodologias ordenaram os animais na mesma posição.

2.3.1 Análise de Agrupamento hierárquica e não-hierárquica

Nas análises de agrupamento foram utilizados os EBVs e GEBVs com acurácia acima de 0,2 obtidos pelos métodos BLUP e ssGBLUP, respectivamente, em que informações fenotípicas, genotípicas e de pedigree de toda a população foram utilizadas, sem a separação de população treinamento e validação. Os resultados destas análises foram utilizados para agrupar os animais com base nos valores genéticos para as características estudadas. Por meio da análise de agrupamento hierárquico, o número de grupos em que a população foi dividida foi determinado. A medida de similaridade entre os animais foi a distância Euclidiana e o algoritmo de agrupamento para formar os grupos foi o de Ward (1963).

Após definido o número de agrupamentos, a análise não-hierárquica pelo método k-means foi realizada a fim de caracterizar o agrupamento dos animais para as características reprodutivas estudadas. Os grupos de animais que melhor atenderiam determinados objetivos de seleção foram determinados. As análises de agrupamento foram executadas por programação em R (R Core Team, 2020).

3 RESULTADOS E DISCUSSÃO

As diferentes medidas de habilidade de predição média dos grupos de validação cruzada utilizando os métodos BLUP e ssGBLUP estão apresentados na Tabela 3.

De maneira geral ao avaliar o quadrado médio do erro de predição (MSE), nota-se que o método genômico é o que possui maior ajuste do modelo aos dados. Contudo, para a característica IPP, um valor levemente menor de MSE foi estimado pelo método BLUP.

Tabela 3. Acurácia de predição medida pela correlação de Pearson (r) entre o fenótipo corrigido e o EBV ou GEBV, quadrado médio do erro de predição (MSE) e coeficiente de regressão (b_1) do fenótipo corrigido sobre o EBV ou GEBV, de acordo com os métodos BLUP e ssGBLUP

Características	BLUP		ssGBLUP		
	Média	SD	Média	SD	
r	CE365	0,359	0,060	0,383	0,044
	CE450	0,409	0,029	0,463	0,032
	PG	0,237	0,031	0,268	0,051
	IPP	0,147	0,082	0,125	0,049
	PAC	0,062	0,040	0,068	0,014
MSE	CE365	3,630	0,190	3,526	0,146
	CE450	6,296	0,631	5,884	0,570
	PG	33,172	0,466	32,108	0,589
	IPP	14,618	1,555	14,696	1,582
	PAC	708,945	94,203	707,740	97,394
b_1	CE365	1,272	0,224	1,191	0,145
	CE450	1,295	0,131	1,256	0,072
	PG	0,859	0,119	0,866	0,177
	IPP	1,056	0,647	0,750	0,337
	PAC	0,490	0,307	0,466	0,093

CE365: Circunferência escrotal aos 365 dias de idade; CE450: Circunferência escrotal aos 450 dias de idade; IPP: Idade ao primeiro parto; PAC: Produtividade acumulada; PG: Período de gestação, como característica do bezerro

As acurácias de predição variaram de 0,06 a 0,41 e de 0,07 a 0,46 para os métodos BLUP e ssGBLUP, respectivamente. Boddhireddy et al. (2014) encontraram acurácias de predição de 0,59, 0,59, 0,50, 0,64 e 0,54 para CE365, CE450, PG, IPP e PAC, respectivamente, utilizando modelo bayesiano (BayesC). No entanto, esses autores definiram grupos de validação cruzada baseados em métodos de agrupamento de idênticos por descendência e a habilidade de predição foi obtida pela correlação do GEBV com EBV.

O estudo desenvolvido por Costa et al. (2019) com características reprodutivas de fêmeas da raça Nelore reportou acurácias entre 0,22 e 0,24; entre 0,27 e 0,30 e entre 0,52 e 0,55 para reconcepção de novilhas, IPP e ocorrência de parto precoce,

respectivamente. Os mesmos autores utilizaram a correlação entre GEBV e o fenótipo corrigido para habilidade de predição, sendo o mesmo indicador de habilidade de predição do presente estudo. Ainda, os métodos utilizados por Costa et al. (2019) foram GBLUP e bayesianos (IBLASSO e BayesC π), concluindo que apesar dos métodos bayesianos resultarem em maior acurácia de predição quando comparados ao GBLUP, os métodos não diferem significativamente entre si. Em contradição a esta conclusão, Silva et al. (2016) avaliando os métodos BLUP, GBLUP, ssGBLUP e BayesC π para características de eficiência alimentar com estimativas de herdabilidade variando entre 0,11 e 0,45, em bovinos Nelore, obtiveram predições mais acuradas e consistentes pelo método ssGBLUP, sendo que a habilidade de predição foi obtida com a mesma equação do presente estudo.

Alguns fatores influenciam a acurácia da predição genômica como a magnitude da herdabilidade, o número de animais na população de referência e a relação entre o tamanho do genoma e o nível de endogamia, que pode estar relacionado a formação de haplotipos (herança de blocos) (Makowsky, 2011; Oldenbroek e Waaij, 2015). As metodologias de predição utilizadas e a abordagem utilizada para definição dos grupos de treino e validação também podem interferir as acurácias de predição. Três abordagens de validação cruzada foram utilizadas por Silva et al. (2016), a primeira considerando animais mais jovens para validação, a segunda considerando subconjuntos de animais menos relacionados para validação, e a terceira definindo os conjuntos de forma aleatória. A última abordagem citada, forneceu as predições mais acuradas. Pérez-Cabal et al. (2012) em seu estudo de predição genômica em bovinos de leite, também encontraram acurácias mais altas quando utilizaram a estratégia de validação cruzada aleatória para dividir os grupos de treino e validação de forma aleatória.

De acordo com Costa et al. (2019) a habilidade de predição é proporcional a magnitude da herdabilidade da característica. As características CE450 e CE365, que apresentaram acurácias de predição maiores por ambos os métodos (0,41 e 0,46 e 0,36 e 0,38, BLUP e ssGBLUP, respectivamente) foram as características com maiores estimativas de herdabilidade, sendo iguais a 0,42 e 0,33, respectivamente. De fato, estudos envolvendo outras características relataram maiores acurácias de predição para aquelas com maiores estimativas de herdabilidade (*i.e.* Silva et al., 2016).

Para características produtivas com estimativas de herdabilidade entre 0,20 e 0,41, as acurácias de predição dos métodos BLUP e ssGBLUP variaram entre 0,23 e 0,34 e entre 0,27 e 0,39, respectivamente (Lourenço et al., 2015). Para a característica reprodutiva de facilidade de parto que possui herdabilidade menor (0,12), os mesmos autores relataram habilidade de predição de 0,12 e de 0,13 dos métodos BLUP e ssGBLUP, respectivamente. O aumento da magnitude da herdabilidade de características alvo resulta em aumento da acurácia de predição.

Segundo Calus et al. (2013), características com menores estimativas de herdabilidade requerem uma população de referência maior. As características reprodutivas comumente apresentam estimativas de herdabilidade entre 0,1 e 0,5 (Grossi et al, 2008; Moreira et al, 2015) e de acordo com Oldenbroek e Waaij (2015), para duas características com herdabilidade alta e baixa, respectivamente, atingirem o mesmo nível de acurácia é necessário um grande aumento no tamanho da população de referência. Assim, a população referência pode estar influenciando nos valores de acurácia de predição obtidos no presente estudo, uma vez que as populações treinamento das características estudadas são compostas por uma pequena parcela de animais genotipados, sendo o restante constituído por animais que apresentam apenas informações de fenótipo e pedigree.

A baixa habilidade de predição observada para PAC por ambos os métodos (0,06 e 0,07, para BLUP e ssGBLUP, respectivamente) também pode ter influência do número de informações genóticas na população referência, uma vez que PAC é a característica que apresenta o menor número de informações. O tamanho da população de referência é um fator limitante para capturar a acurácia do GEBV, estudo desenvolvido por Erbe et al. (2010), sugere que na seleção genômica com população de validação menor que 300 animais as estimativas começam apresentar alto erro associado.

Outro fator influenciador sobre a acurácia de predição da PAC é a magnitude da herdabilidade (0,18) que é um fator restritivo para acurácia da seleção genômica (Oldenbroek e Waaij, 2015). Além disso, a PAC resulta de um índice, que depende das características IPP, habilidade de permanência no rebanho e peso dos bezerras desmamados (Lôbo et al., 2000). A IPP apresentou estimativa de herdabilidade de 0,14 no presente estudo e a estimativa da herdabilidade da característica habilidade de permanência de animais pertencentes ao mesmo banco de dados é de 0,29. Portanto, a acurácia de predição para PAC pode estar sendo limitada pelos resultados

da IPP e pelo baixo número de informações genômicas do presente estudo. Embora as acurácias de predição para PAC apresentadas no presente trabalho são baixas (Tabela 3), os resultados são condizentes com estudos relatados pela literatura, em que o método de predição ssGBLUP obtém maior habilidade de predição (0,63-0,81) quando comparado ao método BLUP (0,58-0,75) (Onogi et al., 2014).

