UNIVERSIDADE ESTADUAL PAULISTA

CÂMPUS DE JABOTICABAL

SELEÇÃO GENÔMICA PARA PRECOCIDADE SEXUAL DE MACHOS DA RAÇA NELORE

Ester Ferreira Felipe

Zootecnista

2021

UNIVERSIDADE ESTADUAL PAULISTA

CÂMPUS DE JABOTICABAL

SELEÇÃO GENÔMICA PARA PRECOCIDADE SEXUAL DE MACHOS DA RAÇA NELORE

Discente: Ester Ferreira Felipe

Orientador: Prof. Dr. Maurício Mello de Alencar

Coorientador: Dr. Rafael Espigolan

Dissertação de Mestrado apresentada ao Programa de Ciência Animal da Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Genética e Melhoramento Animal.

2021

Felipe, Ester Ferreira
F315s Seleção genômica para precocidade sexual de machos da raça Nelore / Ester Ferreira Felipe. – – Jaboticabal, 2021
43 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientador: Maurício Mello de Alencar
Coorientador: Rafael Espigolan

1. Genética animal. 2. Melhoramento genético. 3. Puberdade. 4. Bovino de corte. 5. Touro. 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).

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: SELEÇÃO GENÔMICA PARA PRECOCIDADE SEXUAL DE MACHOS DA RAÇA NELORE

AUTORA: ESTER FERREIRA FELIPE
ORIENTADOR: MAURICIO MELLO DE ALENCAR
COORIENTADOR: RAFAEL ESPIGOLAN

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

Pós-doutorando RAFAEL ESPIGOLAN (Participaçao Virtual)
Universidade de São Paulo (FZEA/USP) / Pirassununga/SP

Rafael Espigolan

Profa. Dra. CARINA UBIRAJARA FARIA (Participaçao Virtual)
Universidade Federal de Uberlândia / Uberlândia/MG

p/ Rafael Espigolan

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

p/ Rafael Espigolan

Jaboticabal, 10 de agosto de 2021

DADOS CURRICULARES DO AUTOR

Ester Ferreira Felipe, filha de José Antônio do Carmo Felipe e Leida Márcia Ferreira Neves, irmã de Arthur Ferreira Felipe, nasceu em Uberlândia, Minas Gerais, no dia 18 de outubro de 1995. Em 14 de abril de 2014 ingressou no curso de Zootecnia da Universidade Federal de Uberlândia (UFU), onde participou e foi bolsista do Programa de Educação Tutorial do curso de Zootecnia (PET - Zootecnia) orientado pela Profa. Dra. Natascha Almeida Marques da Silva e Profa. Dra. Janine França, de 01 de dezembro de 2014 a 31 de agosto de 2018, além de ser integrante do Grupo de Estudos de Parasitos Animais (GEPAN) orientado pela Profa. Dra. Fernanda Rosalinski Moraes, de 01 de março de 2016 a 30 de junho de 2016, integrante do Grupo de Estudos em Melhoramento Genético Animal (GEMEGA) orientado pela Profa. Dra. Carina Ubirajara de Faria, de 01 de agosto de 2016 a 31 de agosto de 2018, e monitora na disciplina Melhoramento Genético Animal I de 01 de abril de 2018 a 14 de julho de 2018. Realizou estágio obrigatório na Associação Nacional de Criadores e Pesquisadores (ANCP) de 05 de agosto de 2018 a 05 de dezembro de 2018. Colou grau dia 21 de dezembro de 2018 para obtenção do título de Zootecnista. Em 11 de março de 2019 iniciou o curso de Mestrado do Programa de Pós-Graduação em Genética e Melhoramento Animal, atualmente Programa de Pós-Graduação em Ciência Animal, da Faculdade de Ciências Agrárias e Veterinárias – UNESP/Câmpus de Jaboticabal, para se habilitar ao Título de Mestre em Genética e Melhoramento Animal, onde defendeu e foi aprovada a obter o título em 10 de agosto de 2021.

“Não há tempo, tão breve é a vida, para discussões banais, desculpas, amarguras, prestações de conta. Só há tempo para amar, e mesmo para isso, é só um instante.”

Mark Twain

Aos meus pais, José Antônio do Carmo Felipe e Leida Márcia Ferreira Neves, por serem meu porto seguro e minha força.

Dedico e Ofereço

AGRADECIMENTOS

A Deus por tudo.

Ao Programa de Pós-graduação em Genética e Melhoramento Animal da Faculdade de Ciências Agrarias e Veterinárias (FCAV/UNESP) pela oportunidade de realizar mais esta etapa da minha vida.

Ao meu orientador Prof. Dr. Maurício Mello de Alencar pela oportunidade, confiança, paciência e orientação.

Ao meu coorientador, Rafael Espigolan, pela coorientação e dedicação, que foi fundamental para a conclusão deste trabalho, além de todo apoio, por me ouvir, confortar e motivar.

Ao Prof. Dr. Fernando Baldi pelos ensinamentos, apoio, auxílio, paciência e valiosas sugestões neste trabalho.

A Me. Marisol Londoño Gil, pelas correções, apoio e colaboração com este trabalho.

Aos membros do comitê de avaliação do projeto de pesquisa, da banca de qualificação e da banca de defesa da dissertação de mestrado, Dra. Carina Ubirajara de Faria, Dr. Danisio Prado Munari e Dr. Fernando Baldi, pela disponibilidade para comporem minha banca e pelas sugestões e o incentivo que contribuíram para a melhoria deste trabalho.

A Associação Nacional de Criadores e Pesquisadores (ANCP) e aos seus funcionários, que considero meus amigos de coração, por fornecerem os dados para realização deste trabalho, contribuírem com minha formação profissional e terem sido minha família em Ribeirão Preto.

Ao Grupo de Estudos em Reprodução Animal (GERA-MS) da Universidade Federal do Mato Grosso do Sul (UFMS) orientado pela Profa. Dra. Eliane Vianna da Costa e Silva e as fazendas que monitoraram a Idade à Puberdade de Machos por fornecerem os dados e contribuírem para existência deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, e pela prorrogação da mesma devido a pandemia do COVID-19.

Aos meus pais, José Antônio e Leida Márcia, por me permitirem escolher qual caminho seguir, e ao meu irmão Arthur Felipe, que junto aos meus pais não mediu esforços para me apoiar e também foi fonte de amor, carinho e força durante toda minha vida.

Ao meu amor e melhor amigo, Arthur Estevam, pela ajuda, incentivo, suporte, paciência, força e, principalmente, pelo seu amor.

Aos meus amigos da Universidade Estadual Paulista (UNESP) e da Universidade de São Paulo (USP) por tornarem o mestrado mais leve, me auxiliarem em todas as minhas dificuldades, por serem meu porto seguro da área acadêmica, pelo carinho e amizade de vocês.

Aos meus amigos, professores e funcionários da Universidade Federal de Uberlândia (UFU) que contribuíram e ainda contribuem com minha formação profissional, o apoio e o incentivo de vocês foram fundamentais para eu chegar onde estou.

Aos meus professores do ensino médio e fundamental da Escola SESI Guiomar de Freitas Costa por colaborarem com meu ensino, por serem fonte de aprendizado e incentivo e um agradecimento especial para os professores de matemática Edson e Silvânia, que acreditavam em mim e no meu potencial e por até hoje me acompanharem e comemorarem minhas vitórias comigo.

Aos meus amigos e familiares, incluso os familiares do Arthur Estevam que também considero minha família, por compreenderem a ausência nos momentos importantes e únicos e por me encherem de amor sempre que possível, sem vocês jamais teria essa conquista.

Aos Mestres da Instituição Religiosa Perfect Liberty pelas orientações e suporte, e por toda minha família da PL por serem exemplos de fé e me darem todo amor e carinho quando precisei.

A minha psicóloga, Luiza Dias, por me amparar e contribuir indiretamente com desenvolvimento deste trabalho, com certeza tenho sorte em tê-la em minha vida.

A todos que contribuíram para eu chegar até aqui.

SUMÁRIO

CAPÍTULO 1 – CONSIDERAÇÕES GERAIS.....1

1.1. INTRODUÇÃO.....1
1.2. OBJETIVO GERAL.....3
1.2.1. Objetivos específicos.....3
1.3. REVISÃO DE LITERATURA.....3
1.3.1. Puberdade de machos.....3
1.3.2. Idade à Puberdade de Machos (IPM).....4
1.3.3. Seleção Genômica.....6
1.3.4. Modelos Bayesianos.....7
1.3.4.1. Bayes A.....7
1.3.4.2. Bayes B.....8
1.3.4.3. Bayes C.....9
1.3.4.4. Bayesian Lasso.....9
1.3.5. GBLUP.....10
1.3.6. ssGBLUP.....10
1.4. REFERÊNCIAS.....11

CAPÍTULO 2 – SELEÇÃO GENÔMICA UTILIZANDO GBLUP E MODELOS BAYESIANOS PARA PRECOCIDADE SEXUAL DE MACHOS DA RAÇA NELORE.....16

2.1. INTRODUÇÃO.....17
2.2. MATERIAL E MÉTODOS.....19
2.2.1. Informações gerais.....19
2.2.2. Estimação do efeito dos marcadores.....20
2.2.3. Estimação dos pseudo-fenótipos.....21
2.2.4. Predição dos valores genômicos.....21
2.2.5. Estratégia de treinamento e validação.....21
2.2.6. Critérios de comparação dos modelos.....23
2.3. RESULTADO E DISCUSSÃO.....23
2.4. CONCLUSÃO.....26
2.5. REFERÊNCIAS.....26

CAPÍTULO 3 – SELEÇÃO GENÔMICA UTILIZANDO MODELOS BAYESIANOS E SSGBLUP PARA PRECOCIDADE SEXUAL DE MACHOS DA RAÇA NELORE... 31

3.1. INTRODUÇÃO 32
3.2. MATERIAL E MÉTODOS 33
3.2.1. Informações gerais 33
3.2.2. Estimação do efeito dos marcadores 35
3.2.3. Estimação dos pseudo-fenótipos 35
3.2.4. Estratégia de treinamento e validação 36
3.2.5. Predição dos valores genômicos 37
3.2.6. Critérios de comparação dos modelos 37
3.3. RESULTADO E DISCUSSÃO 37
3.4. CONCLUSÃO 39
3.5. REFERÊNCIAS 40

SELEÇÃO GENÔMICA PARA PRECOCIDADE SEXUAL DE MACHOS DA RAÇA NELORE

RESUMO – O objetivo deste trabalho foi avaliar a aplicação de diferentes modelos e testar diferentes cenários de treinamento e validação para a predição de valores genômicos para Idade à Puberdade de Machos (IPM), em bovinos da raça Nelore. O conjunto de dados utilizado pertence às fazendas participantes do Programa Nelore Brasil, da Associação Nacional de Criadores e Pesquisadores (ANCP), sendo dados de 11.095 animais fenotipados para IPM e 37.146 animais genotipados com o painel CLARIFIDE® Nelore 3.0 contendo 27.821 marcadores do tipo *single nucleotide polymorphisms* (SNP). Os efeitos dos marcadores foram estimados a partir dos dados genômicos, considerando diferentes distribuições *a priori* para os efeitos e variâncias dos SNPs. Os modelos estudados foram: *Bayes A*, *Bayes B*, *Bayes C*, *Bayesian LASSO*, GBLUP (*Genomic Best Linear Unbiased Prediction*) e ssGBLUP (*single-step Genomic Best Linear Unbiased Prediction*). A habilidade de predição das diferentes metodologias foi comparada testando diferentes cenários de treinamento e validação e por meio das correlações entre os pseudo-fenótipos (EBV (valor genético estimado) e Y* (fenótipo ajustado para os efeitos fixos)) e o valor genômico direto predito (DGV). Este trabalho teve como objetivo avaliar a aplicação dos modelos Bayesianos, GBLUP e ssGBLUP e testar diferentes cenários de treinamento e validação. Não houve diferença na habilidade de predição e sobretudo no viés, entre os modelos bayesianos, e estes são mais vantajosos para realizar a seleção genômica para IPM quando comparados ao GBLUP, sendo menos viesados e possuindo maior habilidade de predição. Não houve diferença na habilidade de predição entre os modelos bayesianos e o ssGBLUP, entretanto os modelos Bayes A, Bayes B e Bayes C apresentaram DGVs menos viesados. A metodologia mais adequada para predizer os valores genômicos da IPM foi a validação cruzada.

Palavras-chave: GBLUP, idade à puberdade, modelos bayesianos, ssGBLUP

GENOMIC SELECTION FOR SEXUAL PRECOCITY IN NELLORE MALES

ABSTRACT – The objective of this work was to evaluate the application of different models and test different training and validation scenarios for the prediction of genomic values for Age at Puberty of Males (IPM) in Nellore cattle. The data set used belongs to farms participating in the Nelore Brazil Program, of the National Association of Breeders and Researchers (ANCP), with data from 11,095 animals phenotyped for IPM and 37,146 animals genotyped with the CLARIFIDE® Nellore 3.0 panel containing 27,821 single-type markers nucleotide polymorphisms (SNP). The effects of the markers were estimated from genomic data, considering different a priori distributions for the effects and variances of the SNPs. The models studied were: Bayes A, Bayes B, Bayes C, Bayesian LASSO, GBLUP (Genomic Best Linear Unbiased Prediction) and ssGBLUP (single-step Genomic Best Linear Unbiased Prediction). The predictive ability of the different methodologies was compared by testing different training and validation scenarios and through the correlations between the pseudo-phenotypes (EBV (estimated breeding value) and Y* (phenotype adjusted for fixed effects)) and the direct genomic value predicted (DGV). This work aimed to evaluate the application of Bayesian, GBLUP and ssGBLUP models and test different training and validation scenarios. There was no difference in prediction ability, and especially in bias, between Bayesian models, and these are more advantageous to perform genomic selection for IPM when compared to GBLUP, being less biased and having greater predictive ability. There was no difference in prediction ability between the Bayesian models and the ssGBLUP, however the Bayes A, Bayes B and Bayes C models presented less biased DGVs. The most adequate methodology to predict IPM genomic values was cross-validation.

Keywords: GBLUP, age at puberty, bayesian models, ssGBLUP

CAPÍTULO 1 – Considerações Gerais

1.1. Introdução

No sistema de produção de carne bovina, as características de reprodução, carcaça, qualidade da carne e eficiência alimentar são as que têm maior efeito na lucratividade (Hayes et al., 2013), sendo as características reprodutivas as principais determinantes da eficiência biológica e econômica do sistema (Baldi et al., 2008). Por isso, ao selecionar os animais, é ideal que os produtores considerem utilizar as características reprodutivas como critério de seleção, para melhorar a lucratividade e a eficiência biológica dos sistemas de gado de corte (Bonamy et al., 2018).