Os resultados de acurácia de predição da análise BLUP para IPP (0,15 – Tabela 3) foram maiores do que a acurácia obtida pelo modelo ssGBLUP (0,12 – Tabela 3). Costa et al., 2014 relataram maiores acurácias de predição para IPP, variando de 0,27 (GBLUP) a 0,30 (Bayes LASSO e BayesC π), porém a estimativa de herdabilidade encontrada pelos autores foi de 0,20, também um pouco maior do que a relatada no presente estudo (0,14). Os autores ressaltam que as características estudadas (IPP, reconcepção de novilhas e ocorrência de prenhez precoce) em fêmeas Nelore, apresentam herdabilidades de baixa a moderada, e assim, são fortemente afetadas pelo componente ambiental.

Irano et al. (2013), usando o método GBLUP encontraram baixa acurácia de predição da seleção genômica (0,20) para IPP utilizando um número de animais genotipados 8 vezes menor que o presente estudo, porém a habilidade de predição foi obtida pela correlação entre os EBV e GEBV, que pode interferir na maior acurácia encontrada quando comparada ao presente estudo. Ademais, os autores sugerem a realização de estudos utilizando outros métodos, além de relatar a necessidade de maior número de animais genotipados para predições genômicas mais acuradas.

Mota et al. (2018) relataram baixa eficiência da GS para IPP estudando uma população de fêmeas Nelore por meio de métodos bayesianos, os resultados foram justificados pelo pequeno número da amostra utilizada (714 animais) e pelas baixas estimativas de herdabilidade (0,11 a 0,25). Embora a amostra utilizada no presente estudo seja maior em relação aos estudos mencionados anteriormente, poucos animais possuem progênes, o que poderia afetar os resultados.

Em bovinos, a seleção genômica tem proporcionado ganhos na acurácia de predição quando o método ssGBLUP é utilizado (Lourenço et al., 2013; Lourenço et al., 2015; Silva et al., 2016; Costa et al., 2019). De maneira geral, resultados semelhantes foram observados para o presente estudo, em que, com exceção da característica IPP, que apresentou maior acurácia de predição pelo método BLUP, as acurácias de predição obtidas pelo método ssGBLUP foram maiores do que as acurácias obtidas pelo método BLUP.

A correlação de Spearman entre os 20 primeiros animais do pedigree com maiores estimativas de GEBV e EBV não foi significativa ($p > 0,05$) para as características estudadas, ou seja, as metodologias BLUP e ssGBLUP apresentam diferenças quanto a ordenação dos animais segundo seus EBVs e GEBVs, respectivamente. Ademais, a proporção de animais que coincidiram nas 20 primeiras posições, que representam animais com maiores GEBVs e EBVs, conforme o método utilizado para predição, foi de 62%, 68%, 77%, 87% e 97% para PAC, PG, IPP, CE450 e CE365, respectivamente. Esses resultados indicam, que a seleção de animais superiores para as características de interesse pode ser diferente quando a informação genômica é considerada.

A predição genômica acrescenta a informação de SNP a análise tradicional, que considera apenas informações de fenótipo das características de interesse e registros de pedigree (Legarra et al., 2014). Desta forma, a análise genômica permite capturar uma maior variância quando comparado a análise tradicional (Miller, 2010; Lôbo et al., 2011; Ryu and Lee, 2014) e ainda, possibilita prever fenótipos precisos em idades precoces (Makowsky et al., 2011). Desta maneira, o presente estudo indica que a inclusão da informação de marcadores do tipo SNP para predição genômica tem potencial na melhoria da habilidade de predição para características reprodutivas de bovinos da raça Nelore. No entanto, acreditamos que aumentar a população de referência e utilizar genótipos de animais mais representativos poderia modificar os resultados, obtendo melhores acurácias do GEBV.

De acordo com o resultado da análise hierárquica (Apêndice G), dividiram-se os animais em quatro grupos. Para as análises não-hierárquicas utilizando os GEBVs, 19.451 animais foram divididos nos grupos 1 (3.994 animais), 2 (5.457 animais), 3 (5.625 animais) e 4 (4.375 animais). As análises utilizando os EBVs continham 18.588 animais, divididos nos grupos 1 (4.118 animais), 2 (4.912 animais), 3 (5.650 animais) e 4 (3.908 animais).

De maneira geral, o perfil dos grupos foi semelhante, sendo que na análise não-hierárquica, os animais do grupo 3 foram os que apresentaram GEBVs (Figura 1a) e EBVs (Figura 1b) mais favoráveis para todas as características, uma vez que apresentaram os maiores valores para CE365, CE450 e PAC e menores para IPP e PG, quando comparado aos demais grupos. Os valores centrais de agrupamento, chamados de centroides, foram mais favoráveis, ou seja, superiores, nas análises

quando os GEBVs foram utilizados, exceto para característica PG que apresentou centroide mais favorável quando o EBV foi utilizado.

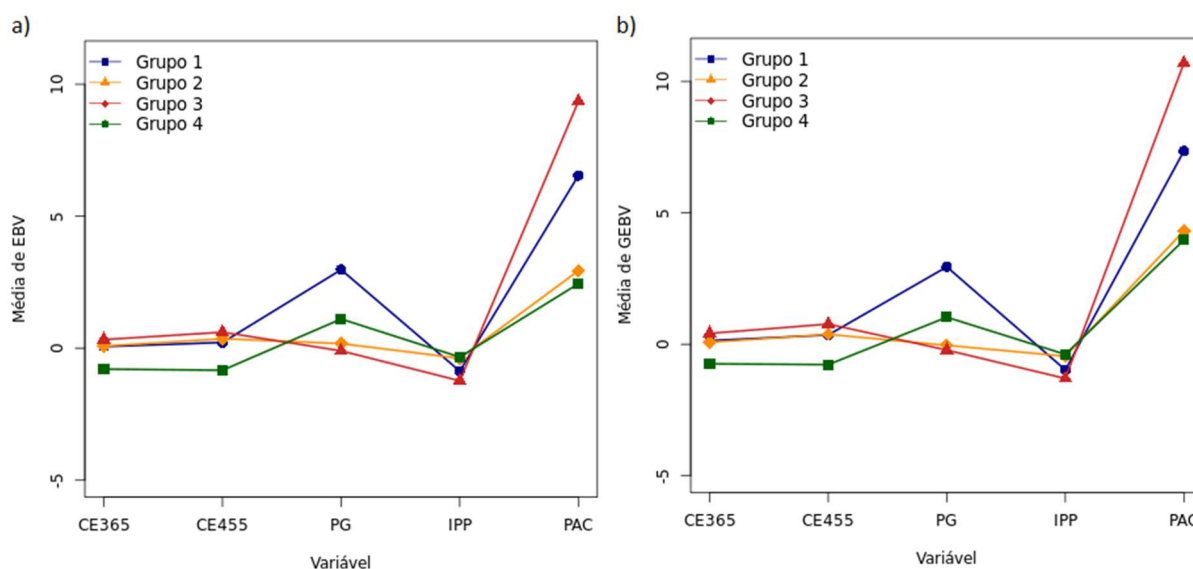


Figura 1. Análise de agrupamento não-hierárquica obtida pelo método K-means, (a) utilizando os valores genéticos (EBV) e (b) os valores genéticos genômicos (GEBV), para as características de circunferência escrotal aos 365 de idade (CE365) e aos 450 (CE450) dias de idade, período de gestação (PG), como característica do bezerro, idade ao primeiro parto (IPP) e produtividade acumulada (PAC)

Quando o grupo 3 obtido pelo GEBV foi comparado ao grupo 3 obtido pelo EBV, diferenças em relação aos animais que compõe cada grupo foram observadas. Nas análises de agrupamentos não-hierárquicas, de um total de 5.650 e 5.625 animais pertencentes ao grupo 3 quando considerados EBV e GEBV, respectivamente, 4.892 animais pertencem simultaneamente ao grupo 3 (GEBV e EBV). Assim, é observada diferença nos valores genéticos dos animais quando a informação genômica é incorporada as análises, sendo que, se todos os animais do grupo 3 obtido com o GEBV fossem selecionados, 86,97% dos animais com EBV também seriam selecionados.

As análises de agrupamento não hierárquicas são alternativas úteis na seleção de bovinos de corte para características produtivas e reprodutivas, podendo direcionar um ganho genético equilibrado e favorável (Paz et al., 2014), no entanto não substituem os índices de seleção. Grupioni e al. (2016) utilizaram análises não-hierárquicas para identificar animais com padrão genético interessante, definido por valores genéticos favoráveis para características de desempenho reprodutivo de

bovinos de corte. Ainda, padrões da relação entre características reprodutivas (IPP, PG, eficiência reprodutiva e intervalo de parto) em bovinos de leite foram evidenciados por análises de agrupamento não hierárquicas (Mello et al., 2020).

4 CONCLUSÃO

A inclusão da informação genômica melhora a habilidade de predição de características reprodutivas. A utilização do ssGBLUP para a prática da obtenção de valores genéticos para as características CE365, CE450, PAC, IPP e PG no programa de melhoramento genético da raça Nelore é indicada. Apesar da acurácia de predição para IPP ser maior pelo método BLUP, os métodos de predição genômica, como o ssGBLUP possibilita a seleção de animais jovens para essa característica, o que representa um importante ganho genético. As análises de agrupamento possibilitam a obtenção de ganho genético equilibrado para as características estudadas.

5 REFERÊNCIAS

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** 93:743-752.

Aguilar I, Misztal I, Legarra A, Tsuruta S, Davis R, (2011) Efficient computations of genomic relationship matrix and other matrices used in the single-step evaluation. **Journal of Animal Breeding and Genetics** 128:422-8.

Bodhireddy P, Prayaga K, Barros P, Lôbo R, and Denise S, (2014) Genomic Predictions of Economically Important Traits in Nelore Cattle of Brazil. In: 10th World Congress of Genetics Applied to Livestock Production. **Presentations...** Vancouver.

Boichard D, Ducrocq V, Croiseau P, Fritz S, (2016) Genomic Selection In Domestic Animals: Principles, Applications And Perspectives. **Comptes Rendus Biologies** 339:274-277.

Calus MPL, Berry DP, Banos G, Haas Y, Veerkamp RF, (2013) Genomic selection: the option for new robustness traits? **Advances in Animal Biosciences** 3:618–625.

Cammack KM, Thomas MG, Enns RM, (2009) REVIEW: Reproductive Traits and Their Heritabilities in Beef Cattle. **Professional Animal Scientists** 25:517-528.

Carvalho R, (2014) Genomic Selection in Nelore Cattle in Brazil. In: 10th World Congress of Genetics Applied to Livestock Production. **Presentations...** Vancouver.

Christensen OF, Lund MS, (2010) Genomic prediction when some animals are not genotyped. **Genetics Selection Evolution** 42:2.

Costa RB, Irano N, Diaz IDPS, Takada L, Carneiro R, Baldi F, Oliveira HN, Tonhati H, Albuquerque LG (2014) Prediction of Genomic Breeding Values for Reproductive Traits in Nelore Heifers. In: 10th World Congress of Genetics Applied to Livestock Production. **Proceedings...** Vancouver.

Costa RB, Irano N, Diaz IDPS, Takada L, et al, (2019) Prediction of genomic breeding values for reproductive traits in Nelore heifers. **Theriogenology** 1-27.

Erbe M, Pimentel ECG, Sharifi AR, and Simianer H, (2010) Assessment of Cross-validation Strategies for Genomic Prediction in Cattle. In: 9th World Congress of Genetics Applied to Livestock Production, edited by German Society for Animal Science. **German Society for Animal Science**, Leipzig, Germany, p. 129–132.

Grossi DA, Frizzas OG, Paz CCP, Bezerra LAF, Lôbo RB, Oliveira JA, Munari DP, (2008) Genetic associations between accumulated productivity, and reproductive and growth traits in Nelore cattle. **Livestock Science** 117:139–146

Gruppioni NV, Freitas LA, Stafuzza NB, Bezerra LAF, Lôbo RB, Munari DP (2006) Non-hierarchical cluster analysis for performance na reproductive traits in Guzera beef cattle. In: 1st International Meeting of Advances in Animal Science. **Resumos...** Jaboticabal-SP.

Guarini A R, Lourenco Dal, Brito LF, Sargolzaei M, Baes C F, Miglior F, Misztal I, and Schenkel FS, (2018) Comparison of genomic predictions for lowly heritable traits using multi-step and single-step genomic best linear unbiased predictor in Holstein cattle. **Jornal of Dairy Science** 101:8076–8086.

Irano N, Costa RB, Terakado, APN, Carneiro R, Baldi F, Albuquerque LG, (2013) Acurácia da seleção genômica para idade ao primeiro parto de fêmeas da raça Nelore. In: X Simpósio Brasileiro de Melhoramento Animal. **Resumos...** Uberaba: SBMA, p. 1-3.

Legarra A, Aguilar I, Misztal I, (2009) A relationship matrix including full pedigree and genomic information. **Jornal of Dairy Science** 92:4656-4663.

Legarra A, Christensen OL, Aguilar I, Misztal I, (2014) Single Step, a general approach for genomic selection. **Livestock Science** 1-8.

Lôbo R, Bezerra LAF, Oliveira HN, Garnero AV, Schwengber EB, Marcondes CR, (2000) Avaliação genética de animais jovens, touros e matrizes: **Sumário 2000**. SP, Ribeirão Preto. Brasil

Lôbo RB, Nkrumah D, Grossi DA, Barros PS, Paiva P, Bezerra LAF, Oliveira HN and Silva MVB, (2011) Implementation of DNA Markers to Produce Genomically - Enhanced EPDs in Nelore Cattle. nced EPDs in Nelore Cattle. **Acta Scientiae Veterinariae** 39:23-27.

Lourenço DAL, Misztal I, Wang H, Aguilar I, TSuruta S, and Bertrand JK, (2013) Prediction accuracy for a simulated maternally affected trait of beef cattle using different genomic evaluation models. **Journal of Animal Science** 91:4090-4098.

Lourenço DAL, Tsuruta S, Fragomeni B. O, Masuda Y, Aguilar I, Legarra A, Bertrand JK, Amen TS, Wang L, Moser DW, and Misztal I, (2015) Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. **Journal of Animal Science** 93:2653-2662.

Makowsky R, Pajewski NM, Klimentidis YC, Vazquez AI, Duarte CW, Allison DB, Campos GDL (2011), Beyond Missing Heritability: Prediction of Complex Traits. **PLOS Genetics** 7:1-9.

Mello RRC, Sinedino LDP, Ferreira JE, Sousa SLG, Mello MRB (2020) Principal component and cluster analyses of production and fertility traits in Red Sindhi dairy cattle breed in Brazil. **Tropical Animal Health and Production** 52:273-281.

Miller S, (2010) Genetic improvement of beef cattle through opportunities in genomics. **Revista Brasileira de Zootecnia** 39:247-255.

Misztal I, Tsuruta S, Strabel T, auvray B, Druet T, Lee DH, (2002) “**BLUPF90 and related programs (BGF90)**”, in Proceedings of the 7th World Congress on Genetics Applied to Livestock Production. 28:21-22.

Moreira HL, Buzanskas ME, Munari DP, Canova ÉB, Lôbo RB, Paz CCP, (2015) Reproductive traits selection in nelore beef cattle. **Ciência e Agrotecnologia** 39:4:355-362.

Mota RR, Silva FF, Guimarães SEF, et al. (2018) Benchmarking Bayesian genome enabled-prediction models for age at first calving in Nelore cows. **Livestock Science** 211:75-79.

Muir WM, (2007) Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. **Journal of Animal Breeding and Genetics** 124:342-355.

Oldenbroek K and Waaij LVD, (2015). Textbook Animal Breeding and Genetics for BSc students. **Centre for Genetic Resources The Netherlands and Animal Breeding and Genomics Centre**. Groen Kennisnet: <https://wiki.groenkennisnet.nl/display/TAB/>

Onogi A, Ogino A, Komatsu T, Shoji N, Simizu K, kurogi K, Yasumori T, Togashi K, Iwata H, (2014) Genomic prediction in Japanese Black cattle: Application of a single-step approach to beef cattle. **Journal of Animal Science** 92:1931-1938.

Paz CCP, Moreira HL, Savegnago RP, Buzanskas ME, Lôbo RB, Munari DP (2014) Non-hierarchical cluster analysis with the breeding values for reproductive and productive traits in beef cattle. In: 10th World Congress of Genetics Applied to Livestock Production. **Proceedings...** Vancouver.

Pérez-Cabal MA, Vazquez AI., Gianola D, Rosa GJM, Weigel KA, (2012). Accuracy of genome-enabled prediction in a dairy cattle population using different cross-validation layouts. **Frontiers in Genetics** 3:27

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL [https://www.R-project.org/.](https://www.R-project.org/)

Resende MDV, Silva FF, Lopes PS, Azevedo CF, (2012) Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência Bayesiana (MCMC), Regressão Aleatória Multivariada e Estatística Espacial. Viçosa: Universidade Federal de Viçosa, 291 p.

Roso VM, Schenkel FS, (2006) AMC – **A computer program to assess the degree of connectedness among contemporary groups**, in 8th World Congress on Genetics Applied to Livestock Production. Belo Horizonte, MG. V. 34.

Ryu J, and Lee C, (2014) Genomic Heritability of Bovine Growth Using a Mixed Model. **Asian-Australasian Journal of Animal Sciences**, 27:1521-1525.

Sargolzaei M, Chesnais JP and Schenkel FS (2014) A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, 15:471-478.

Silva RMO, Fragomeni BO, et al. (2016) Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population. **Journal of Animal Science** 94:3613-3623.

Van Eenennaam AL, Weigel KA, Young AE, Cleveland MA, Dekkers JCM (2014) Applied Animal Genomics: Results from the Field. **Annual Review of Animal Biosciences** 2(1):105-139.

Vanraden PM, (2008) Efficient methods to compute genomic predictions. **Journal Dairy Science** 91:4414-4423.

Ward JH, (1963) Hierarchical grouping to optimize an objective function. **Journal of the American Statistical Association** 58:236-244.

CAPÍTULO 4 - CONSIDERAÇÕES GERAIS

Estudos de Associação ampla do genoma (GWAS) viabilizam a identificação e caracterização de genes e vias biológicas importantes associadas a características reprodutivas de circunferência escrotal aos 365 dias de idade (CE365) e 450 dias de idade (CE450) de idade, período de gestação (PG), como característica do bezerro, idade ao primeiro parto (IPP), produtividade acumulada (PAC), parto precoce até 30 meses de idade (P30) e habilidade de permanência da vaca no rebanho (“Stayability” – STAY) em bovinos da raça Nelore.