Selecionar animais que alcançam a puberdade reprodutiva precocemente tem como finalidade a utilização destes antecipadamente na reprodução, o que encurta o intervalo de gerações e, consequentemente, potencializa o ganho genético do rebanho e produz maior número de animais. A seleção de animais precoces propicia o retorno antecipado de dinheiro ao produtor, a redução das categorias improdutivas do rebanho e o aumento da fertilidade e produtividade da população (Siddiqui et al., 2008; Silveira et al., 2010; Fortes et al., 2012 e Mello, 2014).

A principal característica reprodutiva utilizada nos programas de melhoramento para os machos é a circunferência escrotal (CE) avaliada em diferentes idades em virtude da herdabilidade moderada a alta, correlação favorável com as características reprodutivas das fêmeas, de carcaça e crescimento e à sua alta repetibilidade (Siqueira et al., 2013; Buzanskas et al., 2017; Bonamy et al., 2018 e Schmidt et al., 2019). Entretanto, esta medida não assegura que um macho com alta CE em idade precoce terá uma boa qualidade seminal quando atingir a maturidade sexual, porque esta característica não retrata a realidade do parênquima testicular (Bailey et al., 1996). Por isso, na literatura há estudos sobre outros métodos que podem ser usados para a avaliar a precocidade dos machos.

Costa e Silva et al. (2017) desenvolveram a característica Idade à Puberdade de Machos (IPM) que recomenda realizar, além da circunferência escrotal, a ultrassonografia testicular e o exame andrológico a partir da desmama no animal, que assim, definir-se-ia a precocidade sexual levando em consideração o conceito de

puberdade de Wolf et al. (1965), no qual o animal é considerado púbere quando apresenta em seu ejaculado motilidade progressiva maior ou igual a 10% e concentração espermática total de, no mínimo, 50 milhões de espermatozoides. A IPM apresenta variação genética aditiva suficiente para responder à seleção, tem herdabilidade alta, é correlacionada geneticamente de maneira favorável com a precocidade sexual feminina e a probabilidade de permanência no rebanho das fêmeas, e não influencia as características de crescimento e carcaça, possibilitando ainda melhor identificação dos touros jovens candidatos à reprodução em comparação com aqueles selecionados com base na circunferência escrotal (Silva Neto et al., 2020).

O que limita o produtor de utilizar a IPM como critério de seleção é a dificuldade de medir fenotipicamente esta característica, pois requer a realização de exames andrológicos e de ultrassonografia testicular desde a desmama até o animal atingir a puberdade. Entretanto, isto pode ser solucionado com a aplicação da seleção genômica para IPM, pois essa é especialmente útil para características que são onerosas e difíceis de mensurar nos candidatos à seleção (Goddard e Hayes, 2009; Hayes et al., 2013 e Carvalheiro, 2014).

A seleção genômica (SG) refere-se ao uso de marcadores genômicos para estimar os valores genéticos genômicos dos animais (Hayes et al., 2013) e é uma ferramenta que permite acelerar o ganho genético, viabilizando a identificação de animais genéticos superiores de forma mais precoce e precisa, quando comparada à seleção tradicional, baseada em pedigree e fenótipo, incentivando a utilização dos animais jovens e, consequentemente, diminuindo o intervalo de gerações (Carvalheiro, 2014 e Albuquerque et al., 2017). Diversas abordagens estatísticas, incluindo variações do modelo de regressão BLUP e Bayesianas, têm sido propostas para estimar os efeitos de marcadores na SG. Estas abordagens diferem com base nas suposições que fazem sobre a distribuição dos efeitos dos SNPs (Van Eenennaam et al., 2014) e não há consenso sobre o melhor método a ser abordado, tornando-se indispensável pesquisar a metodologia mais apropriada para cada conjunto de dados.

Há carência de estudos de seleção genômica para a IPM, em razão desta ser uma característica desenvolvida recentemente. Portanto, são necessários pesquisas

no campo da SG, contemplando a comparação de metodologias de predição genômica para verificar qual delas é mais adequada aos dados de IPM, em especial da raça Nelore, visto que esta representa a maior parte do rebanho brasileiro.

1.2. Objetivo geral

O objetivo deste trabalho foi avaliar a aplicação de diferentes modelos para a predição de valores genéticos genômicos, tais como BayesA, BayesB, BayesC, Bayesian LASSO, GBLUP e ssGBLUP, para a Idade à Puberdade de Machos (IPM) em bovinos da raça Nelore.

1.2.1. Objetivos específicos

- Avaliar a acurácia de predição genômica dos métodos BayesA, BayesB, BayesC, Bayesian LASSO, GBLUP e ssGBLUP para a característica Idade à Puberdade dos Machos.
- Analisar estratégias de treinamento e validação para predizer os valores para a característica Idade à Puberdade dos Machos.

1.3. Revisão de literatura

1.3.1. Puberdade de machos

A estrutura básica dos testículos permanece sem modificações desde a diferenciação sexual das gônadas no início da vida fetal até o início da puberdade, que é regulado pela maturação do eixo hipotalâmicoadeno-hipofisário. O hormônio liberador das gonadotrofinas (GnRH) é transportado pelos capilares-porta, do hipotálamo para as células da hipófise anterior, onde ele causará a liberação do hormônio folículo estimulante (FSH) e o hormônio luteinizante (LH), que irão estimular as gônadas, o FSH e os andrógenos mantêm a função gametogênica, enquanto o LH coordena a secreção de testosterona pelas células de Leydig (Hafez e Hafez, 2004).

Os machos pré-púberes, em resposta à estimulação gonadotrófica, secretam testosterona progressivamente até atingirem os níveis de testosterona adulto, cada pulso de LH é seguido de um pico de testosterona cerca de 30 a 45 minutos depois, e à medida que a puberdade avança, o aumento da testosterona no sangue provoca um decréscimo na secreção de gonadotrofinas pelo mecanismo de retroalimentação negativa. Na puberdade, os gonócitos migram para a periferia dos túbulos, diferenciam-se em espermatogônias, ajudando as células a produzir células de Sertoli, as quais permanecem presentes durante toda a vida sexual, sendo seu número um fator limitante na produção espermática (Hafez e Hafez, 2004).

Hafez e Hafez (2004) afirmam que na puberdade todos os componentes do sistema reprodutor atingem o estágio avançado de desenvolvimento, tornando o sistema como um todo funcional, podendo dizer que é quando o animal demonstra a capacidade de liberar os seus gametas e apresenta o comportamento sexual. Outra definição para puberdade, dada por Wolf et al (1965), é o momento em que o primeiro ejaculado contém 50 milhões de espermatozoides com, no mínimo, 10% de motilidade progressiva.

Após a puberdade, os testículos continuam crescendo, o número de espermatozoides no ejaculado aumenta até 18 a 24 meses de idade (Hafez e Hafez, 2004) e a maturidade sexual do touro é atingida com a estabilização da qualidade seminal, onde o ejaculado deve possuir os valores mínimos recomendados pelo CBRA (2013).

1.3.2. Idade à Puberdade de Machos (IPM)

A idade à puberdade de machos pode ser estimada de acordo com as várias definições de puberdade que existem na literatura. Queiroz (2014) propõe o uso do Hormônio Anti-Mulleriano (AMH) como marcador biológico da precocidade sexual dos machos, já que em seu estudo foi constatado que nos animais púberes os níveis de AMH são reduzidos e os da testosterona se elevam. Por outro lado, Cardilli et al. (2014) e Mello (2014) sugerem a utilização da ultrassonografia testicular para avaliar o desenvolvimento do testículo e, consequentemente, identificar os animais

sexualmente mais precoces, já que os animais púberes apresentaram maior ecogenicidade do parênquima testicular do que os pré-púberes na mesma idade.

Outra abordagem para obtenção do fenótipo diretamente nos machos, e que é o método abordado neste trabalho, é descrita por Costa e Silva et al. (2017) que recomendam realizar, além da circunferência escrotal e da ultrassonografia testicular, o exame andrológico a partir da desmama, que assim definiriam a precocidade sexual considerando o conceito de puberdade de Wolf et al. (1965), no qual o animal estará púbere quando apresentar em seu ejaculado motilidade progressiva maior ou igual a 10% e concentração espermática total de, no mínimo, 50 milhões de espermatozoides.

As avaliações ultrassonográficas de testículo são realizadas a partir da desmama em sequência de quatro avaliações intercaladas por 90 dias e realizadas sempre com o mesmo equipamento, sonda linear de 7,5 mHz, em que se adquire uma imagem no plano longitudinal-lateral, onde a densidade de pixels será lida em software próprio e a curva analisada individualmente. A coleta seminal por eletroejaculação a partir da circunferência escrotal de 19 cm é realizada com técnica diferente da utilizada para animais mais velhos, sendo crucial mais atenção e suavidade nos estímulos, e é importantíssimo o volume coletado do ejaculado, pois a definição de idade a puberdade adotada é a de Wolf et al. (1965). Baseados nestas avaliações, os animais são classificados em superprecoces, quando atingem a puberdades antes dos 14 meses, precoces, quando tornam-se púberes entre 15 e 16 meses e tradicionais, quando entram na puberdade a partir dos 17 meses (Stafuzza et al., 2019).

Selecionar os animais pela idade à puberdade de machos é uma forma de realizar a seleção direta para a precocidade sexual. Entretanto, há uma preocupação se a seleção para esta característica pode prejudicar o animal no decorrer da vida e se esta compensa financeiramente para o produtor. Silveira et al. (2010) e Costa e Silva et al. (2017) afirmam que a precocidade sexual é essencial para aumentar a lucratividade do produtor, e em um estudo com fêmeas, Terakado et al. (2015) obtiveram resultados positivos com relação à permanência no rebanho, ao desmame de bezerros mais pesados, à maior chance de bezerros nascidos vivos e à menor mortalidade pré e pós-natal para os animais que entraram na puberdade mais cedo.

1.3.3. Seleção Genômica

A seleção genômica (SG) refere-se ao uso de marcadores genômicos para estimar os valores genéticos dos animais (Hayes et al., 2013). Este tipo de seleção depende da disponibilidade de uma grande população de animais fenotipados com genótipos, chamada de população referência (Van Eenennaam et al., 2014), porque é a análise estatística desta população que estima os efeitos para cada marcador, de modo que o efeito de cada marcador seja predito simultaneamente com os outros marcadores (Goddard e Hayes, 2009), permitindo o desenvolvimento de equações de predição genômica que podem ser usadas para estimar valor genético aprimorado pela genômica (GEBV) de indivíduos baseados somente em seu genótipo (Hayes et al., 2013).

Diversas metodologias para estimar os efeitos dos marcadores e predizer os valores genéticos genômicos são descritos na literatura, oferecendo uma possibilidade de escolher o melhor método para o estudo da característica avaliada. Os métodos de seleção genômica definem as equações de predição utilizadas para estimar os efeitos dos SNPs, baseada em uma população de treinamento e para realizar a validação dessa equação, a mesma é aplicada em uma população (validação) de indivíduos com informação genotípica (Goddard e Hayes, 2009).

Graças a essas metodologias e sua incorporação em programas de melhoramento genético, a confiabilidade dos valores genéticos aumentou em níveis de até 75% em animais jovens (sem registros fenotípicos) e o custo das avaliações genéticas diminuiu, principalmente nesses animais (Schaeffer, 2006). O intervalo de geração também diminuiu, uma vez que as avaliações genômicas podem ser usadas em um grande grupo de animais candidatos à seleção (Schefers e Weigel, 2012; VanRaden et al., 2009; Wiggans et al., 2016) e graças a esta metodologia, a taxa de ganho genético em características de importância econômica dobrou em relação à obtida com metodologias convencionais (Wiggans et al., 2016; Schefers e Weigel, 2012). Esse ganho genético extra é muito favorável para empresas pecuárias que incluem informações genômicas na seleção de seus animais, uma vez que podem ser mais eficientes economicamente e mais produtivos (Hayes et al., 2009; Schefers e Weigel, 2012).

A implementação prática da seleção genômica em uma população deve ser sempre monitorada, onde a atualização frequente da população de treinamento com novos animais genotipados, deve ser feita para permitir o aumento da acurácia de predição em animais jovens, visto que a mesma depende da relação entre a população de animais dos grupos de treinamento e validação (Neves et al., 2014). Outros fatores também podem ter efeito sobre a habilidade de predição, como a herdabilidade, número de animais na população de treinamento, densidade de marcadores (SNPs) e metodologias utilizadas na predição (predição (Makowsky, 2011; Neves et al., 2014; Oldenbroek e Waaij, 2014; Silva et al., 2016).

1.3.4. Modelos Bayesianos

Quando as características são influenciadas por um SNP de maior efeito, os métodos bayesianos têm uma vantagem sobre a seleção baseada em SNPs. Métodos bayesianos como Bayes A, Bayes B, Bayes C e Bayesian Lasso, entre outros, incluem informações *a priori* sobre o efeito dos SNPs, mesclando as informações da distribuição *a priori* e os dados para fazer as estimativas, por meio de aproximações não lineares, que geralmente detectam SNPs/QTLs com variâncias diversas em cada lócus (de grandes a pequenas), tomando amostras da distribuição por meio de algoritmos de cadeias Markov Monte Carlo (MCMC) (Zhang et al., 2016).

1.3.4.1. *Bayes A*

Com essa metodologia, as informações são modeladas em duas etapas: há um modelo no nível dos dados e outro no nível das variâncias dos SNPs. O modelo no nível de dados é um BLUP, mas as variâncias são Var (g_{ij}) = $\sigma^2 g_i$, diferentes para cada SNP e são estimadas pelo modelo de variância dos SNPs. O segundo nível considera que a distribuição *a priori* das variâncias dos efeitos dos SNPs segue uma distribuição em escala qui-quadrada invertida com *a priori* $(\sigma_{gi}^2) \sim X^{-2}(v, S)$, onde S é um parâmetro de escala e v é o número de graus de liberdade. A distribuição a posteriori das variâncias dos efeitos cromossômicos σ_{gi}^2, combina a informação prévia dos efeitos e a dos dados. Uma vantagem de usar uma distribuição qui-quadrado

invertida como *a priori* para as variâncias e usar uma normal para os dados é que a distribuição posterior também é uma qui-quadrada invertida. Assim, as variâncias dos efeitos dos SNPs (σ_{gi}^2) para a distribuição a posteriori dados os efeitos dos SNPs, $P(\sigma_{gi}^2|g_i)$, são distribuídas como uma qui-quadrada invertida escalada com parâmetros $S + \mathrm{g}_i'\mathrm{g}_i$ e $v + \mathrm{n}_i$ graus de liberdade, ou seja, $P(\sigma_{gi}^2|g_i) = X_{scaled}^{-2}(v + \mathrm{n}_i, S + g_i'g_i)$, onde n_i é o número de efeitos de haplótipo do segmento (Meuwissen et al., 2001)**.**

Este modelo inclui a informação *a priori* das variâncias do erro σ_e^2, que é distribuída como uma qui-quadrada não informativa χ^2(-2;0), para a média μ uma distribuição uniforme e não informativa, enquanto que a distribuição *a priori* dos efeitos dos haplótipos dentro do segmento cromossômico i são distribuídos como um $g_i \sim \mathrm{N}(0; \sigma_{gi}^2)$ normal. A estimativa dos efeitos cromossômicos g_i, estima os efeitos de SNPs com variâncias diferentes para cada SNP, ou seja, assume uma variância diferente para cada efeito g_i, (Meuwissen et al., 2001).