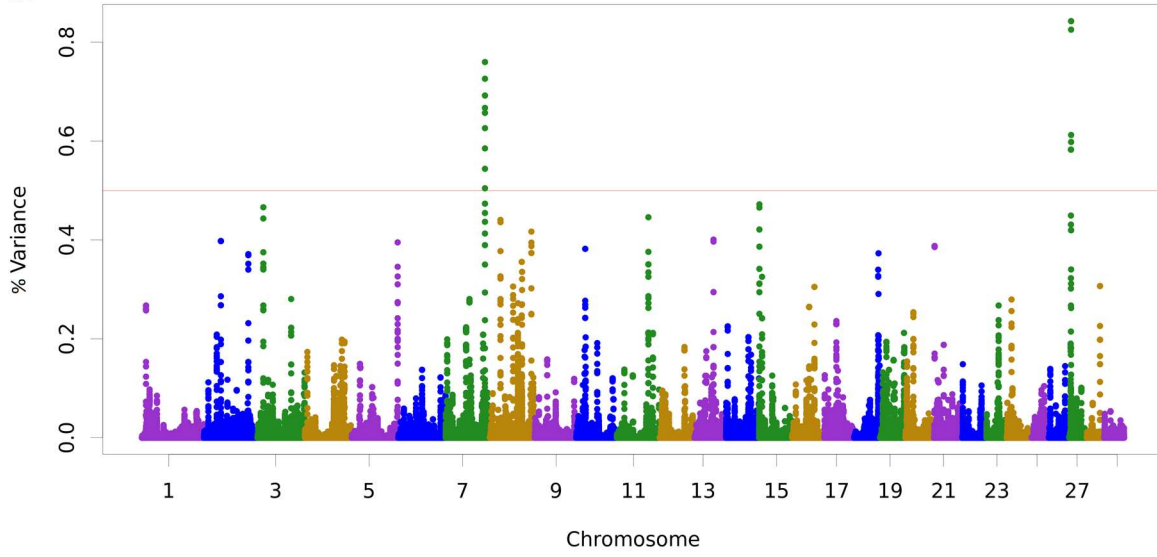
Genes candidatos com funções relacionadas a reprodução foram identificados em associação com mais de uma característica (tais como, CAMK1D, TASP1, ACOXL, RAB11FIP5 e SFXN5), indicando seu efeito pleiotrópico e evidenciando a importante função deles para eficiência reprodutiva. As vias biológicas de alfa-oxidação de ácidos graxos, de regulação negativa de adipócitos e de sinalização de esfingolipídios apresentam relação com as características de eficiência reprodutiva em bovinos da raça Nelore do presente estudo.

A inclusão da informação de marcadores do tipo SNP para predição genômica tem potencial na melhoria da habilidade de predição para características reprodutivas de bovinos da raça Nelore. Entretanto, o sucesso da GS depende do aumento do número de animais que compõe a população de referência e da definição eficiente dos animais a serem genotipados, sendo que os mesmos devem ser representativos da população.

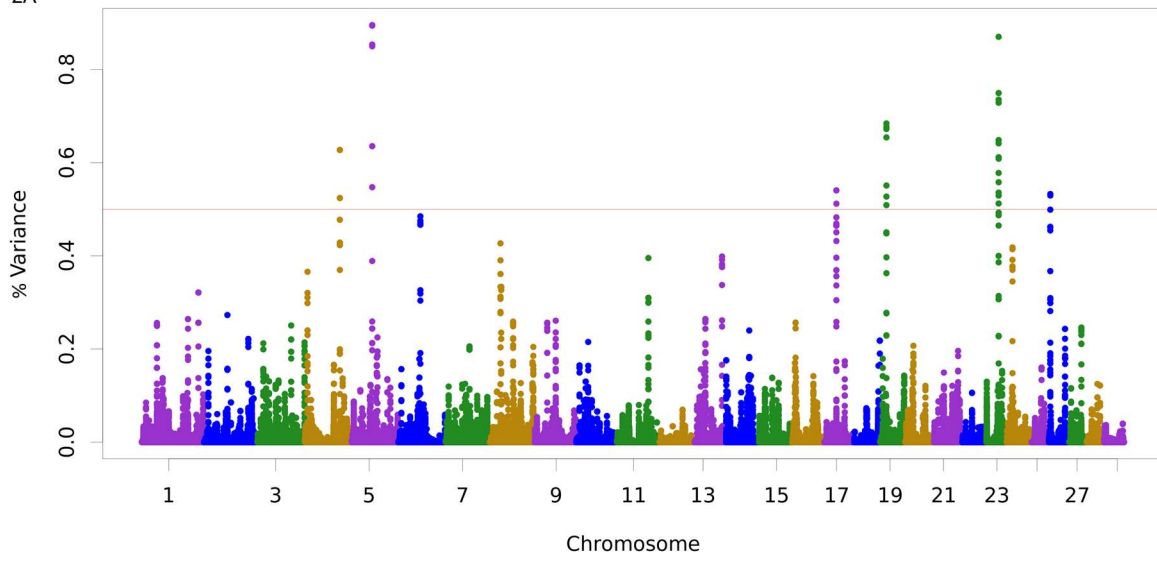
A informação genômica traz ganhos para os programas de melhoramento genético principalmente reduzindo os intervalos de gerações, uma vez que permitem a mensuração do mérito genético dos animais em idades precoces, sem a necessidade de esperar que os animais atinjam idade reprodutiva e produtiva para fazê-la. A efetivação das análises de GWAS e GS pelos programas de melhoramento animal são de grande importância no monitoramento da endogamia, prospecção de genes candidatos e identificação de animais geneticamente superiores, gerando o aumento na velocidade do progresso genético.

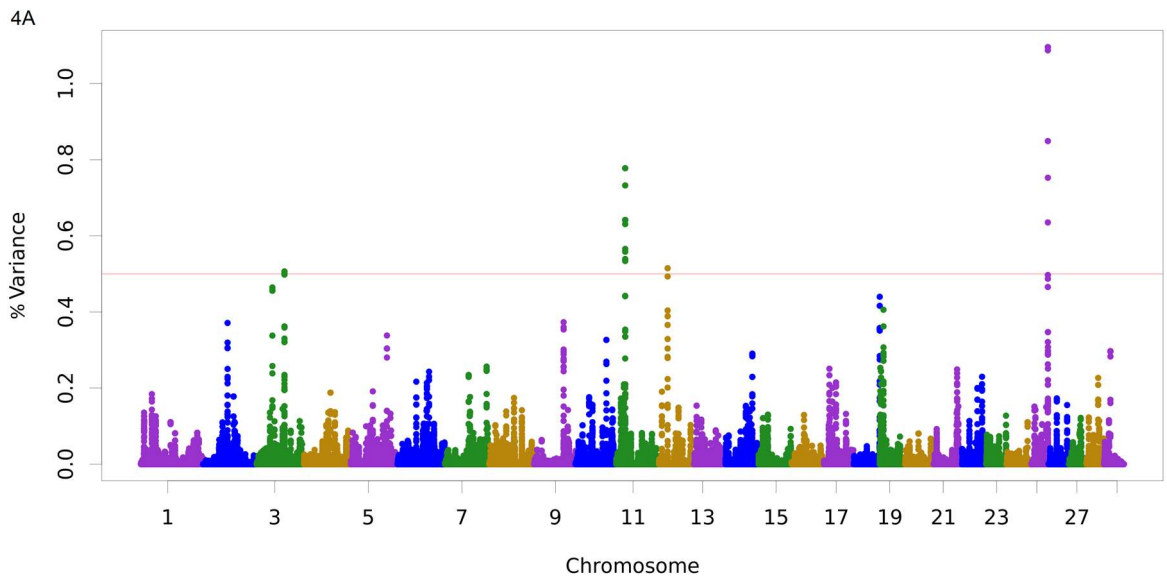
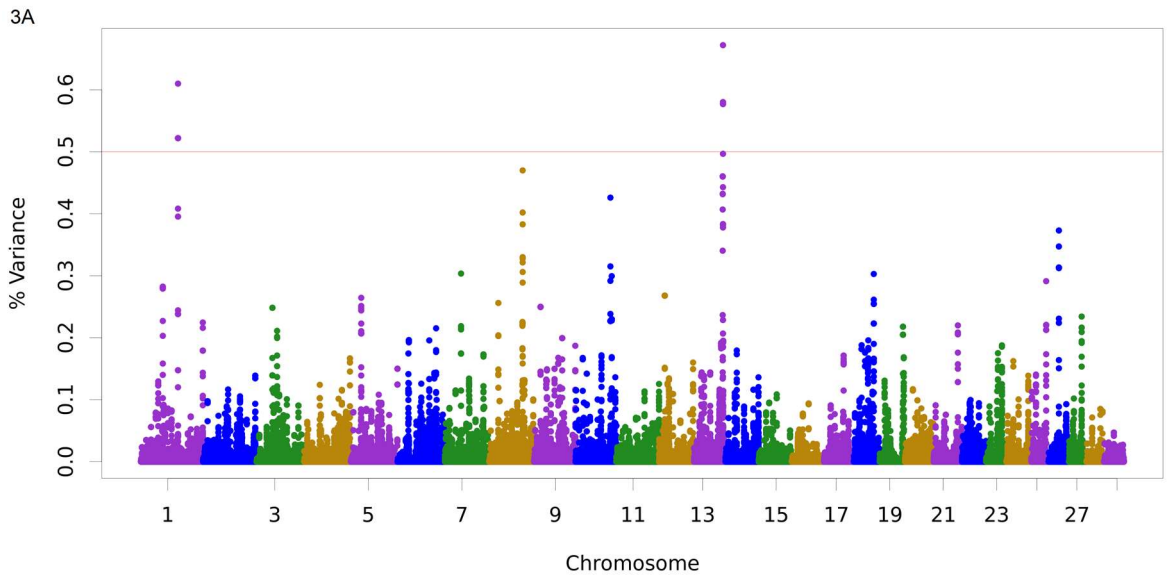
APÊNDICES