1.3.4.2. *Bayes B*

O método *Bayes B* assume com probabilidade π que uma porcentagem de SNPs tem efeito zero sobre a variância da característica, ou seja, $P(g_i|\pi, \sigma_{gi}^2) = 0$ com probabilidade (π), e para o resto do SNPs $(1 - \pi)$, uma grande variância diferente é assumida para cada g_i, quer dizer, $P(g_i|\pi, \sigma_{gi}^2) \sim \mathrm{N}(0; \sigma_{gi}^2)$ com probabilidade $(1 - \pi)$. Portanto, as distribuições *a priori* são consideradas as mesmas da metodologia *Bayes A*, mas condicionadas à probabilidade $(1 - \pi)$ de que uma porcentagem dos marcadores tenha efeitos e variâncias. Agora g_i e σ_{gi}^2 são misturados e distribuídos com forma desconhecida. A estratégia de amostragem neste caso é realizada com a amostragem de Gibbs para μ e σ_e^2 e uma amostragem conjunta g_i e σ_{gi}^2 com o algoritmo Metropolis-Hastings (Meuwissen et al., 2001; Goddard et al., 2010).

1.3.4.3. *Bayes C*

Nesta metodologia, assume-se que a maioria dos QTLs não está em desequilíbrio de ligação com os SNPs, de forma que a maioria dos SNPs possui variância e efeito zero, com probabilidade π. A mesma variação é assumida para o resto dos marcadores. Em *Bayes C*, assim como em *Bayes B*, assume-se que a probabilidade (1- π) de que um SNP tenha efeito diferente de zero seja conhecida (Zaabza et al., 2017).

Uma modificação nesta metodologia é chamada *BayesC*π, onde π é desconhecido e estimado a partir dos dados (Kizilkaya et al., 2010; Habier et al., 2011), π é estimado considerando que é distribuído como um $\pi \sim \text{Uniforme}(0;1)$. A modelagem dos SNPs é a mesma de *Bayes C*. $P\left(g_i \middle| \pi, \sigma_{gi}^2\right) = 0$ com uma probabilidade $(1-\pi)$; e $P\left(g_i \middle| \pi, \sigma_{gi}^2\right) \sim \text{N}(0; \sigma_{gi}^2)$ onde $P(\sigma_{gi}^2) \sim X^{-2}(v, S)$ com probabilidade π. Neste caso, as distribuições *a posteriori* de π são conhecidas. Os diferentes parâmetros do modelo são estimados por métodos MCMC (Zaabza et al., 2017).

1.3.4.4. *Bayesian Lasso*

O método Bayesian LASSO (*Least Absolute Shrinkage and Selection Operator)* foi proposto por Tibshirani (1996) e modificado por Legarra et al. (2011), desenvolvendo o *Improved Bayesian Lasso* (IBLASSO). Nesta metodologia a distribuição marginal a *priori* para os efeitos dos SNPs é uma dupla exponencial com parâmetro $\lambda^2 \sim Gamma$, no qual a variância residual (σ_e^2) assume uma distribuição qui-quadrado invertida escalonada a *priori*, com df_e graus de liberdade e um parâmetro de escala S_e(Park e Casella, 2008).

De acordo com Park e Casella (2008), assume-se uma distribuição para variância $t_i^2 | \lambda^2 \sim Exp(\lambda^2)$, onde $\lambda^2 \sim Gamma(\varphi_1, \varphi_2)$. Sendo estimada a variância aditiva como:

$$\sigma_u^2 = \sum_{i=1}^{m} \frac{t_i^2 \sigma_e^2}{\sigma_i^2} \, 2p_i(1-p_i)$$

A distribuição a *priori* do efeito do marcador é uma exponencial. λ é o parâmetro "*sharpness*", o parâmetro que determina a forma da distribuição do efeito do SNP. A *priori* de λ é uma distribuição gama, delimitada entre 0 e 10^7. A parametrização consiste em estimar as variâncias individuais dos SNPs (t_i^2), condicional à regularização do parâmetro λ (VanRaden, 2008).

1.3.5. GBLUP

O método GBLUP assume que os efeitos de todos os SNPs são amostrados a partir da mesma distribuição normal; os efeitos de todos os marcadores são considerados pequenos com variância igual.

O BLUP genômico é definido pelo modelo:

$$\boldsymbol{y} = \mathbf{1}\boldsymbol{\mu} + \boldsymbol{Zg} + \boldsymbol{e}$$

Onde $\boldsymbol{y}$ é o vetor dos dados, $\mathbf{1}$ é um vetor de n um, μ é a média populacional, $\boldsymbol{Z}$ é a matriz de incidência que associa os registros com os efeitos dos marcadores, $\boldsymbol{g}$ é o vetor de efeitos SNPs que se supõe ser normalmente distribuído com $\boldsymbol{g} \sim N(0, \boldsymbol{G}\sigma_g^2)$, onde σ_g^2 é variância genética aditiva e $\boldsymbol{G}$ é a matriz das relações genômicas; e é o vetor do erro normal, $e \sim N(0, \sigma_e^2)$, onde σ_e^2 é a variância do erro. A matriz das relações genômicas pode ser definida como $\boldsymbol{G} = \frac{\boldsymbol{X'X}}{\sum_{i=1}^{m} p_i(1-p_i)}$, onde $\boldsymbol{X}$ é a matriz especificada para os genótipos em cada lócus, e p_i é a frequência de alelo menor para cada SNP (Zaabza et al., 2017).

1.3.6. ssGBLUP

A metodologia ssGBLUP permite a combinação das matrizes de relações genômicas ($\mathbf{G}$) e de parentesco ($\mathbf{A}$) na avaliação genética. Misztal et al. (2009) sugeriram que uma matriz de relações do numerador ($\mathbf{A}$) pode ser modificada para uma matriz ($\mathbf{H}$) que inclui relações baseadas em pedigree. A inversa da matriz $\mathbf{H}$ é obtida de acordo com Aguilar et al. (2010) da seguinte forma:

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

Onde $\boldsymbol{A^{-1}}$ é a inverso da matriz $\mathbf{A}$, $\boldsymbol{A_{22}^{-1}}$ é o inverso da matriz de parentesco de animais genotipados e $\boldsymbol{G^{-1}}$ é a inversa da matriz de relações genômicas $\mathbf{G}$, construída como em VanRaden (2008), usando as frequências dos alelos e escalada para $\boldsymbol{A_{22}}$, de forma que a média dos elementos diagonais e não diagonais de $\boldsymbol{G}$ sejam iguais às médias dos elementos diagonais e não diagonais de $\boldsymbol{A_{22}}$ (Misztal et al., 2009).

1.4. Referências

AGUILAR, I.; MISZTAL, I.; JOHNSON, D.L.; LEGARRA, A.; TSURUTA, S.; LAWLOR, T. J. (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**, [S.L.], v. 93, n. 2, p. 743-752. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2009-2730.

ALBUQUERQUE, L. G.; FERNANDES JÚNIOR, G. A.; CARVALHEIRO, R. (2017) Beef cattle genomic selection in tropical environments. In: ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS CONFERENCE, 22., 2017, Townsville. **Proceedings... .** [s. L.]: AAABG. p. 255-263. Disponível em: <http://www.aaabg.org/>. Acesso em: 19 maio 2019.

BAILEY, T. L.; MONKE, D.; HUDSON, R. S.; WOLFE, D. F.; CARSON, R. L.; RIDDELL, M. G. (1996) Testicular shape and its relationship to sperm production in mature Holstein bulls. **Theriogenology**, [S.L.], v. 46, n. 5, p. 881-887. Elsevier BV. http://dx.doi.org/10.1016/s0093-691x(96)00245-2.

BALDI, F.; ALENCAR, M. M.; FREITAS, A. R.; BARBOSA, R. T. (2008) Parâmetros genéticos para características de tamanho e condição corporal, eficiência reprodutiva e longevidade em fêmeas da raça Canchim. **Revista Brasileira de Zootecnia**, [s.l.], v. 37, n. 2, p.247-253. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/s1516-35982008000200010.

BONAMY, M.; KLUSKA, S.; PERIPOLLI, E.; LEMOS, M. V. A.; AMORIM, S. T.; VACA, R. J.; LÔBO, R. B.; CASTRO, L. M.; FARIA, C. U.; FERRARI, F. B.; BALDI, F. (2018) Genetic association between different criteria to define sexual precocious heifers with growth, carcass, reproductive and feed efficiency indicator traits in Nellore cattle using genomic information. **Journal Of Animal Breeding And Genetics**, [S.L.], v. 136, n. 1, p. 15-22. Wiley. http://dx.doi.org/10.1111/jbg.12366.

BUZANSKAS, M. E.; PIRES, P. S.; CHUD, T. C. S.; BERNARDES, P. A.; ROLA, L. D.; SAVEGNAGO, R. P.; LÔBO, R. B.; MUNARI, D. P. (2017) Parameter estimates for reproductive and carcass traits in Nelore beef cattle. **Theriogenology**, [S.L.], v. 92, p. 204-209. Elsevier BV. http://dx.doi.org/10.1016/j.theriogenology.2016.09.057.

CARDILLI, D. J.; TONIOLLO, G. H.; PASTORE, A. A.; CANOLA, J. C.; OLIVEIRA, J. A.; MERCADANTE, M. E. Z. (2014) Precocidade sexual em bovinos Nelore avaliada por ultrassonografia testicular. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, [s.l.], v. 66, n. 4, p.1296-1298. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/1678-6736.

CARVALHEIRO, R. (2014) Genomic Selection in Nelore Cattle in Brazil. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 10., Vancouver. **Proceedings...** Vancouver: WCGALP, 2014. Disponível em: <http://www.wcgalp.org/proceedings/2014>. Acesso em: 25 maio 2019.
COLÉGIO BRASILEIRO DE REPRODUÇÃO ANIMAL (2013) **Manual para exame andrológico e avaliação de sêmen animal.** 3.ed. Belo Horizonte: CBRA, 104 p.

COSTA E SILVA, E. V.; QUEIROZ, V. L. D.; COSTA FILHO, L. C. C.; FERREIRA, B. X.; FORTUNATO, A. A. A. D., SOUZA, C. C.; OLIVEIRA, V. M.; ZUCCARI, C. E. S. N. (2017) Ultrassonografia testicular como recurso de diagnóstico na andrologia bovina. In: NOGUEIRA, E.; SILVA, J. C. B.; NICACIO, A. C.; MINGOTI, G. Z. (Org.). **Ultrassonografia na reprodução e avaliação de carcaças em bovinos.** 1ed. Brasilia: EMBRAPA, p. 113 – 133. 2017, v. Cap 4.

FORTES, M. R. S.; LEHNERT, S. A.; BOLORMAA, S.; REICH, C.; FORDYCE, G.; CORBET, N. J.; WHAN, V.; HAWKEN, R. J.; REVERTER, A. (2012) Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science**, [s.l.], v. 52, n. 3, p.143-150. CSIRO Publishing. http://dx.doi.org/10.1071/an11165.

GODDARD, M. E.; HAYES, B. J. (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics**, [s.l.], v. 10, n. 6, p.381-391. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/nrg2575.

GODDARD, M. E.; HAYES, B. J.; MEUWISSEN, T. H. E. (2010) Genomic selection in livestock populations. **Genetics Research**, [S.L.], v. 92, n. 5-6, p. 413-421. Cambridge University Press (CUP). http://dx.doi.org/10.1017/s0016672310000613.

HABIER, D.; FERNANDO, R. L.; KIZILKAYA, K.; GARRICK, D. J. (2011) Extension of the bayesian alphabet for genomic selection. **Bmc Bioinformatics**, [s.l.], v. 12, n. 1. Springer Science and Business Media LLC. http://dx.doi.org/10.1186/1471-2105-12-186.

HAFEZ, E. S. E.; HAFEZ, B. (2004) **Reprodução Animal**. Tradução de Renato Campanarut Barnabe. 7. ed. Barueri: Manole. 513 p.

HAYES, B. J.; BOWMAN, P. J.; CHAMBERLAIN, A. J.; GODDARD, M. E. (2009) Invited review: Genomic selection in dairy cattle. **Journal Of Dairy Science**, [s.l.], v. 92, n. 2, p.433-443. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2008-1646.

HAYES, B. J.; LEWIN, H. A.; GODDARD, M. E. (2013) The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. **Trends In Genetics**, [s.l.], v. 29, n. 4, p.206-214. Elsevier BV. http://dx.doi.org/10.1016/j.tig.2012.11.009.

KIZILKAYA, K.; FERNANDO, R. L.; GARRICK, D. J. (2010) Genomic prediction of simulated multibreed and purebred performance using observed fifty thousand single nucleotide polymorphism genotypes1. **Journal Of Animal Science**, [S.L.], v. 88, n. 2, p. 544-551. Oxford University Press (OUP). http://dx.doi.org/10.2527/jas.2009-2064.

LEGARRA, A.; ROBERT-GRANIÉ, C.; CROISEAU, P.; GUILLAUME, F.; FRITZ, S. (2011) Improved Lasso for genomic selection. **Genetics Research**, [S.L.], v. 93, n. 1, p. 77-87. Hindawi Limited. http://dx.doi.org/10.1017/s0016672310000534.

MAKOWSKY, R.; PAJEWSKI, N. M.; KLIMENTIDIS, Y. C.; VAZQUEZ, A. I.; DUARTE, C. W.; ALLISON, D. B.; CAMPOS, G. L. (2011) Beyond Missing Heritability: prediction of complex traits. **Plos Genetics**, [s.l.], v. 7, n. 4. Public Library of Science (PLoS). http://dx.doi.org/10.1371/journal.pgen.1002051.

MELLO, R. R. C. (2014) Puberdade e maturidade sexual em touros bovinos. **ACSA - Agropecuária Científica no Semiárido**, [s.l.], v. 10, n. 3, p.11-28.

MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. (2001) Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. **Genetics**, S.L., v. 157, n. 4, p. 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**, [S.L.], v. 92, n. 9, p. 4648-4655. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2009-2064

NEVES, H. H. R.; CARVALHEIRO, R.; O'BRIEN, A. M. P.; UTSUNOMIYA, Y. T.; CARMO, A. S.; SCHENKEL, F. S.; SÖLKNER, J.; MCEWAN, J. C.; TASSELL, C. P. V.; COLE, J. B; SILVA, M. V. G. B.; QUEIROZ, S. A.; SONSTEGARD, T. S.; GARCIA, J. F. (2014) Accuracy of genomic predictions in Bos indicus (Nellore) cattle. **Genetics Selection Evolution**, [s.l.], v. 46, n. 1, artigo 17. Springer Nature. http://dx.doi.org/10.1186/1297-9686-46-17

OLDENBROEK, K.; WAAIJ, L. V. D. (2014) **Textbook Animal Breeding**: animal breeding and genetics for bsc students: Wageningen University And Research. 311 p.

PARK, T.; CASELLA, G. (2008) The Bayesian Lasso. **Journal Of The American Statistical Association**, [S.L.], v. 103, n. 482, p. 681-686. Informa UK Limited. http://dx.doi.org/10.1198/016214508000000337.

QUEIROZ, V. L. D. (2014) **Hormônio anti-Mulleriano e puberdade em machos Nelore**. 59p. Dissertação (Mestrado em Ciência Animal) – Faculdade de medicina

veterinária e Zootecnia, Universidade federal de mato Grosso do Sul, Campo Grande, 2014.

SCHAEFFER, L. R. (2006) Strategy for applying genome-wide selection in dairy cattle. **Journal Of Animal Breeding And Genetics**, [S.L.], v. 123, n. 4, p. 218-223. Wiley. http://dx.doi.org/10.1111/j.1439-0388.2006.00595.x.

SCHEFERS, J. M.; WEIGEL, K. A. (2012) Genomic selection in dairy cattle: integration of dna testing into breeding programs. **Animal Frontiers**, [S.L.], v. 2, n. 1, p. 4-9. Oxford University Press (OUP). http://dx.doi.org/10.2527/af.2011-0032.

SCHMIDT, P. I.; CAMPOS, G. S.; ROSO, V. M.; SOUZA, F. R. P.; BOLIGON, A. A. (2019) Genetic analysis of female reproductive efficiency, scrotal circumference and growth traits in Nelore cattle. **Theriogenology**, [S.L.], v. 128, p. 47-53. Elsevier BV. http://dx.doi.org/10.1016/j.theriogenology.2019.01.032.

SIDDIQUI, M. A.; BHATTACHARJEE, J.; DAS, Z. C.; ISLAM, M. M.; ISLAM, M. A.; HAQUE, M. A.; PARRISH, J. J.; SHAMSUDDIN, M. (2008) Crossbred Bull Selection for Bigger Scrotum and Shorter Age at Puberty with Potentials for Better Quality Semen. **Reproduction In Domestic Animals**, [s.l.], 43(1), p.74-79. Wiley. http://dx.doi.org/10.1111/j.1439-0531.2007.00857.x.

SILVA NETO, J. B.; PERIPOLLI, E.; COSTA E SILVA, E. V.; ESPIGOLAN, R.; NEIRA, J. D. R.; SCHETTINI, G.; COSTA FILHO, L. C. C.; BARBOSA, F. B.; MACEDO, G. G.; COSTA-BRUNES, L.; LOBO, R. B.; PEREIRA, A. S. C.; BALDI, F. (2020) Genetic correlation estimates between age at puberty and growth, reproductive, and carcass traits in young Nelore bulls. **Livestock Science**, [S.L.], v. 241. Elsevier BV. http://dx.doi.org/10.1016/j.livsci.2020.104266.

SILVA, R. M. O.; FRAGOMENI, B. O.; LOURENCO, D. A. L.; MAGALHÃES, A. F. B.; IRANO, N.; CARVALHEIRO, R.; CANESIN, R. C.; MERCADANTE, M. E. Z.; BOLIGON, A. A.; BALDI, F. S. (2016) Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population1. **Journal Of Animal Science**, [s.l.], v. 94, n. 9, p. 3613-3623. Oxford University Press (OUP). http://dx.doi.org/10.2527/jas.2016-0401.

SILVEIRA, T. S.; SIQUEIRA, J. B.; GUIMARÃES, S. E. F.; PAULA, T. A. R.; MIRANDA NETO, T.; GUIMARÃES, J. D. (2010) Maturação sexual e parâmetros reprodutivos em touros da raça Nelore criados em sistema extensivo. **Revista Brasileira de Zootecnia**, Viçosa, v. 39, n. 3, p.503-511. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/s1516-35982010000300008.

SIQUEIRA, J. B.; GUIMARÃES, J. D.; PINHO, R. O. (2013) Relação entre perímetro escrotal e características produtivas e reprodutivas em bovinos de corte: uma revisão. **Revista Brasileira de Reprodução Animal**, Belo Horizonte, v. 37, n. 1, p. 3-13.

STAFUZZA, N. B.; COSTA E SILVA, E. V.; SILVA, R. M. O.; COSTA FILHO, L. C. C.; BARBOSA, F. B.; MACEDO, G. G.; LÔBO, R. B.; BALDI, F. (2019) Genome-wide

association study for age at puberty in young Nelore bulls. **Journal Of Animal Breeding And Genetics**, v. 137, n. 2, p. 234-244. Wiley. http://dx.doi.org/10.1111/jbg.12438.

TERAKADO, A. P. N.; PEREIRA, M. C.; YOKOO, M. J.; ALBUQUERQUE, L. G. (2015) Evaluation of productivity of sexually precocious Nelore heifers. **Animal**, [s.l.], v. 9, n. 06, p.938-943. Cambridge University Press (CUP). Disponível em: <http://dx.doi.org/10.1017/s1751731115000075>. Acesso em: 12 maio 2019.

TIBSHIRANI, R. (1996) Regression Shrinkage and Selection Via the Lasso. **Journal Of The Royal Statistical Society**: Series B (Methodological), [S.L.], v. 58, n. 1, p. 267-288. Wiley. http://dx.doi.org/10.1111/j.2517-6161.1996.tb02080.x.

VAN EENENNAAM, A. L.; WEIGEL, K. A.; YOUNG, A. E.; CLEVELAND, M. A.; DEKKERS, J. C. M. (2014) Applied Animal Genomics: Results from the Field. **Annual Review Of Animal Biosciences**, [s.l.], v. 2, n. 1, p.105-139. Annual Reviews. http://dx.doi.org/10.1146/annurev-animal-022513-114119.

VANRADEN, P.M. (2008) Efficient Methods to Compute Genomic Predictions. **Journal Of Dairy Science**, [S.L.], v. 91, n. 11, p. 4414-4423. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2007-0980.

VANRADEN, P. M.; VAN TASSELL, C. P.; WIGGANS, G. R.; SONSTEGARD, T. S.; SCHNABEL, R. D.; TAYLOR, J. F.; SCHENKEL, F. S. (2009) Invited Review: reliability of genomic predictions for north american holstein bulls. **Journal Of Dairy Science**, [S.L.], v. 92, n. 1, p. 16-24. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2008-1514.

WIGGANS, G. R.; COOPER, T. A.; VANRADEN, P. M.; VAN TASSELL, C. P.; BICKHART, D. M.; SONSTEGARD, T. S. (2016) Increasing the number of single nucleotide polymorphisms used in genomic evaluation of dairy cattle. **Journal Of Dairy Science**, [S.L.], v. 99, n. 6, p. 4504-4511. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2015-10456.

WOLF, F. R.; ALMQUIST, J. O.; HALE, E. B. (1965) Prepuberal Behavior and Puberal Characteristics of Beef Bulls on High Nutrient Allowance. **Journal Of Animal Science**, [s.l.], v. 24, n. 3, p.761-765. Oxford University Press (OUP). http://dx.doi.org/10.2527/jas1965.243761x.

ZAABZA, H. B.; GARA, A. B.; REKIK, B. (2017) Bayesian Modeling in Genetics and Genomics. In: TEJEDOR, Javier Prieto. **Bayesian Inference**. Rijeka: Intechopen. Cap. 10. p. 207-220.

ZHANG, X.; LOURENCO, D.; AGUILAR, I.; LEGARRA, A.; MISZTAL, I. (2016) Weighting Strategies for Single-Step Genomic BLUP: an iterative approach for accurate calculation of gebv and gwas. **Frontiers In Genetics**, [S.L.], v. 7. Frontiers Media SA. http://dx.doi.org/10.3389/fgene.2016.00151.

CAPÍTULO 2 – Seleção genômica utilizando GBLUP e modelos bayesianos para precocidade sexual de machos da Raça Nelore

RESUMO – O objetivo deste trabalho foi avaliar a aplicação de diferentes modelos e testar diferentes cenários de treinamento e validação para a predição de valores genômicos para Idade à Puberdade de Machos (IPM), em bovinos da raça Nelore. O conjunto de dados analisados veio de fazendas participantes do Programa Nelore Brasil, da Associação Nacional de Criadores e Pesquisadores (ANCP), com 7.660 animais fenotipados para IPM e 37.146 animais genotipados com o painel CLARIFIDE® Nelore 3.0 contendo 27.821 marcadores do tipo *single nucleotide polymorphisms* (SNP). Os efeitos dos marcadores foram estimados a partir dos dados genômicos, considerando diferentes distribuições *a priori* para os efeitos e variâncias dos SNPs. Os modelos estudados foram: *Bayes A*, *Bayes B*, *Bayes C*, *Bayesian LASSO* e GBLUP. A habilidade de predição das diferentes metodologias foi comparada testando diferentes cenários de validação e por meio da correlação entre o pseudo-fenótipo EBV (valor genético estimado) e o DGV (valor genômico direto predito) nas populações de validação. Não houve diferença na habilidade de predição e sobretudo no viés, entre os modelos bayesianos, e estes são mais vantajosos para realizar a seleção genômica para a Idade a Puberdade de machos do que o GBLUP, sendo menos viesados e possuindo uma maior habilidade de predição. A metodologia mais adequada para predizer os valores genômicos do IPM é validação cruzada com o maior número de animais, mesmo que diminua a média da acurácia do EBV.

Palavras-chave: GBLUP, idade à puberdade, modelos bayesianos, reprodução

2.1. Introdução

As características reprodutivas e de fertilidade somadas afetam mais o lucro do sistema de produção de bovinos de corte do que as características de crescimento, por isso, ao selecionar os animais, é ideal que os produtores considerem utilizar as características reprodutivas e associadas com a precocidade sexual como critério de seleção, para melhorar a lucratividade e a eficiência biológica dos sistemas de gado de corte (Brumatti et al., 2011; Bonamy et al., 2018). A seleção de animais que alcançam a puberdade precocemente tem como objetivo a utilização destes antecipadamente na reprodução, o que diminui o intervalo de gerações e, consequentemente, produz maior número de animais e potencializa o ganho genético do rebanho. Nas raças zebuínas, particularmente na raça Nelore, o macho que entra em puberdade antes dos 15 meses de idade já pode ser utilizado na segunda estação de monta após seu nascimento, uma antes da que era prevista, possibilitando eliminar a sua fase de recria e gerar economia considerável na produção de touros, além de acelerar o processo de melhoramento genético do rebanho (Silva Neto et al., 2020).

Os programas de melhoramento genético de raças zebuínas de corte utilizam a circunferência escrotal (CE), a idade ao primeiro parto (IPP) e a probabilidade de parto precoce (3P) como características indicadoras de precocidade sexual das fêmeas e dos machos. No entanto, a seleção nos machos tem maior impacto no progresso genético pelo fato de touros sofrerem maior pressão de seleção e produzirem maior número de progênies quando comparados às fêmeas. Contudo, e também pelo fato de as características reprodutivas dos machos apresentarem correlação genética positiva e favorável com as características reprodutivas das fêmeas, realizar a seleção utilizando características diretamente associadas à reprodução dos machos é uma opção mais favorável a ser realizada, para aumentar a taxa de ganho genético anual das características reprodutivas (Fortes et al., 2012; Johnston et al., 2014; Silva Neto et al., 2020).

Por isso, Costa e Silva et al. (2017) desenvolveram uma abordagem alternativa para avaliar a precocidade sexual dos machos, que além dos registros da circunferência escrotal, combina o exame clínico andrológico e a ultrassonografia de testículos. A nova característica que utiliza esta abordagem é chamada de Idade à

Puberdade de Machos (IPM) e possui herdabilidade moderada a alta, variação genética aditiva suficiente para responder à seleção, não influencia as características de crescimento e carcaça, e está correlacionada geneticamente de maneira favorável com a probabilidade de permanência no rebanho e a precocidade sexual das fêmeas, além de permitir uma melhor identificação dos touros jovens candidatos à reprodução em comparação com aqueles selecionados com base na circunferência escrotal (Costa e Silva et al., 2017; Stafuzza et al., 2019; Silva Neto et al., 2020).

O problema é que pelo fato da IPM ser uma característica de difícil mensuração e de alto custo dificulta o produtor a utilizá-la em larga escala como critério de seleção. Entretanto, uma alternativa para viabilizar a seleção para esta característica é aplicar a seleção genômica (SG), pois essa é especialmente útil para características que são onerosas e difíceis de mensurar nos candidatos à seleção (Goddard e Hayes, 2009; Hayes et al., 2013 e Carvalheiro, 2014). A SG consiste em uma alternativa adequada para este tipo de situação e para aumentar a intensidade de seleção, já que é capaz de testar grupos maiores de animais potencialmente superiores geneticamente do que as estruturas tradicionais de teste de progênie (Matthews et al., 2019).

A SG estima os efeitos dos marcadores por meio de uma população de referência em que todos os animais genotipados possuem fenótipos ou pseudo-fenótipos e, posteriormente, com a estimação do efeito dos SNPs, prediz os valores genômicos (Genomic Estimated Breeding Value – GEBVs) de animais que possuem apenas genótipo (Hayes et al., 2009). Os métodos bayesianos e o método GBLUP, são muito utilizados para predizer o mérito genético dos indivíduos em programas de melhoramento, sendo a principal diferença entre eles a forma de assumir a distribuição dos efeitos dos SNPs (Hayes e Goddard, 2010), tendo influência sobre a distribuição dos efeitos dos QTL e seu desequilíbrio de ligação com os marcadores genéticos (Van Eenennaam et al., 2014).

Pelo fato da idade à puberdade de machos ter sido uma característica desenvolvida recentemente, são necessários estudos genômicos para verificar qual metodologia de predição genômica é mais adequada aos dados de IPM. Por isso, o objetivo deste trabalho é avaliar a habilidade de predição dos métodos BayesA, BayesB, BayesC, Bayesian LASSO e GBLUP para a IPM, assim como determinar a

estratégia de validação mais adequada para predizer os valores genômicos para IPM na raça Nelore.