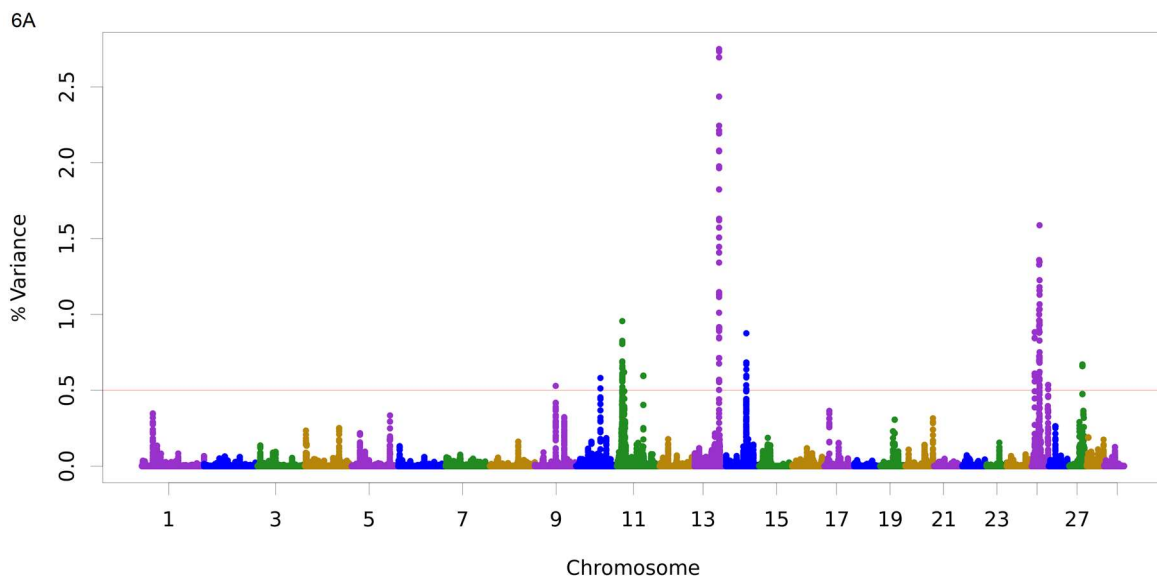
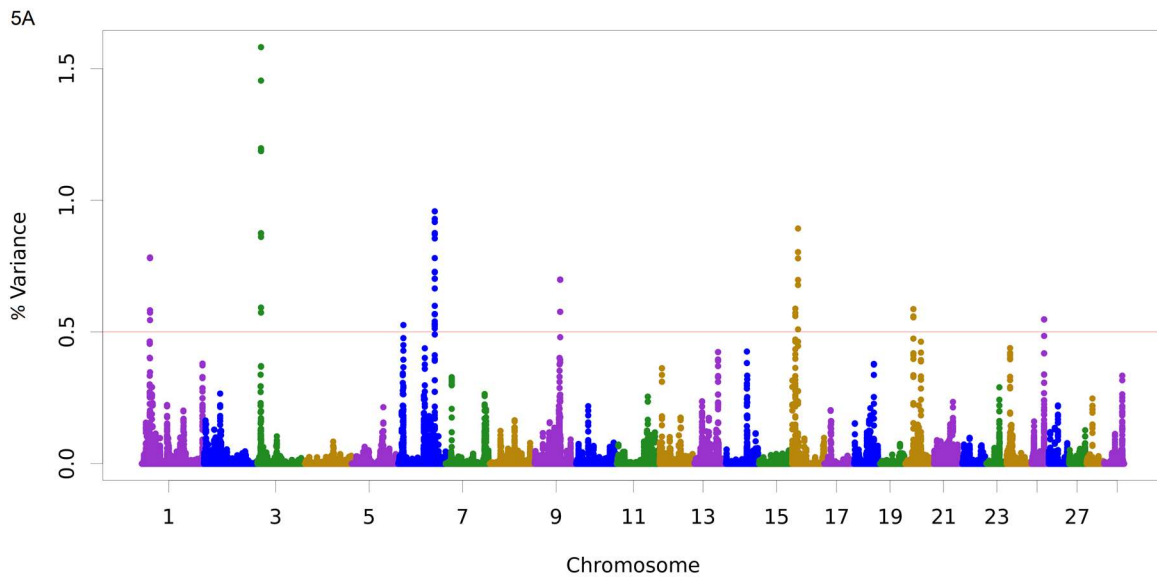
1A

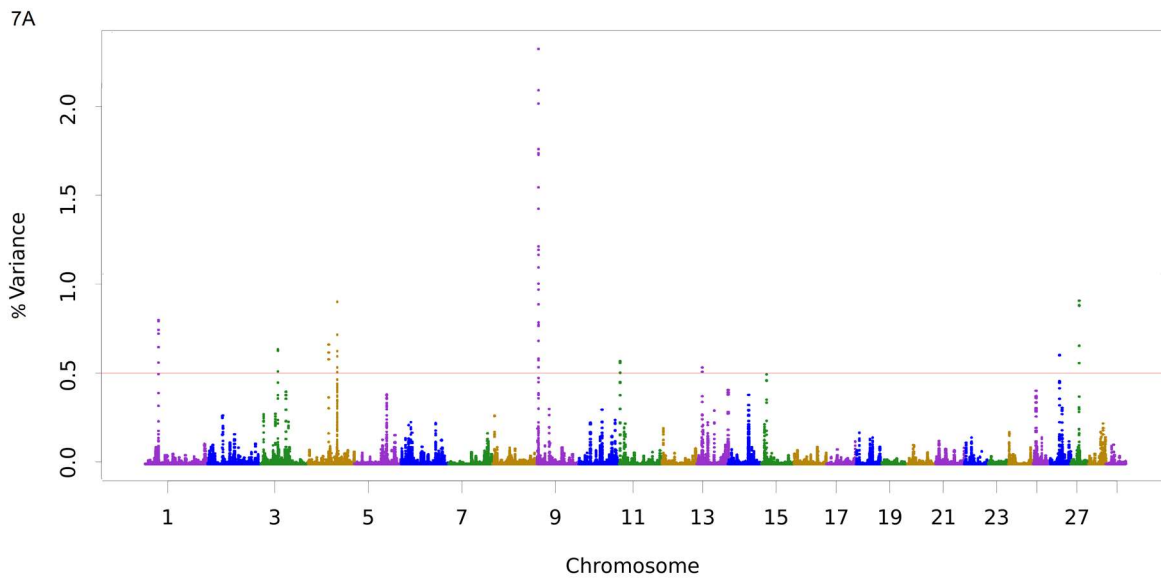


2A









Supplementary material A. Additive genetic variance explained by windows of 10 adjacent SNP distributed by chromosomes for: 1A Scrotal circumference at 365 days of age, 2A Scrotal circumference at 450 days of age, 3A Gestation length, as a calf trait, 4A Age at first calving, 5A Accumulated productivity, 6A Early calving until 30 months and 7A Stayability