2.2. Material e métodos

2.2.1. Informações gerais

Para a realização deste trabalho foram utilizadas informações de bovinos da raça Nelore participantes do Programa Nelore Brasil da Associação Nacional de Criadores e Pesquisadores (ANCP). A característica Idade à Puberdade de Machos (IPM) foi avaliada utilizando dados provenientes de 7.660 touros fenotipados para IPM e dados genômicos de 37.146 animais genotipados com painel de baixa densidade (CLARIFIDE® Nelore 3.0) contendo 27.821 marcadores do tipo SNP.

A IPM foi obtida por meio das avaliações realizadas nos bezerros a partir da desmama utilizando exame clínico andrológico. As avaliações ultrassonográficas de testículo foram efetuadas a partir da desmama em sequência de quatro avaliações intercaladas por 90 dias e executadas sempre com o mesmo equipamento, caracterizado por uma sonda linear de 7,5 MHz, em que se adquiriu uma imagem no plano longitudinal-lateral, na qual a densidade de pixels foi lida no software Biosoft Toolbox (Biotronics, Inc.) e a curva analisada individualmente. A coleta seminal foi realizada por eletroejaculação a partir da circunferência escrotal de 19 cm, com técnica diferente da utilizada para animais mais velhos, tendo sido crucial mais atenção e suavidade nos estímulos e importante o volume coletado do ejaculado, pois a definição de idade à puberdade adotada, baseada no trabalho de Wolf et al. (1965), considerou o animal púbere quando este apresentou em seu ejaculado motilidade progressiva ≥10% e concentração espermática total de, pelo menos, 50 milhões de espermatozoides.

Os Grupos de Contemporâneos (GC) foram formados considerando a fazenda, ano, estação de nascimento e grupo de manejo. Na análise de consistência dos fenótipos, feitos no Software R (R Core Team, 2020), foram eliminadas observações com medidas de três desvios-padrão acima ou abaixo da média e grupos de

contemporâneos contendo menos de cinco animais ou contendo progênies de apenas um touro.

Os valores genéticos genômicos dos animais foram estimados utilizando diferentes modelos de predição. O modelo pode ser representado da seguinte forma:

$$\boldsymbol{Y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{e}$$

Em que: $\boldsymbol{Y}$ = vetor de observações; $\boldsymbol{X}$ = matriz de incidência para os efeitos fixos; $\boldsymbol{b}$ = vetor de efeitos fixos (GC); $\boldsymbol{Z}$ = matriz de incidência dos efeitos aleatórios; $\boldsymbol{a}$ = vetor dos efeitos aleatórios genéticos aditivos diretos dos animais; e $\boldsymbol{e}$ = vetor de efeitos residuais.

Os critérios para o controle de qualidade dos marcadores foram baseados na posição dos mesmos, frequência alélica mínima (MAF), taxa de leitura (CallRate) e teste exato de Fisher para Equilíbrio de Hardy-Weinberg (EHW). Portanto, marcadores com posição desconhecida ou redundante no genoma, bem como SNPs com MAF < 0,01, CallRate < 0,90 e P_{EHW} < 0,15 e não-autossômicos, foram excluídos.

Após realizada a consistência dos dados fenotípicos, permaneceram 7.610 animais com dados para IPM. As estatísticas descritivas, variâncias e herdabilidade da característica IPM são mostradas na Tabela 1.

TABELA 1 - Estatísticas descritivas, variâncias e herdabilidade da característica Idade à Puberdade de Machos (IPM - meses) da raça Nelore

Média	Mediana	Mín	Máx	Desvio Padrão	Coeficiente de Variação	Variância Residual (SE)	Variância Genética (SE)	Herdabilidade (SD)
17,56	16,57	9,20	22,00	3,66	20,82	7,33 (0,37)	3,86 (0,51)	0,35 (0,04)

Mín: Mínimo, Máx: Máximo, SE = SD = erro padrão aproximado.

2.2.2. Estimação dos efeitos dos marcadores

A estimativa dos efeitos dos marcadores foi realizada a partir dos dados genômicos, considerando diferentes distribuições *a priori* para os efeitos e variâncias dos SNPs. As análises genômicas que utilizaram os métodos Bayesianos, BayesA, BayesB, BayesC e Bayesian LASSO, foram executadas no Programa R (R Core Team, 2020), utilizando-se os pacotes “BGLR” e “psych”, e as que utilizaram o método GBLUP foram executadas nos programas da família BLUPF90 (Misztal et al., 2002, 2014).

2.2.3. Estimação dos pseudo-fenótipos

O valor genético estimado (EBV) foi usado como pseudo-fenótipo, considerando o GC (fazenda, ano e estação de nascimento e grupo de manejo) como efeito fixo. O coeficiente de correlação entre o pseudo-fenótipo com o DGV (valor genômico direto) foi utilizado para avaliar e comparar a acurácia da predição genômica dos métodos BayesA, BayesB, BayesC, Bayesian LASSO e GBLUP.

2.2.4. Predição dos valores genômicos

O DGV foi predito a partir dos efeitos dos SNPs estimados utilizando a seguinte equação:

$$\mathrm{DGV} = \sum_{\mathrm{j}}^{\mathrm{n}} \mathrm{X}_{\mathrm{ij}}\hat{\mathrm{g}}_{\mathrm{j}}$$

Em que: **n** é o número de SNPs, $\boldsymbol{X_{ij}}$ é o genótipo do i-ésimo animal para o j-ésimo SNP (0, 1 ou 2 para os genótipos AA, AB e BB, respectivamente) e $\boldsymbol{\hat{g}_j}$ é o vetor de efeito do j-ésimo SNP.

2.2.5. Estratégia de treinamento e validação

A população foi dividida em dois grupos (treinamento e validação), e os efeitos dos marcadores estimados na população de treinamento foram utilizados para prever o DGV na população de validação. Em razão de a estrutura dos dados deste trabalho, optamos por realizar vários cenários de análises de acordo com as populações de treinamento e validação. As diferentes populações de treinamento (PT) e validação (PV) utilizadas na avaliação genômica para IPM da população de estudo são mostradas nas Tabelas 2 e 3.

As PT (Tabela 2) de 1 a 5 possuem 799 animais sem progênie que possuem fenótipo, genótipo e EBV mais *X* animais com progênie que possuem genótipo e EBV, sendo que *X* varia em cada população de acordo com acurácia do EBV dos animais. A PT 6 é formada por 1.233 animais escolhidos aleatoriamente de um grupo de 1.644 animais que possuem EBV, fenótipo e genótipo. A PT 7 é formada por 1.928 animais

escolhidos aleatoriamente de um grupo de 2.570 animais que possuem EBV e genótipo.

A PV (Tabela 3) 1 é formada por 800 animais sem progênie que possuem fenótipo, EBV e genótipo. A PV 2 é formada por 411 animais escolhidos aleatoriamente de um grupo de 1.644 animais que possuem EBV, fenótipo e genótipo. A PV 3 é formada por 643 animais escolhidos aleatoriamente de um grupo de 2.570 animais que possuem EBV e genótipo.

TABELA 2 – Estrutura das populações de treinamento utilizadas na avaliação genômica da característica Idade à Puberdade de Machos (IPM) da raça Nelore

População de treinamento	Acurácia média EBV	Número amostral	Pseudo-fenótipo
PT1	0,25	1.770	EBV
PT2	0,28	1.442	EBV
PT3	0,30	1.225	EBV
PT4	0,32	1.041	EBV
PT5	0,32	964	EBV
PT6	0,29	1.233	EBV
PT7	0,26	1.928	EBV

PT = População de treinamento; EBV = Valor genético estimado.

TABELA 3 – Estrutura das populações de validação utilizadas na avaliação genômica da característica Idade à Puberdade de Machos (IPM) da raça nelore

População de validação	Acurácia média EBV	Número amostral	Pseudo-fenótipo
PV1	0,27	800	EBV
PV2	0,29	411	EBV
PV3	0,26	643	EBV

PV = População de validação; EBV = Valor genético estimado.

Na Tabela 4 são mostrados os diferentes cenários utilizados para realizar as avaliações e predições genômicas para a característica IPM.

Nos cenários 6 e 7 foi feito uma validação cruzada com 4 *fold* e 10 repetições aleatórias.

TABELA 4 – Cenários de Treinamento e Validação usados para as predições genômicas da característica idade à puberdade de machos (IPM) da raça Nelore

Cenário	População de Treinamento	População de Validação	Predição	Correlação com DGV
1	PT1	PV1	DGV	EBV
2	PT2	PV1	DGV	EBV
3	PT3	PV1	DGV	EBV
4	PT4	PV1	DGV	EBV
5	PT5	PV1	DGV	EBV
6	PT6	PV2	DGV	EBV

7	PT7	PV3	DGV	EBV

PT = População de treinamento; PV = População de validação; DGV = Valor genômico direto; EBV = Valor genético estimado.

2.2.6. Critérios de comparação dos modelos

A habilidade de predição das diferentes metodologias propostas no presente trabalho foi comparada por meio da correlação entre o pseudo-fenótipo (EBV) e o DGV predito pelos métodos BayesA, BayesB, BayesC, Bayesian LASSO e GBLUP nas populações de validação.

Para avaliar a extensão do viés da predição foi comparada a regressão entre a variável resposta (EBV) e o DGV, considerando somente os resultados obtidos dentro do grupo de validação para cada modelo de predição proposto, de forma que valores mais próximos de 1 (um) foram considerados menos viesados.

2.3. Resultado e discussão

Na Tabela 5 são apresentadas as habilidades de predição dos diversos cenários realizados neste trabalho.

TABELA 5 – Habilidade de predição (hp) e o coeficiente de regressão (cr) dos cenários de acordo com diferentes modelos e populações de treinamento e validação para idade à puberdade de machos (IPM) da raça Nelore

	Bayes A		***Bayes B***		***Bayes C***		***Bayesian Lasso***		**GBLUP**	
Cenário	**hp**	**cr**	**hp**	**cr**	**hp**	**Cr**	**hp**	**cr**	**hp**	**Cr**
1	0,63	0,99	0,63	0,99	0,63	0,99	0,63	1,00	0,60	1,22
2	0,60	0,93	0,60	0,93	0,60	0,93	0,60	0,94	0,58	1,00
3	0,58	0,91	0,58	0,91	0,58	0,91	0,58	0,91	0,56	0,84
4	0,58	0,92	0,58	0,92	0,58	0,92	0,58	0,93	0,56	0,78
5	0,57	0,93	0,57	0,91	0,57	0,92	0,57	0,92	0,55	0,73
6	0,64	1,00	0,63	1,01	0,64	1,00	0,64	1,01	-	-
7	0,71	0,99	0,71	0,99	0,71	0,99	0,71	0,99	-	-

GBLUP: Genomic Best Linear Unbiassed Predictor.

Não houve diferença na habilidade de predição e sobretudo no viés, entre os modelos bayesianos. Ao realizar a comparação dos métodos bayesianos com GBLUP

(Tabela 5), pode-se notar que os métodos bayesianos foram superiores na hp e menos viesados, com exceção apenas do cenário dois, onde o GBLUP foi menos viesado. De acordo com um estudo de simulação de Meuwissen (2009), o GBLUP requer mais registros para atingir uma alta precisão do que um método Bayesiano (*BayesB*), e quando o número de registros é pequeno e a densidade do marcador é alta, os métodos bayesianos parecem superiores ao GBLUP. Algo semelhante foi encontrado por Mehrban et al. (2017) que, fazendo uma comparação da precisão e viés obtido para peso da carcaça com os métodos *BayesC, BayesLasso* e GBLUP, revelaram a superioridade do modelo *BayesC*.

Silva et al. (2021) compararam a habilidade de predição dos métodos BLUP, GBLUP, ssGBLUP, *BayesA, BayesB, BayesC* e *BayesianLASSO* para características indicadoras de composição de carcaça em bovinos da raça Nelore e, assim como neste trabalho, as abordagens bayesianas para as características de escores visuais obtiveram melhor habilidade de predição e viés. Similarmente, Costa et al. (2019), avaliando características reprodutivas das fêmeas em bovinos Nelore e Neves et al. (2014), analisando período de gestação e circunferência escrotal relataram que os modelos de regressão bayesiana foram mais precisos do que o GBLUP.

Quando os diferentes cenários são comparados, nota-se que é mais vantajoso treinar com mais animais (cenário 1 vs. outros cenários, Tabela 5) mesmo que a acurácia do pseudo-fenótipo seja menor, do que treinar com menos animais de maior acurácia, melhorando tanto a habilidade de predição quanto o viés. As populações de referência para seleção genômica geralmente envolvem indivíduos selecionados, o que pode resultar em uma predição viesada dos valores genéticos genômicos preditos (Gowane et al., 2019). Um passo fundamental para evitar viés nas avaliações genômicas é a genotipagem de um bom número de indivíduos para estabelecer a população de referência, pois isso permite aumentar a precisão do DGV (Dehnavi et al., 2018). Em nosso estudo, quando a população de treinamento apresentava o maior número de indivíduos (cenário 1), foram encontradas as maiores habilidades de predição e os menores vieses, e diminuiu a hp e aumentou o viés conforme o tamanho da população de treinamento diminuiu, de acordo com a literatura (Gao et al., 2015; Mehrban et al., 2017) é uma indicação da necessidade de aumentar o número de indivíduos na população em treinamento para obter melhores predições para IPM.

Constata-se também que a validação cruzada, realizada nos cenários 6 e 7, é mais interessante do que determinar a população de treinamento, visto que a maior habilidade de predição é do cenário 7, onde tanto os animais da população de treinamento quanto os da população de validação foram escolhidos aleatoriamente de um grupo de 2.570 animais que possuem EBV e genótipo. Do mesmo modo, Silva et al. (2016), Brunes et al. (2020) e Magnabosco et al. (2020), estudando características de eficiência alimentar, encontraram a maior hp no cenário da validação cruzada, do que nas abordagens que se basearam na idade ou na acurácia do EBV para a formação das populações. Vanraden et al. (2009) observaram, em um estudo com touros Holstein, que a formação das populações considerando a idade dos animais ou acurácia do EBV pode levar a uma maior variação no número de animais nas populações de treinamento e validação quando comparado ao método aleatório e isto pode influenciar a hp.

Diferente deste trabalho, Brunes et al. (2020) relatam que, para características de alta herdabilidade, a predição genômica usando validação baseada na idade ou acurácia do EBV pode ser realizada sem perda de hp. Por isso, os autores, recomendam o uso da validação cruzada em casos de populações formadas principalmente por animais jovens e características que possuem pouco registro de dados.

Contudo, para a estrutura de dados apresentada neste trabalho, os modelos bayesianos são melhores para predizer os valores genômicos do que o GBLUP *a priori*, tendo vantagem na habilidade de predição e sobretudo no viés, bem como, a validação cruzada com maior número de animais, mesmo com a média da acurácia do EBV menor, é a metodologia mais precisa e consistente para predizer os valores genômicos da característica IPM.

A IPM é uma característica onerosa e difícil de mensurar, por isso é mensurada em poucos animais, geralmente em touros líderes, então tem-se uma coleta de fenótipos seletiva e, consequentemente, isso aumenta o viés e limita o delineamento da população de treinamento e validação. Entretanto, os resultados obtidos neste estudo demonstraram que é possível realizar a seleção genômica para IPM utilizando estratégias de treinamento e validação e metodologias que melhor se adequam a estrutura de dados da característica. A aplicação da SG para IPM é de extrema

importância porque, além de reduzir os custos e facilitar a sua utilização, a SG pode aumentar a confiabilidade dos valores genéticos em níveis de até 75% em animais jovens que não possuem registros fenotípicos, fornecendo, assim, maior segurança para o produtor utilizar IPM como critério de seleção (Schaeffer, 2006).

A IPM possui herdabilidade moderada, indicando que esta característica responderia favoravelmente à seleção além de possibilitar uma melhor identificação dos machos que estão púberes em comparação àqueles selecionados com base na circunferência escrotal. O ganho genético do rebanho é esperado, porque a IPM possui correlação favorável com a precocidade sexual das fêmeas e a permanência no rebanho das mesmas, além de não afetar o crescimento e as características de carcaça (Silva Neto et al., 2020). Os resultados deste trabalho promovem a utilização da IPM como critério de seleção concomitantemente com a seleção genômica para ter maior confiabilidade dos valores genéticos dos animais.

2.4. Conclusão

Não houve diferença na habilidade de predição e sobretudo no viés, entre os modelos bayesianos, e estes são mais vantajosos para realizar a seleção genômica para a IPM do que o GBLUP, sendo menos viesados e possuindo uma maior habilidade de predição. A metodologia mais adequada para predizer os valores genômicos do IPM é validação cruzada com o maior número de animais, mesmo que diminua a média da acurácia do EBV dos animais.

2.5. Referências

BONAMY, M.; KLUSKA, S.; PERIPOLLI, E.; LEMOS, M. V. A.; AMORIM, S. T.; VACA, R. J.; LÔBO, R. B.; CASTRO, L. M.; FARIA, C. U.; FERRARI, F. B.; BALDI, F. (2018) Genetic association between different criteria to define sexual precocious heifers with growth, carcass, reproductive and feed efficiency indicator traits in Nellore cattle using genomic information. **Journal Of Animal Breeding And Genetics**, [S.L.], v. 136, n. 1, p. 15-22. Wiley. http://dx.doi.org/10.1111/jbg.12366.

BRUMATTI, R. C.; FERRAZ, J. B. S.; ELER, J. P.; FORMIGONNI, I. B. (2011) Desenvolvimento de índice de seleção em gado corte sob o enfoque de um modelo bioeconômico. **Archivos de Zootecnia**, [s.l.], v. 60, n. 230, p.205-213. Instituto de

Salud Carlos III/BNCS/SciELO Espana. http://dx.doi.org/10.4321/s0004-05922011000200005.

BRUNES, L. C.; BALDI, F.; LOPES, F. B.; NARCISO, M. G.; LOBO, R. B.; ESPIGOLAN, R.; COSTA, M. F. O.; MAGNABOSCO, C. U. (2020) Genomic prediction ability for feed efficiency traits using different models and pseudo-phenotypes under several validation strategies in Nelore cattle. **Animal**, [S.L.], n.100085. Elsevier BV. http://dx.doi.org/10.1016/j.animal.2020.100085.

CARVALHEIRO, R. (2014) Genomic Selection in Nelore Cattle in Brazil. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 10., Vancouver. **Proceedings...** Vancouver: WCGALP, 2014. Disponível em: <http://www.wcgalp.org/proceedings/2014>. Acesso em: 25 maio 2019.

COSTA E SILVA, E. V.; QUEIROZ, V. L. D.; COSTA FILHO, L. C. C.; FERREIRA, B. X.; FORTUNATO, A. A. A. D.; SOUZA, C. C.; OLIVEIRA, V. M.; ZUCCARI, C. E. S. N. (2017) Ultrassonografia testicular como recurso de diagnóstico na andrologia bovina. In: NOGUEIRA, E.; BORGES SILVA, J. C.; NICACIO, A. C.; MINGOTI, G. Z. (ed.). **Ultrassonografia na reprodução e avaliação de carcaças em bovinos**. Brasília, DF: EMBRAPA. Cap. 4. p. 113-133. E-book : il. color.

COSTA, R. B.; IRANO, N.; DIAZ, I. D. P. S.; TAKADA, L.; HERMISDORFF, I. C.; CARVALHEIRO, R.; BALDI, F.; OLIVEIRA, H. N.; TONHATI, H.; ALBUQUERQUE, L. G. (2019) Prediction of genomic breeding values for reproductive traits in Nellore heifers. **Theriogenology**, [S.L.], v. 125, p. 12-17. Elsevier BV. http://dx.doi.org/10.1016/j.theriogenology.2018.10.014.

DEHNAVI, E.; MAHYARI, S. A.; SCHENKEL, F. S.; SARGOLZAEI, M. (2018) The effect of using cow genomic information on accuracy and bias of genomic breeding values in a simulated Holstein dairy cattle population. **Journal Of Dairy Science**, [s.l.], v. 101, n. 6, p. 5166-5176. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2017-12999.

FORTES, M. R. S.; LEHNERT, S. A.; BOLORMAA, S.; REICH, C.; FORDYCE, G.; CORBET, N. J.; WHAN, V.; HAWKEN, R. J.; REVERTER, A. (2012) Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science**, [s.l.], v. 52, n. 3, p.143-150. CSIRO Publishing. http://dx.doi.org/10.1071/an11165.

GAO, N.; LI, J.; HE, J.; XIAO, G.; LUO, Y.; ZHANG, H.; CHEN, Z.; ZHANG, Z. (2015) Improving accuracy of genomic prediction by genetic architecture based priors in a Bayesian model. **Bmc Genetics**, [s.l.], v. 16, n. 1. Springer Science and Business Media LLC. http://dx.doi.org/10.1186/s12863-015-0278-9.

GODDARD, M. E.; HAYES, B. J. (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics**, [s.l.], v. 10, n. 6, p.381-391. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/nrg2575.

GOWANE, G. R.; LEE, S. H.; CLARK, S.; MOGHADDAR, N.; AL-MAMUN, H. A.; WERF, J. H. J. (2019) Effect of selection and selective genotyping for creation of reference on bias and accuracy of genomic prediction. **Journal Of Animal Breeding And Genetics**, [s.l.], v. 136, n. 5, p. 390-407. Wiley. http://dx.doi.org/10.1111/jbg.12420.

HAYES, B. J.; BOWMAN, P. J.; CHAMBERLAIN, A. J.; GODDARD, M. E. (2009) Invited review: Genomic selection in dairy cattle. **Journal Of Dairy Science**, [s.l.], v. 92, n. 2, p.433-443. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2008-1646.

HAYES, B. J.; GODDARD, M. (2010) Genome-wide association and genomic selection in animal breedingThis article is one of a selection of papers from the conference “Exploiting Genome-wide Association in Oilseed Brassicas: a model for genetic improvement of major oecd crops for sustainable farming⁝.. **Genome**, [S.L.], v. 53, n. 11, p. 876-883. Canadian Science Publishing. http://dx.doi.org/10.1139/g10-076.

HAYES, B. J.; LEWIN, H. A.; GODDARD, M. E. (2013) The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. **Trends In Genetics**, [s.l.], v. 29, n. 4, p.206-214. Elsevier BV. http://dx.doi.org/10.1016/j.tig.2012.11.009.

JOHNSTON, D. J.; CORBET, N. J.; BARWICK, S. A.; WOLCOTT, M. L.; HOLROYD, R. G. (2014) Genetic correlations of young bull reproductive traits and heifer puberty traits with female reproductive performance in two tropical beef genotypes in northern Australia. **Animal Production Science**, [s.l.], v. 54, n. 1, p.74-84. CSIRO Publishing. http://dx.doi.org/10.1071/an13044.

MAGNABOSCO, C. U.; LOPES, F.; MAGNABOSCO, V.; LOBO, R.; PEREIRA, L.; ESPIGOLAN, R.; BRUNES, L. (2020) PSXII-22 Genomic prediction accuracy for feed efficiency related traits using different pseudo-phenotypes, prediction and validation methods in Nellore cattle. **Journal Of Animal Science**, [S.L.], v. 98, n. 4, p. 245-246. Oxford University Press (OUP). http://dx.doi.org/10.1093/jas/skaa278.446.

MATTHEWS, D.; KEARNEY, J. F.; CROMIE, A. R.; HELY, F. S.; AMER, P. R. (2019) Genetic benefits of genomic selection breeding programmes considering foreign sire contributions. **Genetics Selection Evolution**, [s.l.], v. 51, n. 1. Springer Science and Business Media LLC. http://dx.doi.org/10.1186/s12711-019-0483-5.

MEHRBAN, H.; LEE, D. H.; MORADI, M. H.; ILCHO, C.; NASERKHEIL, M.; IBÁÑEZ-ESCRICHE, N. (2017) Predictive performance of genomic selection methods for carcass traits in Hanwoo beef cattle: impacts of the genetic architecture. **Genetics Selection Evolution**, [S.L.], v. 49, n. 1. Springer Science and Business Media LLC. http://dx.doi.org/10.1186/s12711-016-0283-0.

MEUWISSEN, T. H. E. (2009) Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. **Genetics Selection Evolution**, [S.L.], v. 41, n. 1. Springer Science and Business Media LLC. http://dx.doi.org/10.1186/1297-9686-41-35.

MISZTAL, I.; TSURUTA, S.; LOURENCO, D.; AGUILAR, I.; LEGARRA, A.; VITEZICA, Z. (2014) **Manual for BLUPF90 family of programs**. University of Georgia, Athens, USA, May.

MISZTAL, I.; TSURUTA, S.; STRABEL, T.; AUVRAY, B.; DRUET, T.; LEE, D. H. (2002) **BLUPF90 and related programs (BGF90)**. Pages 1-2 in Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, August, 2002. Session 28.

NEVES, H. H.; CARVALHEIRO, R.; O'BRIEN, A. M. P.; UTSUNOMIYA, Y. T.; CARMO, A. S.; SCHENKEL, F. S.; SÖLKNER, J.; MCEWAN, J. C.; VAN TASSELL, C. P.; COLE, J. B.; SILVA, M. V. G. B.; QUEIROZ, S. A.; SONSTEGARD, T. S.; GARCIA, J. F. (2014) Accuracy of genomic predictions in Bos indicus (Nellore) cattle. **Genetics Selection Evolution**, [S.L.], v. 46, n. 1. Springer Science and Business Media LLC. http://dx.doi.org/10.1186/1297-9686-46-17.

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/.

SCHAEFFER, L. R. (2006) Strategy for applying genome-wide selection in dairy cattle. **Journal Of Animal Breeding And Genetics**, [S.L.], v. 123, n. 4, p. 218-223. Wiley. http://dx.doi.org/10.1111/j.1439-0388.2006.00595.x.

SILVA NETO, J. B.; PERIPOLLI, E.; COSTA E SILVA, E. V.; ESPIGOLAN, R.; NEIRA, J. D. R.; SCHETTINI, G.; COSTA FILHO, L. C. C.; BARBOSA, F. B.; MACEDO, G. G.; COSTA-BRUNES, L.; LOBO, R. B.; PEREIRA, A. S. C.; BALDI, F. (2020) Genetic correlation estimates between age at puberty and growth, reproductive, and carcass traits in young Nelore bulls. **Livestock Science**, [S.L.], v. 241. Elsevier BV. http://dx.doi.org/10.1016/j.livsci.2020.104266.

SILVA, R. M. O.; FRAGOMENI, B. O.; LOURENCO, D. A. L.; MAGALHÃES, A. F. B.; IRANO, N.; CARVALHEIRO, R.; CANESIN, R. C.; MERCADANTE, M. E. Z.; BOLIGON, A. A.; BALDI, F. S.; MISZTAL, I.; ALBUQUERQUE, L. G. (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**, [S.L.], v. 94, n. 9, p. 3613-3623. Oxford University Press (OUP). http://dx.doi.org/10.2527/jas.2016-0401.

SILVA, R. P.; ESPIGOLAN, R.; BERTON, M. P.; LÔBO, R. B.; MAGNABOSCO, C. U.; PEREIRA, A. S. C.; BALDI, F. (2021) Genomic prediction ability for carcass composition indicator traits in Nellore cattle**. Livestock Science**, [S.L.], v. 245. Elsevier BV. http://dx.doi.org/10.1016/j.livsci.2021.104421.

STAFUZZA, N. B.; COSTA E SILVA, E. V.; SILVA, R. M. O.; COSTA FILHO, L. C. C.; BARBOSA, F. B.; MACEDO, G. G.; LÔBO, R. B.; BALDI, F. (2019) Genome-wide association study for age at puberty in young Nelore bulls. **Journal Of Animal Breeding And Genetics**, v. 137, n. 2, p. 234-244. Wiley. http://dx.doi.org/10.1111/jbg.12438.

VAN EENENNAAM, A. L.; WEIGEL, K. A.; YOUNG, A. E.; CLEVELAND, M. A.; DEKKERS, J. C. M. (2014) Applied Animal Genomics: Results from the Field. **Annual Review Of Animal Biosciences**, [s.l.], v. 2, n. 1, p.105-139. Annual Reviews. http://dx.doi.org/10.1146/annurev-animal-022513-114119.

VANRADEN, P. M.; VAN TASSELL, C. P.; WIGGANS, G. R.; SONSTEGARD, T. S.; SCHNABEL, R. D.; TAYLOR, J. F.; SCHENKEL, F. S. (2009) Invited review: Reliability of genomic predictions for North American Holstein bulls. **Journal of Dairy Science**, [S.L.], v. 92, n. 1, p. 16-24. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2008-1514.

WOLF, F. R.; ALMQUIST, J. O.; HALE, E. B. (1965) Prepuberal Behavior and Puberal Characteristics of Beef Bulls on High Nutrient Allowance. **Journal Of Animal Science**, [s.l.], v. 24, n. 3, p.761-765. Oxford University Press (OUP). http://dx.doi.org/10.2527/jas1965.243761x.