Trait	
Chromosome: Position(pb)	
Gene	
Scrotal circumference at 455 days of age	
19:16748888-16777564	
ASIC2	
26:489143-4913438	
UBE2D1, CISD1, IPMK, ZWINT, ENSBTAG00000023971, ENSBTAG00000048299, ENSBTAG00000047692	
Gestation length, as a calf trait	
13:75605755-75651541	
CDH22	
Age at first calving	
11:21246-21281912	
ENSBTAG00000038257, ZC3H6, ZC3H8, FBLN7, TMEM87B, ENSBTAG00000044282, RF00401, MERTK, RF00156, ANAPC1, RF00026, ENSBTAG00000023163, BCL2L11, ACOXL, BUB1, TPC3, ENSBTAG00000045538, NPHP1, MALL, MAL, ENSBTAG00000034657, MRPS5, ZNF514, ZNF2, ENSBTAG00000047503, PROM2, ENSBTAG00000017955, FAHD2A, GPAT2, ADRA2B, ASTL, DUSP2, STARD7, TMEM127, CIAO1, SNRNP200, ENSBTAG00000047659, ITPRIPL1, NCAPH, NEURL3, ARID5A, KANSL3, FER1L5, LMAN2L, CNM4, CNM3, ANKRD23, ANKRD39, SEMA4C, FAM178B, COX5B, ACTR1B, ZAP70, TMEM131, VWA3B, CNGA3, INPP4A, COA5, UNC50, MGAT4A, RF00425, TSGA10, C2orf15, LIPT1, MITD1, MRPL30, LYG2, TXNDC9, EIF5B, REV1, ENSBTAG00000012449, LONRF2, CHST10, NMS, PDCL3, PTPN13, ENSBTAG00000016794, NPAS2, RPL31, TBC1D8, CNOT11, RF00578, RNF149, CREG2, RFX8, MAP4K4, IL1R2, RF00594, IL1R1, RF00001, ENSBTAG00000046579, IL1RL2, IL1RL1, IL18R1, IL18RAP, SLC9A4, SLC9A2, MFSD9, TMEM182, MRPS9, GPR45, TGFBRAP1, C11H2orf49, FHL2, TACR1, POLE4, HK2, ENSBTAG00000044983, SEMA4F, M1AP, DOK1, LOXL3, HTRA2, AUP1, DQX1, TLX2, PCGF1, LBX2, CCDC142, MRPL53, MOGS, WBP1, INO80B, RTKN, WDR54, C11H2orf81, MGC152281, DCTN1, SLC4A5, ENSBTAG00000010072, MTHFD2, MOB1A, BOLA3, TET3, DGUOK, ACTG2, STAMBP, C2orf78, DUSP11, bta-mir-2295, TPRKB, NAT8, ALMS1, EGR4, FBXO41, CCT7, PRADC1, SMYD5, NOTO, ENSBTAG00000008035, RF00100, RAB11FIP5, bta-mir-2294, SFXN5, EMX1, SPR, EXOC6B, ENSBTAG00000046679, ENSBTAG00000039263, CYP26B1, DYSF, ZNF638, PAIP2B, NAGK, TEX261, ANKRD53, ATP6V1B1, VAX2, CD207, CLEC4F, FIGLA, ADD2, TGFA, FAM136A, XDH, SRD5A2, MEMO1, ENSBTAG00000033887, DPY30, SPAST, SLC30A6, NLRC4, YIPF4, BIRC6, TTC27, ENSBTAG00000037357, LTBP1, RASGRP3, ENSBTAG00000045057, FAM98A, RF00066, RF00015, CRIM1, FEZ2, VIT, STRN, HEATR5B, GPATCH11, EIF2AK2, SULT6B1, CEBPZ, NDUFAF7, PRKD3, QPCT, CDC42EP3, RMDN2, CYP1B1, ATL2, RF01161, HNRNPLL, ENSBTAG00000044545, ENSBTAG00000026809, GALM, SRSF7, ENSBTAG00000012984, GEMIN6, DHX57, MORN2, ARHGEF33	
11:21287816-21311113	
ARHGEF33	

continuação...

Accumulated productivity

1:2054452-20570066

PAXBP1, SYNJ1, SYNJ1, CFAP298, EVA1C, URB1, MRAP, MIS18A, HUNK, SCAF4, SOD1, TIAM1, TIAM1, bta-mir-2284i, bta-mir-2284x, ENSBTAG00000004530, KRTAP11-1, KRTAP7-1, KRTAP8-1, ENSBTAG00000044154, RF00026, KRTAP13-1, ENSBTAG00000037933, KRTAP27-1, KRTAP24-1, CLDN8, CLDN17, ENSBTAG00000030124, GRIK1, GRIK1, ENSBTAG00000045873, RF00026, BACH1, MAP3K7CL, CCT8, USP16, RWDD2B, RWDD2B, LTN1, ENSBTAG00000038433, N6AMT1, RF00026, ENSBTAG00000040155, RF00432, ADAMTS5, ADAMTS1, RF00026, ENSBTAG00000005742, RF00026, APP, RF00026, GABPA, ATP5PF, ATP5PF, JAM2, MRPL39, bta-mir-155, RF00026, RF00026, ENSBTAG00000038900, ENSBTAG00000048165, RF00100, NCAM2, RF00001, ENSBTAG00000039704, RF00026, RF00001, RF00001, TMPRSS15, TMPRSS15, CHODL, C1H21orf91, RF00026, ENSBTAG00000047943, BTG3, CXADR, RF00100, ENSBTAG00000040418, bta-mir-125b-2, MIRLET7C, bta-mir-99a, ENSBTAG00000047183, RF00026

3:7752656-7785494

ATF6

6:1442287-14434633

MARCH1, TMA16, TKTL2, NPY5R, NPY1R, NAF1, RF00026, BBS7, CCNA2, EXOSC9, ANXA5, RF00012, RF00026, QRFPR, NDNF, PRDM5, RF00026, RF00001, ENSBTAG00000037426, ENSBTAG00000047900, ENSBTAG00000024874, ENSBTAG00000035129, ENSBTAG00000035129, MAD2L1, ENSBTAG00000035764, ENSBTAG00000035764, ENSBTAG00000040228, ENSBTAG00000037311, PDE5A, FABP2, C6H4orf3, USP53, RF00001, MYOZ2, SYNPO2, SEC24D, METTL14, PRSS12, RF00399, ENSBTAG00000010010, NDST3, NDST3, ENSBTAG00000034844, RF00001, RF00001, RF00001, RF00001, RF00001, RF00001, RF00001, RF00001, ENSBTAG00000034849, TRAM1L1, ENSBTAG00000000119, ENSBTAG00000009280, NDST4, UGT8, ENSBTAG00000011299, ARSJ, CAMK2D, CAMK2D, ENSBTAG00000002392, ENSBTAG00000002392, ENSBTAG00000002392, LARP7, bta-mir-302b, bta-mir-302c, bta-mir-302a, bta-mir-302d, bta-mir-367, ENSBTAG00000026936, ENSBTAG00000013198, NEUROG2, ALPK1, ALPK1, ALPK1, TIFA, AP1AR

6:94344402-94375956

MRPL1

6:94388672-94426934

MRPL1

16:1706924-17085278

SOX13, ETNK2, REN, KISS1, GOLT1A, PLEKHA6, PLEKHA6, PLEKHA6, PPP1R15B, PIK3C2B, PIK3C2B, RF00334, MDM4, ENSBTAG00000026068, ENSBTAG00000034611, LRRN2, RF01169, NFASC, CNTN2, TMEM81, RBBP5, DSTYK, TMCC2, TMCC2, NUA2, ENSBTAG00000047727, KLHDC8A, LEMD1, bta-mir-135b, CDK18, MFSD4A, ELK4, ELK4, SLC45A3, NUCKS1, RAB29, SLC41A1, PM20D1, SLC26A9, RAB7B, RHEX, AVPR1B, FAM72A, SRGAP2, IKBKE, RASSF5, EIF2D, DYRK3, MAPKAPK2, IL10, IL19, IL19, IL20, PIGR, PIGR, PIGR, FCAMR, FCMR, IL24, ENSBTAG00000046946, C16H1orf116, YOD1, PFKFB2, PFKFB2, C4BPB, C4BPA, C4BPA, ENSBTAG00000039196, ENSBTAG00000047621, CD55, CD55, CD55, ENSBTAG00000045783, CR2, CR2, ENSBTAG00000023177, ENSBTAG00000024647, ENSBTAG00000047228, ENSBTAG00000040409, CFH, ENSBTAG00000047966, KCNT2, ENSBTAG00000044933, ENSBTAG00000025113, ENSBTAG00000020953, RF00019, ENSBTAG00000014862, RF00004, ENSBTAG00000012765, CDC73, ENSBTAG00000047703, ENSBTAG00000046060, B3GALT2, GLRX2, TROVE2, UCHL5, ENSBTAG00000044446, RGS2, RGS13, RGS1, RGS21, RGS18, RF00001, RF00156, BRINP3, ENSBTAG00000045416, RF00001, RF00026

continuação...