CAPÍTULO 3 – Seleção genômica utilizando modelos bayesianos e ssGBLUP para precocidade sexual de machos da raça nelore

RESUMO – O objetivo deste trabalho foi avaliar a aplicação de diferentes modelos e testar diferentes cenários de treinamento e validação para a predição de valores genômicos para Idade à Puberdade de Machos (IPM), em bovinos da raça Nelore. A IPM foi analisada utilizando dados de bovinos da raça Nelore participantes do Programa Nelore Brasil da Associação Nacional de Criadores e Pesquisadores (ANCP). Os dados são referentes a 11.095 touros fenotipados para IPM e dados genômicos de 37.141 animais genotipados com painel de baixa densidade (CLARIFIDE® Nelore 3.0) contendo 27.821 marcadores do tipo *single nucleotide polymorphisms* (SNP). Estimaram-se os efeitos dos marcadores a partir dos dados genômicos, levando em consideração diferentes distribuições *a priori* para os efeitos e variâncias dos SNPs. As metodologias estudadas foram: *Bayes A*, *Bayes B*, *Bayes C*, *Bayesian LASSO* e ssGBLUP. A habilidade de predição dos modelos foi comparada por meio da correlação entre o pseudo-fenótipo (Y*, fenótipo corrigido para os efeitos fixos) e o DGV predito nas populações de validação pelas diferentes metodologias propostas no presente estudo. Não houve diferença na habilidade de predição entre os modelos bayesianos e ssGBLUP, entretanto os modelos Bayes A, Bayes B e Bayes C foram menos viesados. No geral, a maior diferença na habilidade de predição dos valores genéticos genômicos para IPM foi entre os cenários, sendo a validação cruzada a metodologia que apresentou maior habilidade de predição e DGVs menos viesados.

Palavras-chave: idade à puberdade, reprodução, single-step, validação cruzada

3.1. Introdução

A idade em que os animais entram na puberdade afeta o sistema de produção de bovinos de corte. Um touro precoce pode participar da segunda estação de monta após seu nascimento, uma antes da que era prevista, permitindo eliminar sua fase de recria e reduzir substancialmente os custos da produção (Silva Neto et al., 2020). Além disso, utilizar a precocidade sexual como critério de seleção reduz o intervalo de gerações, aumenta a fertilidade e a produtividade do rebanho e propicia maior lucratividade e eficiência biológica do sistema (Baldi et al., 2008; Siddiqui et al., 2008; Fortes et al., 2012; Bonamy et al., 2018; Stafuzza et al., 2019).

A circunferência escrotal (CE) avaliado em diferentes idades é a principal característica reprodutiva dos machos que os programas de melhoramento utilizam como critério de seleção. No entanto, esta característica não retrata a realidade do parênquima testicular e, por isso, não é assegurado que o um macho com alta CE em idade precoce terá uma boa qualidade seminal quando atingir a maturidade sexual (Bailey et al., 1996). Silva et al. (2002), estudando a circunferência escrotal e a qualidade do sêmen, constataram que touros da raça Nelore com a mesma idade e tamanho testicular apresentaram uma alta variação da motilidade progressiva do sêmen, comprovando que touros com a mesma circunferência escrotal e idade podem produzir sêmen com diferentes taxas de motilidade.

Costa e Silva et al. (2017), combinando medidas de circunferência escrotal, ultrassonografia testicular e análise de amostras de sêmen, desenvolveram uma nova característica para avaliar a precocidade sexual dos machos, chamada de Idade à Puberdade dos Machos (IPM). A IPM possibilita uma melhor identificação dos touros jovens precoces e possui herdabilidade moderada a alta, variação genética aditiva suficiente para responder à seleção e correlação favorável com a permanência das fêmeas no rebanho e com a precocidade sexual das fêmeas (Costa e Silva et al., 2017; Stafuzza et al., 2019 e Silva Neto et al., 2020).

Entretanto, existe uma resistência por parte dos produtores em utilizar a IPM em larga escala como critério de seleção pela sua difícil mensuração e alto custo. Uma solução para esta dificuldade é aplicar a seleção genômica (SG) para IPM, porque os custos da genotipagem diminuíram drasticamente nos últimos anos e é uma

ferramenta adequada para características que são onerosas e difíceis de mensurar (Goddard e Hayes, 2009; Hayes et al., 2013 e Carvalheiro, 2014). Além disso, acrescentar informações genômicas à característica auxiliará a desvendar genes e mecanismos moleculares que influenciam a precocidade sexual, além de aumentar a confiabilidade dos valores genéticos dos touros jovens (Stafuzza et al., 2019).

A SG refere-se à utilização de marcadores genômicos para determinar os valores genéticos genômicos dos animais (Hayes et al., 2013) e é um mecanismo que possibilita aumentar o ganho genético, reduzir o intervalo de gerações e aumentar a intensidade de seleção, pelo fato da viabilização da identificação de animais genéticos superiores de forma mais precoce e com maior confiabilidade quando comparada a seleção tradicional (Carvalheiro, 2014; Albuquerque et al., 2017; Stafuzza et al., 2019). Diferentes modelos estatísticos têm sido propostos para estimar os efeitos de marcadores na SG, incluindo variações do modelo de regressão BLUP e Bayesianas, sendo a principal diferença entre as metodologias a forma de assumir a distribuição dos efeitos dos SNPs (Hayes e Goddard, 2010; Van Eenennaam et al., 2014). Não há concordância sobre qual a melhor abordagem a ser utilizada, sendo imprescindível pesquisar o modelo estatístico mais adequado para cada conjunto de dados.

Considerando que a IPM é uma característica desenvolvida recentemente, há poucos estudos genômicos que a abordam. Por essa razão, o objetivo deste trabalho foi avaliar a aplicação dos modelos Bayes A, Bayes B, Bayes C, Bayesian Lasso e ssGBLUP e testar diferentes cenários de treinamento e validação para analisar qual modelo e estratégia são mais adequados para predizer os valores genômicos para Idade à Puberdade de Machos (IPM), em bovinos da raça Nelore.

3.2. Material e métodos

3.2.1. Informações gerais

Foram usados dados da característica IPM de bovinos da raça Nelore participantes do Programa Nelore Brasil da Associação Nacional de Criadores e Pesquisadores (ANCP). Os dados são referentes a 11.095 touros fenotipados para

IPM e dados genômicos de 37.141 animais genotipados com painel de baixa densidade (CLARIFIDE® Nelore 3.0) contendo 27.821 marcadores do tipo SNP.

A IPM foi mensurada por meio das avaliações efetuadas nos bezerros a partir da desmama utilizando exame clínico andrológico. A partir da desmama as avaliações ultrassonográficas de testículo foram executadas em sequência de quatro avaliações intercaladas por 90 dias e realizadas sempre com o mesmo equipamento, caracterizado por uma sonda linear de 7,5 MHz, em que se obteve uma imagem no plano longitudinal-lateral, na qual a densidade de pixels foi lida no software Biosoft Toolbox (Biotronics, Inc.) e a curva observada individualmente. A partir da circunferência escrotal de 19 cm realizou-se a coleta seminal por eletroejaculação, com técnica diferente da empregada em animais mais velhos, tendo sido crucial maior delicadeza e atenção nos estímulos e importante o volume coletado do ejaculado, porque a definição de idade à puberdade empregada, baseada no trabalho de Wolf et al. (1965), considerou o animal púbere quando em seu ejaculado houve motilidade progressiva ≥10% e concentração espermática total de, pelo menos, 50 milhões de espermatozoides.

Os Grupos de Contemporâneos (GC) foram formados levando em consideração o grupo de manejo, estação de nascimento, ano e fazenda. Realizou-se no Software R (R Core Team, 2020) a análise de consistência dos fenótipos, na qual foram eliminadas observações com medidas de três desvios-padrão acima ou abaixo da média e grupos de contemporâneos contendo progênies de apenas um touro ou contendo menos de cinco animais.

Os valores genéticos genômicos dos animais foram estimados utilizando diferentes modelos de predição. O modelo pode ser representado da seguinte forma:

$$\boldsymbol{Y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{e}$$

Em que: $\boldsymbol{Y}$ = vetor de observações; $\boldsymbol{X}$ = matriz de incidência para os efeitos fixos; $\boldsymbol{b}$ = vetor de efeitos fixos (GC); $\boldsymbol{Z}$ = matriz de incidência dos efeitos aleatórios; $\boldsymbol{a}$ = vetor dos efeitos aleatórios genéticos aditivos diretos dos animais; e $\boldsymbol{e}$ = vetor de efeitos residuais.

Os critérios para processar o controle de qualidade dos marcadores consideraram o teste exato de Fisher para Equilíbrio de Hardy-Weinberg (EHW), posição dos marcadores, taxa de leitura (*CallRate*) e frequência alélica mínima (MAF).

Portanto, marcadores com P_{EHW} < 0,15, posição desconhecida ou redundante no genoma, CallRate < 0,90, bem como SNPs com MAF < 0,01 e não-autossômicos, foram excluídos.

Após efetivada a consistência dos dados fenotípicos, permaneceram 11.014 animais com dados para IPM. As estatísticas descritivas, variâncias e herdabilidade da característica IPM são mostradas na Tabela 1.

TABELA 1 - Estatísticas descritivas, variâncias e herdabilidade da característica Idade à Puberdade de Machos (IPM - meses) da raça Nelore

Média	Mediana	Mín	Máx	Desvio Padrão	Coeficiente de Variação	Variância Residual (SE)	Variância Genética (SE)	Herdabilidade (SD)
17,13	16,10	8,37	22,00	3,62	21,15	7,03 (0,31)	4,19 (0,43)	0,37 (0,03)

Mín: Mínimo, Máx: Máximo, SE = SD = erro padrão aproximado.

3.2.2. Estimação do efeito dos marcadores

Os efeitos dos marcadores foram estimados a partir dos dados genômicos, levando em consideração diferentes distribuições *a priori* para os efeitos e variâncias dos SNPs. As análises genômicas utilizando os métodos bayesianos, (BayesA, BayesB, BayesC e Bayesian LASSO), foram realizadas no Programa R (R Core Team, 2020), por meio dos pacotes “BGLR” e “psych”, e, para o ssGBLUP, os programas da família BLUPF90 (Misztal et al., 2002, 2014) foram utilizados .

3.2.3. Estimação dos pseudo-fenótipos

O fenótipo ajustado para os efeitos fixos (Y*) foi utilizado como pseudo-fenótipo, considerando o GC (grupo de manejo, estação de nascimento, ano e fazenda) como efeito fixo. O coeficiente de correlação entre o pseudo-fenótipo com o DGV (valor genômico direto) foi utilizado para avaliar e comparar a acurácia da predição genômica dos modelos BayesA, BayesB, BayesC, Bayesian LASSO e ssGBLUP.

3.2.4. Estratégia de treinamento e validação

Pelo fato da estrutura dos dados da característica de IPM, em que poucos animais com progênie possuem fenótipo para a característica, optamos por realizar três cenários de análises. Os animais foram divididos em dois grupos (treinamento e validação) e as diferentes populações de treinamento (PT) e validação (PV) utilizadas na avaliação genômica para IPM são apresentadas nas Tabelas 2 e 3.

A PT 1 (Tabela 2) possui 2.000 animais que possuem fenótipo e genótipo e a PT 2 é formada por 2.039 animais escolhidos aleatoriamente de um grupo de 2.719 animais que possuem fenótipo e genótipo.

TABELA 2 – Estrutura das populações de treinamento utilizadas na avaliação genômica da característica Idade à Puberdade de Machos (IPM) da raça Nelore

População de treinamento	Número amostral	Fenótipo
PT1	2.000	IPM
PT2	2.039	IPM

PT = População de treinamento; IPM = Idade à puberdade de machos.

A PV 1 (Tabela 3) é formada por 719 animais que possuem fenótipo e genótipo e a PV 2 é formada por 680 animais escolhidos aleatoriamente de um grupo de 2.719 animais que possuem fenótipo e genótipo.

TABELA 3 – Estrutura das populações de validação utilizadas na avaliação genômica da característica Idade à Puberdade de Machos (IPM) da raça nelore

População de validação	Número amostral	Fenótipo
PV1	719	IPM
PV2	680	IPM

PV = População de validação; IPM = Idade à puberdade de machos.

Os diferentes cenários utilizados para realizar as avaliações e predições genômicas para a característica IPM são mostrados na Tabela 4.

Tabela 4 – Cenários de Treinamento e Validação usados para as predições genômicas da característica idade à puberdade de machos (IPM) da raça Nelore

Cenário	População de Treinamento	População de Validação	Predição	Correlação com DGV
1	PT1	PV1	DGV	Y*
2	PT2	PV2	DGV	Y*

PT = População de treinamento; PV = População de validação; DGV = Valor genômico direto; Y* = Fenótipo ajustado.

3.2.5. Predição dos valores genômicos

O DGV foi predito a partir dos efeitos dos SNPs estimados utilizando a seguinte equação:

$$\mathrm{DGV} = \sum_{j}^{n} X_{ij}\hat{g}_j$$

Em que: **n** é o número de SNPs, $\boldsymbol{X_{ij}}$ é o genótipo do i-ésimo animal para o j-ésimo SNP (0, 1 ou 2 para os genótipos AA, AB e BB, respectivamente) e $\boldsymbol{\hat{g}_j}$ é o vetor de efeito do j-ésimo SNP.

Os efeitos dos marcadores estimados na população de treinamento foram utilizados para calcular o DGV na população de validação, sendo que no cenário 2 foi feito uma validação cruzada com 4 *fold* e 10 repetições aleatórias.

3.2.6. Critérios de comparação dos modelos

A habilidade de predição dos métodos BayesA, BayesB, BayesC, Bayesian LASSO e ssGBLUP foi comparada por meio da correlação entre o pseudo-fenótipo (Y*) e o DGV predito nas populações de validação pelos diferentes modelos propostos no presente estudo.

O coeficiente de regressão do DGV sobre o Y* foi obtido para estudar o nível de inflação/deflação dos DGVs em relação à variável resposta utilizada, de forma que valores mais próximos de 1 (um) foram considerados menos viesados.

3.3. Resultado e discussão

As habilidades de predição e os coeficientes de regressão dos cenários realizados neste trabalho são apresentadas na Tabela 5.

No geral, não foram verificadas diferenças na habilidade de predição entre os modelos bayesianos e o ssGBLUP, entretanto o Bayesian Lasso e o ssGBLUP obtiveram uma predição mais viesada. No cenário 2, os bayesianos foram superiores na hp e menos viesados quando comparados aos resultados do cenário 1.

TABELA 5 – Habilidade de predição (hp) e o coeficiente de regressão (cr) dos cenários de acordo com diferentes modelos e populações de treinamento e validação para idade à puberdade de machos (IPM) da raça Nelore

	Bayes A		***Bayes B***		***Bayes C***		***Bayesian Lasso***		**ssGBLUP**	
Cenário	**hp**	**cr**	**Hp**	**Cr**	**hp**	**cr**	**hp**	**cr**	**Hp**	**cr**
1	0,42	0,83	0,42	0,84	0,42	0,84	0,42	0,78	0,43	0,77
2	0,50	1,02	0,49	1,02	0,50	1,01	0,49	1,04	-	-

ssGBLUP: single step Genomic Best Linear Unbiassed Predictor.