Heifer early calving until 30 months

11:11320434-11348287

RAB11FIP5

11:11355768-11368368

RAB11FIP5, SFXN5

11:11374317-11381414

SFXN5

11:11410105-11424583

SFXN5

11:11439886-11463529

SFXN5

11:11496633-11536209

EMX1

11:67791729-67805289

ENSBTAG00000047822

13:64212843-64261331

RF00045, ASIP, AHCY

13:6426382-64340236

SPTLC3, RF00273, ISM1, TASP1, TASP1, ESF1, NDUFAF5, SEL1L2, FLRT3, MACROD2, KIF16B, ENSBTAG00000032609, SNRPB2, OTOR, NANP, ENSBTAG00000023169, ENSBTAG00000004354, ENSBTAG00000022922, ENSBTAG00000022922, ENSBTAG00000040046, CCDC3, CAMK1D, RF00619, CDC123, NUDT5, SEC61A2, DHTKD1, UPF2, RF00012, PROSER2, ECHDC3, USP6NL, CELF2, RF00026, ENSBTAG00000045874, ENSBTAG00000048047, ENSBTAG00000048047, GATA3, TAF3, ATP5F1C, KIN, ITIH2, ENSBTAG00000037516, ITIH5, SFMBT2, RF00026, RF00026, PRKCQ, PFKFB3, RBM17, IL2RA, IL2RA, ENSBTAG0000006078, FBH1, ENSBTAG00000036222, ENSBTAG00000036222, ENSBTAG00000017012, ENSBTAG00000022931, ENSBTAG00000022931, ENSBTAG00000034145, ANKRD26, YME1L1, MASTL, ACBD5, ACBD5, ABI1, PDSS1, APBB1IP, CCNY, CREM, CREM, CUL2, PARD3, NRP1, ITGB1, ITGB1, CCDC7, ARL5B, ENSBTAG00000047430, ENSBTAG00000037969, MALRD1, ENSBTAG00000044751, RF00026, PLXDC2, NEBL, ENSBTAG00000003937, SKIDA1, MLLT10, DNAJC1, DNAJC1, ENSBTAG00000044566, ENSBTAG00000025968, BMI1, SPAG6, COMMD3, PIP4K2A, RF00020, RF00026, RF00019, ARMC3, ARMC3, MSRB2, PTF1A, ENSBTAG00000044137, ENSBTAG00000044137, OTUD1, RF00100, RF00425, ENSBTAG00000040232, KIAA1217, ARHGAP21, ENSBTAG00000044011, ENKUR, RF00026, THNSL1, GPR158, MYO3A, GAD2, GJD4, FZD8, ENSBTAG00000047440, ENSBTAG00000044876, RF00619, OPTN, MCM10, ENSBTAG00000047869, PHYH, SEPHS1, BEND7, PRPF18, FRMD4A, FAM107B, CDFN, HSPA14, ENSBTAG00000045609, RF00099, SUV39H2, DCLRE1C, MEIG1, OLAH, ACBD7, RPP38, NMT2, NMT2, FAM171A1, RF00026, ITGA8, MINDY3, PTER, C1QL3, RSU1, CUBN, TRDMT1, VIM, ST8SIA6, HACD1, STAM, RF00153, TMEM236, MRC1, SLC39A12, CACNB2, NSUN6, EPC1, ENSBTAG00000043839, KIF5B, RF00026, ARHGAP12, ZEB1, RF01233, ZNF438, ENSBTAG00000027444, ENSBTAG00000027444, ENSBTAG00000027444, ENSBTAG00000046981, ENSBTAG00000046352, JCAD, MTPAP, MAP3K8, LYZL1, BAMBI, WAC, MPP7, ARMC4, MKX, RAB18, ENSBTAG00000048079, PCSK2, PCSK2, ENSBTAG00000006043, BFSP1, DSTN, RRBP1, BANF2, SNX5, RF00567, MGME1, OVOL2, OVOL2, KAT14, ZNF133, ZNF133, DZANK1, POLR3F, RBBP9, SEC23B, DTD1, ENSBTAG00000012276, RF00001, SCP2D1, SLC24A3, ENSBTAG00000032893, RIN2, bta-mir-2305, NAA20, CRNKL1, CFAP61, INSM1, RALGAPA2, RF00156, KIZ, XRN2, NKX2-4, NKX2-2,

continuação...

ENSBTAG00000035643, PAX1, FOXA2, RF00026, THBD, CD93, NXT1, GZF1, NAPB, CSTL1, CST11, CST11, CST11, MGC133636, ENSBTAG00000027420, CST8, CST8, ENSBTAG00000015949, ENSBTAG00000047137, ENSBTAG00000032687, CST3, SYNDIG1, CST7, CST7, APMAP, ENTPD6, ACSS1, VSX1, PYGB, ABHD12, ANKRD16, ENSBTAG00000035572, ENSBTAG0000002446, GDI2, FAM208B, RF00322, ASB13, ENSBTAG00000027416, CALML5, NET1, UCN3, TUBAL3, TUBAL3, ENSBTAG00000037509, ENSBTAG00000039722, ENSBTAG00000039643, ENSBTAG00000017938, ENSBTAG00000027419, AKR1C3, AKR1C4, AKR1C4, ENSBTAG00000039289, ENSBTAG00000021368, ENSBTAG00000022570, 20ALPHA-HSD, AKR1E2, ENSBTAG00000017191, KLF6, ENSBTAG00000046128, PITRM1, PFKP, RF00402, ADARB2, ENSBTAG00000039356, ENSBTAG00000039356, ENSBTAG00000039356, ENSBTAG00000037833, WDR37, IDI1, GTPBP4, RF00026, LARP4B, DIP2C, RF00026, RF00089, ZMYND11, ZMYND11, PRNP, PRND, RASSF2, RF00026, SLC23A2, TMEM230, PCNA, CDS2, RF00377, ENSBTAG00000020668, PROKR2, GPCPD1, SHLD1, CHGB, TRMT6, MCM8, CRLS1, LRRN4, FERMT1, ENSBTAG00000046562, BMP2, ENSBTAG00000048291, ENSBTAG00000045753, ENSBTAG00000046276, ENSBTAG00000046776, ENSBTAG00000047638, ENSBTAG00000047320, ENSBTAG00000047962, ENSBTAG00000048089, ENSBTAG00000047653, ENSBTAG00000047102, ENSBTAG00000046716, ENSBTAG00000045555, ENSBTAG00000046631, ENSBTAG0000007199, HAO1, HAO1, ENSBTAG00000032156, ADRA1D, SMOX, RNF24, PANK2, bta-mir-103-2, MAVS, AP5S1, CDC25B, CENPB, SPEF1, C13H20orf27, HSPA12B, ENSBTAG00000027412, SIGLEC1, ADAM33, GFRA4, GFRA4, ATRN, C13H20orf194, C13H20orf194, C13H20orf194, SLC4A11, ITPA, DDRGK1, LZTS3, UBOX5, AVP, OXT, MRPS26, PTPRA, VPS16, PCED1A, TMEM239, C13H20orf141, CPXM1, CPXM1, EBF4, IDH3B, IDH3B, NOP56, RF00274, RF00275, RF00594, RF00432, RF00610, TMC2, SNRPB, RF00570, TGM6, TGM3, ENSBTAG00000022534, RF00156, STK35, PDYN, SIRPA, SIRPA, ENSBTAG00000048087, ENSBTAG00000039016, ENSBTAG00000039016, SIRPB1, SIRPB1, ENSBTAG00000011638, ENSBTAG00000027221, GINS1, PCMTD2, MYT1, NPWR2, OPRL1, LKAAEAR1, RGS19, TCEA2, SOX18, PRPF6, RF00026, SAMD10, ZNF512B, ZNF512B, UCKL1, bta-mir-1388, DNAJC5, DNAJC5, TPD52L2, TPD52L2, ABHD16B, ZBTB46, SLC2A4RG, ENSBTAG00000007501, ZGPAT, ZGPAT, ARFRP1, TNFRSF6B, ENSBTAG00000047150, ENSBTAG00000047150, ENSBTAG00000047150, RTEL1, STMN3, GMEB2, FNDC11, SRMS, PTK6, PPDPF, EEF1A2, KCNQ2, KCNQ2, ENSBTAG00000031764, CHRNA4, COL20A1, COL20A1, COL20A1, ARFGAP1, ARFGAP1, BIRC7, BIRC7, YTHDF1, bta-mir-124b, RF00635, BHLHE23, SLC17A9, GID8, DIDO1, ENSBTAG00000046520, COL9A3, OGFR, MRGBP, NTSR1, SLCO4A1, bta-mir-133a-1, bta-mir-1-1, GATA5, RBBP8NL, CABLES2, RPS21, bta-mir-2306, LAMA5, ADRM1, ADRM1, OSBPL2, OSBPL2, OSBPL2, PSMA7, ENSBTAG00000021558, HRH3, MTG2, SS18L1, TAF4, TAF4, CDH4, CDH26, FAM217B, PPP1R3D, SYCP2, PHACTR3, EDN3, ZNF831, PRELID3B, ATP5F1E, TUBB1, CTSZ, NELFCD, GNAS, GNAS, GNAS, GNAS, bta-mir-296, NPEPL1, ENSBTAG00000003872, ENSBTAG00000038838, APCDD1L, VAPB, RAB22A, ANKRD60, C13H20orf85, PMEPA1, ZBP1, ZBP1, ZBP1, PCK1, CTCFL, RBM38, RAE1, SPO11, ENSBTAG00000047773, BMP7, bta-mir-2307, ENSBTAG00000037492, TFAP2C, ENSBTAG00000024608, FAM209A, RTF2, GCNT7, CASS4, RF00001, CSTF1, AURKA, FAM210B, MC3R, SIRPD, ENSBTAG00000006539, NSFL1C, FKBP1A, SDCBP2, SNPH, RAD21L1, TMEM74B, PSMF1, RSPO4, ANGPT4, FAM110A, SLC52A3, SRXN1, TCF15, CSNK2A1, TBC1D20, RBCK1, TRIB3, NRSN2, ENSBTAG00000027386, ZCCHC3, C13H20orf96, DEFB129, ENSBTAG00000048009, DEFB119, DEFB119, DEFB122A, DEFB122, DEFB123, DEFB124, REM1, HM13, ID1, COX4I2, BCL2L1, TPX2, MYLK2, FOXS1, ENSBTAG00000031214, DUSP15, TTLL9, PDRG1, XKR7, CCM2L, HCK, HCK, TM9SF4, PLAGL2, POFUT1, POFUT1, POFUT1, POFUT1, KIF3B, ASXL1, NOL4L, NOL4L, COMMD7, DNMT3B, MAPRE1, SUN5, RF00026, BPIFB2, BPIFB6, BPIFB6, BPIFB6, BPIFB3, BPIFB4, ENSBTAG00000038453, BPIFA2A, BPIFA2C, ENSBTAG00000031375,

continuação...