Os resultados obtidos no presente estudo diferem dos reportados por Magnabosco et al. (2020) e Silva et al. (2016), que abordaram características relacionadas à eficiência alimentar em bovinos Nelore e constataram que o procedimento do ssGBLUP obteve a predição genômica mais precisa e consistente. Ainda, Lourenco et al. (2014) encontraram que o método de um único passo (ssGBLUP) tinha uma vantagem sobre os métodos de várias etapas (Bayes C e GBLUP), principalmente pelo fato de o *single-step* utilizar fenótipos em vez de pseudo-fenótipos e considerar toda a estrutura da população na estimação dos GEBV.

No presente estudo, muitos animais que apresentam fenótipo não são genotipados e, além disso, a maioria dos touros genotipados no treinamento tem pouca ligação com o fenótipo. Esses fatores possivelmente podem limitar a expressão das vantagens descritas para o procedimento do ssGBLUP quando comparado com os modelos *multiple step*. Nesse sentido, Silva et al. (2021), comparando a habilidade de predição dos métodos BLUP, GBLUP, ssGBLUP, Bayes A, Bayes B, Bayes C e Bayesian LASSO para características indicadoras de composição de carcaça em bovinos da raça Nelore, obtiveram melhor habilidade de predição e viés das abordagens bayesianas para as características de escores visuais, no entanto, o ssGBLUP para características de carcaça obtidas por ultrassonografia apresentou melhor habilidade de predição, porém, o viés foi inferior para duas das três características avaliadas por ultrassonografia.

Na comparação dos diferentes cenários propostos, observa-se que é mais vantajoso realizar a validação cruzada (cenário 2), porque há a melhora tanto da habilidade de predição quanto do viés. Possivelmente, isto ocorre em função do alto grau de relacionamento que existe entre os animais nesta análise, demonstrando que determinar de maneira aleatória as populações de treinamento e validação pode gerar

melhores predições quando comparados aos cenários com populações definidas previamente. Nesse sentido, Saatchi et al. (2011), em um estudo com touros Angus, e Brunes et al. (2020), trabalhando com touros Nelore, concluíram que a habilidade de predição tende a ser maior quando os animais da população de validação estão relacionados aos animais da população de treinamento.

Resultados semelhantes aos encontrados no presente trabalho foram reportados por Silva et al. (2016), Brunes et al. (2020) e Magnabosco et al. (2020) que estudaram características de eficiência alimentar e encontraram maior hp no cenário da validação cruzada, dentro de um conjunto de cenários que abordavam diferentes critérios para a formação das populações de treinamento e validação (idade ou acurácia do EBV). No entanto, Brunes et al. (2020) também afirmam que a predição genômica usando validação baseada na idade ou acurácia do EBV pode ser realizada sem perda de hp para as características de alta herdabilidade, por isso, os autores recomendam que a validação cruzada seja utilizada em casos de populações formadas principalmente por animais jovens e características que possuem pouco registro de dados, que o é o caso da IPM.

Dessa maneira, os resultados obtidos neste trabalho constatam que é possível identificar diretamente os touros jovens com maior precocidade sexual, em vez de selecioná-los com base apenas na circunferência escrotal, e encorajam a utilização da característica de IPM como critério de seleção e a aplicação da seleção genômica, para ter maior confiabilidade dos valores genéticos dos touros jovens. A antecipação da puberdade sexual dos machos gerará um excelente progresso genético no rebanho, pelo fato do aumento da intensidade de seleção aplicada nos machos ser maior do que nas fêmeas, e por IPM ser uma característica correlacionada favoravelmente com a precocidade sexual das fêmeas e a permanência no rebanho das mesmas e possuir variação genética aditiva suficiente para responder à seleção (Silva Neto et al., 2020).

3.4. Conclusão

Não houve diferença na habilidade de predição entre os modelos bayesianos e ssGBLUP, entretanto os modelos Bayes A, Bayes B e Bayes C foram menos viesados.

A maior diferença na predição dos valores genômicos foi entre os cenários, sendo a validação cruzada a metodologia mais vantajosa para predizer os valores genômicos da IPM, apresentando maior habilidade de predição e menor viés.

3.5. Referências

ALBUQUERQUE, L. G.; FERNANDES JÚNIOR, G. A.; CARVALHEIRO, R. (2017) Beef cattle genomic selection in tropical environments. In: ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS CONFERENCE, 22., 2017, Townsville. **Proceedings... .** [s. L.]: AAABG. p. 255-263. Disponível em: <http://www.aaabg.org/>. Acesso em: 19 maio 2019.

BAILEY, T. L.; MONKE, D.; HUDSON, R. S.; WOLFE, D. F.; CARSON, R. L.; RIDDELL, M. G. (1996) Testicular shape and its relationship to sperm production in mature Holstein bulls. **Theriogenology**, [S.L.], v. 46, n. 5, p. 881-887. Elsevier BV. http://dx.doi.org/10.1016/s0093-691x(96)00245-2.

BALDI, F.; ALENCAR, M. M.; FREITAS, A. R.; BARBOSA, R. T. (2008) Parâmetros genéticos para características de tamanho e condição corporal, eficiência reprodutiva e longevidade em fêmeas da raça Canchim. **Revista Brasileira de Zootecnia**, [s.l.], v. 37, n. 2, p.247-253. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/s1516-35982008000200010.

BONAMY, M.; KLUSKA, S.; PERIPOLLI, E.; LEMOS, M. V. A.; AMORIM, S. T.; VACA, R. J.; LÔBO, R. B.; CASTRO, L. M.; FARIA, C. U.; FERRARI, F. B.; BALDI, F. (2018) Genetic association between different criteria to define sexual precocious heifers with growth, carcass, reproductive and feed efficiency indicator traits in Nellore cattle using genomic information. **Journal Of Animal Breeding And Genetics**, [S.L.], v. 136, n. 1, p. 15-22. Wiley. http://dx.doi.org/10.1111/jbg.12366.

BRUNES, L. C.; BALDI, F.; LOPES, F. B.; NARCISO, M. G.; LOBO, R. B.; ESPIGOLAN, R.; COSTA, M. F. O.; MAGNABOSCO, C. U. (2020) Genomic prediction ability for feed efficiency traits using different models and pseudo-phenotypes under several validation strategies in Nelore cattle. **Animal**, [S.L.], n.100085. Elsevier BV. http://dx.doi.org/10.1016/j.animal.2020.100085.

CARVALHEIRO, R. (2014) Genomic Selection in Nelore Cattle in Brazil. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 10., Vancouver. **Proceedings...** Vancouver: WCGALP, 2014. Disponível em: <http://www.wcgalp.org/proceedings/2014>. Acesso em: 25 maio 2019.

COSTA E SILVA, E. V.; QUEIROZ, V. L. D.; COSTA FILHO, L. C. C.; FERREIRA, B. X.; FORTUNATO, A. A. A. D.; SOUZA, C. C.; OLIVEIRA, V. M.; ZUCCARI, C. E. S. N.

(2017) Ultrassonografia testicular como recurso de diagnóstico na andrologia bovina. In: NOGUEIRA, E.; BORGES SILVA, J. C.; NICACIO, A. C.; MINGOTI, G. Z. (ed.). **Ultrassonografia na reprodução e avaliação de carcaças em bovinos**. Brasília, DF: EMBRAPA. Cap. 4. p. 113-133. E-book : il. color.

FORTES, M. R. S.; LEHNERT, S. A.; BOLORMAA, S.; REICH, C.; FORDYCE, G.; CORBET, N. J.; WHAN, V.; HAWKEN, R. J.; REVERTER, A. (2012) Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science**, [s.l.], v. 52, n. 3, p.143-150. CSIRO Publishing. http://dx.doi.org/10.1071/an11165.

GODDARD, M. E.; HAYES, B. J. (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. **Nature Reviews Genetics**, [s.l.], v. 10, n. 6, p.381-391. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/nrg2575.

HAYES, B. J.; LEWIN, H. A.; GODDARD, M. E. (2013) The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. **Trends In Genetics**, [s.l.], v. 29, n. 4, p.206-214. Elsevier BV. http://dx.doi.org/10.1016/j.tig.2012.11.009.

HAYES, B.; GODDARD, M. (2010) Genome-wide association and genomic selection in animal breedingThis article is one of a selection of papers from the conference “Exploiting Genome-wide Association in Oilseed Brassicas: a model for genetic improvement of major oecd crops for sustainable farming⁙.. **Genome**, [S.L.], v. 53, n. 11, p. 876-883. Canadian Science Publishing. http://dx.doi.org/10.1139/g10-076.

LOURENCO, D. A. L.; MISZTAL, I.; TSURUTA, S.; AGUILAR, I.; EZRA, E.; RON, M.; SHIRAK, A.; WELLER, J. I. (2014) Methods for genomic evaluation of a relatively small genotyped dairy population and effect of genotyped cow information in multiparity analyses. **Journal Of Dairy Science**, [S.L.], v. 97, n. 3, p. 1742-1752. American Dairy Science Association. http://dx.doi.org/10.3168/jds.2013-6916.

MAGNABOSCO, C. U.; LOPES, F.; MAGNABOSCO, V.; LOBO, R.; PEREIRA, L.; ESPIGOLAN, R.; BRUNES, L. (2020) PSXII-22 Genomic prediction accuracy for feed efficiency related traits using different pseudo-phenotypes, prediction and validation methods in Nellore cattle. **Journal Of Animal Science**, [S.L.], v. 98, n. 4, p. 245-246. Oxford University Press (OUP). http://dx.doi.org/10.1093/jas/skaa278.446.

MISZTAL, I.; TSURUTA, S.; LOURENCO, D.; AGUILAR, I.; LEGARRA, A.; VITEZICA, Z. (2014) **Manual for BLUPF90 family of programs**. University of Georgia, Athens, USA, May.

MISZTAL, I.; TSURUTA, S.; STRABEL, T.; AUVRAY, B.; DRUET, T.; LEE, D. H. (2002) **BLUPF90 and related programs (BGF90)**. Pages 1-2 in Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, August, 2002. Session 28.

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/.

SAATCHI, M.; MCCLURE, M. C.; MCKAY, S. D.; ROLF, M. M.; KIM, J.; DECKER, J. E.; TAXIS, T. M.; CHAPPLE, R. H.; RAMEY, H. R.; NORTHCUTT, S. L.; BAUCK, S.; WOODWARD, B.; DEKKERS, J. C. M.; FERNANDO, R. L.; SCHNABEL, R. D.; GARRICK, D. J.; TAYLOR, J. F. (2011) Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. **Genetics Selection Evolution**, [S.L.], v. 43, n. 1. Springer Science and Business Media LLC. http://dx.doi.org/10.1186/1297-9686-43-40.

SIDDIQUI, M. A.; BHATTACHARJEE, J.; DAS, Z. C.; ISLAM, M. M.; ISLAM, M. A.; HAQUE, M. A.; PARRISH, J. J.; SHAMSUDDIN, M. (2008) Crossbred Bull Selection for Bigger Scrotum and Shorter Age at Puberty with Potentials for Better Quality Semen. **Reproduction In Domestic Animals**, [s.l.], 43(1), p.74-79. Wiley. http://dx.doi.org/10.1111/j.1439-0531.2007.00857.x.

SILVA NETO, J. B.; PERIPOLLI, E.; COSTA E SILVA, E. V.; ESPIGOLAN, R.; NEIRA, J. D. R.; SCHETTINI, G.; COSTA FILHO, L. C. C.; BARBOSA, F. B.; MACEDO, G. G.; COSTA-BRUNES, L.; LOBO, R. B.; PEREIRA, A. S. C.; BALDI, F. (2020) Genetic correlation estimates between age at puberty and growth, reproductive, and carcass traits in young Nelore bulls. **Livestock Science**, [S.L.], v. 241. Elsevier BV. http://dx.doi.org/10.1016/j.livsci.2020.104266.

SILVA, A. E. D. F.; UNANIAN, M. M.; CORDEIRO, C. M. T.; FREITAS, A. R. (2002) Relação da circunferência escrotal e parâmetros da qualidade do sêmen em touros da raça Nelore, PO. **Revista Brasileira de Zootecnia**, [S.L.], v. 31, n. 3, p. 1157-1165. FapUNIFESP (SciELO). http://dx.doi.org/10.1590/s1516-35982002000500012.

SILVA, R. M. O.; FRAGOMENI, B. O.; LOURENCO, D. A. L.; MAGALHÃES, A. F. B.; IRANO, N.; CARVALHEIRO, R.; CANESIN, R. C.; MERCADANTE, M. E. Z.; BOLIGON, A. A.; BALDI, F. S.; MISZTAL, I.; ALBUQUERQUE, L. G. (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**, [S.L.], v. 94, n. 9, p. 3613-3623. Oxford University Press (OUP). http://dx.doi.org/10.2527/jas.2016-0401.

SILVA, R. P.; ESPIGOLAN, R.; BERTON, M. P.; LÔBO, R. B.; MAGNABOSCO, C. U.; PEREIRA, A. S. C.; BALDI, F. (2021) Genomic prediction ability for carcass

composition indicator traits in Nellore cattle**. Livestock Science**, [S.L.], v. 245. Elsevier BV. http://dx.doi.org/10.1016/j.livsci.2021.104421.

STAFUZZA, N. B.; COSTA E SILVA, E. V.; SILVA, R. M. O.; COSTA FILHO, L. C. C.; BARBOSA, F. B.; MACEDO, G. G.; LÔBO, R. B.; BALDI, F. (2019) Genome-wide association study for age at puberty in young Nelore bulls. **Journal Of Animal Breeding And Genetics**, v. 137, n. 2, p. 234-244. Wiley. http://dx.doi.org/10.1111/jbg.12438.

VAN EENENNAAM, A. L.; WEIGEL, K. A.; YOUNG, A. E.; CLEVELAND, M. A.; DEKKERS, J. C. M. (2014) Applied Animal Genomics: Results from the Field. **Annual Review Of Animal Biosciences**, [s.l.], v. 2, n. 1, p.105-139. Annual Reviews. http://dx.doi.org/10.1146/annurev-animal-022513-114119.

WOLF, F. R.; ALMQUIST, J. O.; HALE, E. B. (1965) Prepuberal Behavior and Puberal Characteristics of Beef Bulls on High Nutrient Allowance. **Journal Of Animal Science**, [s.l.], v. 24, n. 3, p.761-765. Oxford University Press (OUP). http://dx.doi.org/10.2527/jas1965.243761x.