ENSBTAG00000011704, BPIFA2B, ENSBTAG00000031373, BPIFA3, BPIFA3, BPIFA1, BPIFB1, BPIFB5, ENSBTAG00000031354, CDK5RAP1, SNTA1, ENSBTAG00000014869, CBFA2T2, NECAB3, C13H20orf144, C13H20orf144, ACTL10, E2F1, PXMP4, ZNF341, ZNF341, CHMP4B, RALY, EIF2S2, RF00045, ASIP, AHCY

13:64358933-64432139

ITCH, ENSBTAG00000046623

13:64450162-64500844

ITCH, DYNLRB1, MAP1LC3A, PIGU

25:18619922-18639119

ACSM3

25:18672747-18714923

REXO5

25:18946238-18966876

DNAH3

Stayability

1:33285504-33305018

CADM2

3:43009131-43028743

CDC14A

4:55984897-56013042

ZNF277

11:1424088-1433201

ACOXL

13:1339928-13434427

PLCB1, PLCB4, LAMP5, PAK5, ENSBTAG00000046074, RF00408, MRPL33, ANKEF1, SNAP25, MKKS, RF00026, SLX4IP, JAG1, SRY, BTBD3, RF00156, ENSBTAG00000045631, SPTLC3, RF00273, ISM1, TASP1, TASP1, ESF1, NDUFAF5, SEL1L2, FLRT3, MACROD2, KIF16B, ENSBTAG00000032609, SNRPB2, OTOR, NANP, ENSBTAG00000023169, ENSBTAG0000004354, ENSBTAG00000022922, ENSBTAG00000022922, ENSBTAG00000040046, CCDC3, CAMK1D, RF00619, CDC123, NUDT5, SEC61A2, DHTKD1, UPF2, RF00012, PROSER2, ECHDC3, USP6NL, CELF2

Supplementary material B. Genes in significant genomics regions associated with reproductive traits in Nelore cattle

Age at first calving				
Category	Term	Genes	P-Value	GO Term
Biological Process	mitochondrial translational elongation	4	3,6E-2	GO:0070125
	mitochondrial translational initiation	4	3,6E-2	GO:0070124
	regulation of intracellular pH	3	3,9E-2	GO:0051453
Cell Component	spindle pole	4	3,6E-2	GO:0000922
Molecular Function	interleukin-1, Type I, activating receptor activity	2	1,9E-2	GO:0004909
	interleukin-1 receptor activity	2	2,8E-2	GO:0004908

Supplementary material C. Enriched GO terms and KEGG pathways significantly enriched (≤ 0.05) from DAVID software for age at first calving

Accumulated productivity				
Category	Term	Genes	P-Value	GO Term
Biological Process	positive regulation of JAK-STAT cascade	4	1,1E-4	GO:0046427
	response to organic substance	3	1,1E-3	GO:0010033
	response to hydrogen peroxide	3	7,8E-3	GO:0042542
	regulation of multicellular organism growth	3	1,2E-2	GO:0040014
	negative regulation of complement activation, classical pathway	2	1,5E-2	GO:0045959
	protein phosphorylation	5	1,9E-2	GO:0006468
	protein localization to paranode region of axon	2	2,9E-2	GO:0002175
	metabolic process	4	3,1E-2	GO:0008152
	relaxation of cardiac muscle	2	3,7E-2	GO:0055119
	relaxation of vascular smooth muscle	2	4,4E-2	GO:0060087
	paranodal junction assembly	2	4,4E-2	GO:0030913
	Cell Component	other organism cell	2	1,5E-2
basolateral plasma membrane		4	3,9E-2	GO:0016323
cytosol		14	4,3E-2	GO:0005829
Molecular function	heparin binding	5	8,2E-3	GO:0008201
	GTPase activator activity	6	1,2E-2	GO:0005096
	pancreatic polypeptide receptor activity	2	2,2E-2	GO:0001602
	[heparan sulfate-glucosamine N-sulfotransferase activity	2	3,0E-2	GO:0015016
	peptide YY receptor activity	2	3,0E-2	GO:0001601
Pathway	Complement and coagulation cascades	5	2,9E-3	-
	Cell adhesion molecules (CAMs)	6	7,7E-3	-
	MicroRNAs in cancer	6	4,8E-2	-

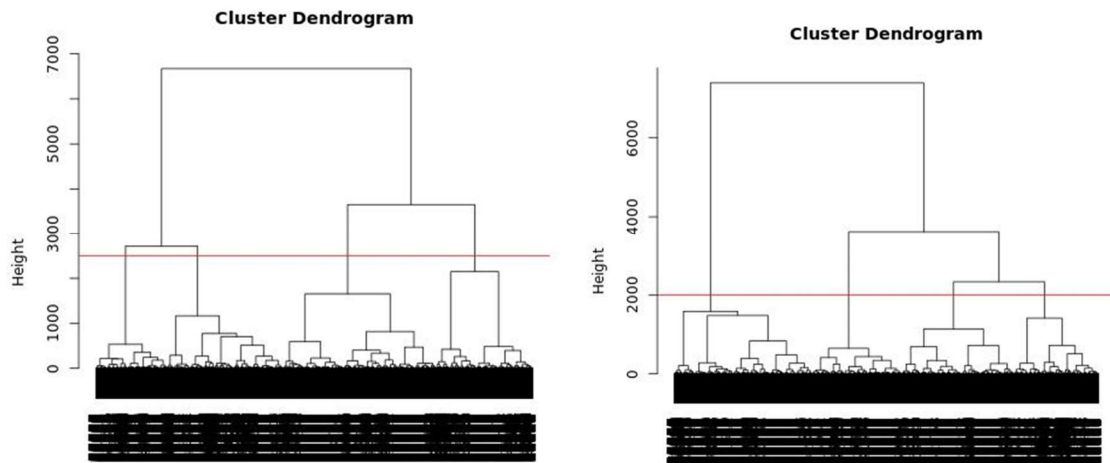
Supplementary material D. Enriched GO terms and KEGG pathways significantly enriched (≤ 0.05) from DAVID software for accumulated productivity

Heifer early calving until 30 months				
Category	Term	Genes	P-Value	GO Term
Biological Process	negative regulation of cysteine-type endopeptidase activity	9	1,3E-11	GO:2000117
	innate immune response	15	9,6E-4	GO:0045087
	fatty acid alpha-oxidation	3	7,8E-3	GO:0001561
	defense response to bacterium	7	1,2E-2	GO:0042742
	negative regulation of transcription, DNA-templated	13	1,8E-2	GO:0045892
	negative regulation of sequence-specific DNA binding transcription factor activity	5	1,9E-2	GO:0043433
	neuron projection extension	3	2,6E-2	GO:1990138
	cytoskeleton-dependent intracellular transport	3	3,6E-2	GO:0030705
	microtubule-based movement	5	4,0E-2	GO:0007018
	negative regulation of fat cell differentiation	4	4,5E-2	GO:0045599
	regulation of sodium ion transmembrane transport	2	4,6E-2	GO:1902305
	Cell Component	cell surface	16	7,7E-3
extracellular region		20	1,9E-2	GO:0005576
microtubule		9	2,1E-2	GO:0005874
extrinsic component of cytoplasmic side of plasma membrane		5	3,4E-2	GO:0031234
Molecular Function	cysteine-type endopeptidase inhibitor activity	10	7,5E-9	GO:0004869
	protease binding	10	3,0E-8	GO:0002020
	lipid binding	14	6,8E-8	GO:0008289
	oxidoreductase activity	12	5,5E-5	GO:0016491
	neurohypophyseal hormone activity	2	4,7E-2	GO:0005185
phytanoyl-CoA dioxygenase activity	2	4,7E-2	GO:0048244	
Pathway	ribosome biogenesis in eukaryotes	7	3,0E-3	-

Supplementary material E. Enriched GO terms and KEGG pathways significantly enriched (≤ 0.05) from DAVID software for heifer early calving until 30 months

Stayability				
Category	Term	Genes	P-Value	GO Term
Biological Process	cerebral cortex development	3	2,1E-3	GO:0021987
Molecular Function	phosphatidylinositol phospholipase C activity	2	2,8E-2	GO:0004435
Pathway	Aldosterone synthesis and secretion	3	9,0E-3	-
	Insulin secretion	3	9,7E-3	-
	Sphingolipid signaling pathway	3	2,0E-2	-
	Oxytocin signaling pathway	3	3,1E-2	-

Supplementary material F. Enriched GO terms and KEGG pathways significantly enriched (≤ 0.05) from DAVID software for stayability



Apêndice G. Dendrograma baseado nos valores genéticos (EBV) (esquerda) e nos valores genéticos genômicos (GEBV) (direita), obtido pela análise de agrupamento hierárquica pelo método de Ward, para circunferência escrotal aos 365 dias (CE365) e 450 dias (CE450) de idade, idade ao primeiro parto (IPP), produtividade acumulada (PAC) e período de gestação (PG), como característica do bezerro, para bovinos da raça Nelore. A linha vermelha indica os grupos que foram divididos a partir da distância Euclidiana