

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**APLICAÇÃO DE FERRAMENTAS GENÔMICAS NA
AVALIAÇÃO GENÉTICA DA IDADE AO PRIMEIRO PARTO
EM BÚFALAS DA RAÇA MURRAH**

Jessica Cristina Gonçalves dos Santos
Zootecnista

2023

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**APLICAÇÃO DE FERRAMENTAS GENÔMICAS NA
AVALIAÇÃO GENÉTICA DA IDADE AO PRIMEIRO PARTO
EM BÚFALAS DA RAÇA MURRAH**

Jessica Cristina Gonçalves dos Santos

Orientador: Prof. Dr. Humberto Tonhati

Coorientador: Dr. Francisco Ribeiro de Araujo Neto

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

2023

S237a

Santos, Jessica Cristina Gonçalves dos

Aplicação de ferramentas genômicas na avaliação genética da idade ao primeiro parto em búfalas da raça Murrah / Jessica Cristina Gonçalves dos Santos. -- Jaboticabal, 2023
59 p. : il., tabs.

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

Orientador: Humberto Tonhati

Coorientador: Francisco Ribeiro de Araujo Neto

1. Melhoramento Genético Animal. 2. Búfalos. 3. Interação
genótipo-ambiente. 4. Genômica. I. Título.

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

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: APLICAÇÃO DE FERRAMENTAS GENÔMICAS NA AVALIAÇÃO GENÉTICA DA IDADE AO PRIMERO PARTO EM BÚFALAS DA RAÇA MURRAH

AUTORA: JÉSSICA CRISTINA GONÇALVES DOS SANTOS

ORIENTADOR: HUMBERTO TONHATI

COORDINADOR: FRANCISCO RIBEIRO DE ARAUJO NETO

Aprovada como parte das exigências para obtenção do Título de Doutora em Genética e Melhoramento Animal, pela Comissão Examinadora:

Prof. Dr. FRANCISCO RIBEIRO DE ARAUJO NETO (Participação Virtual)
Departamento de Zootecnia / Instituto Federal Goiano - Rio Verde/GO



Documento assinado digitalmente

FRANCISCO RIBEIRO DE ARAUJO NETO
Data: 29/05/2023 11:13:34-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. JOSINEUDSON AUGUSTO II DE VASCONCELOS SILVA (Participação Virtual)
Departamento de Melhoramento e Nutrição Animal / FMVZ UNESP Botucatu



Documento assinado digitalmente

JOSINEUDSON AUGUSTO II DE VASCONCELOS
Data: 29/05/2023 19:21:38-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. LEONARDO DE OLIVEIRA SENO (Participação Virtual)
Departamento de Zootecnia / Universidade Federal da Grande Dourados - Dourados



Documento assinado digitalmente

LEONARDO DE OLIVEIRA SENO
Data: 29/05/2023 14:10:35-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. RUSBEL RAUL ASPILCUETA BORQUIS (Participação Virtual)
Departamento de Zootecnia / Universidade Tecnológica Federal do Paraná - Dois Vizinhos



Documento assinado digitalmente

RUSBEL RAUL ASPILCUETA BORQUIS
Data: 29/05/2023 17:36:05-0300
Verifique em <https://validar.iti.gov.br>

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



Documento assinado digitalmente

HENRIQUE NUNES DE OLIVEIRA
Data: 29/05/2023 10:18:28-0300
Verifique em <https://validar.iti.gov.br>

Jaboticabal, 26 de maio de 2023

DADOS CURRICULARES DO AUTOR

Jessica Cristina Gonçalves dos Santos – nascida na cidade de Canarana-MT em 31 de janeiro de 1995, filha de Nelson Gonçalves Constante e Marcia Terezinha dos Santos. Ingressou no curso de Zootecnia da Faculdade de Ciências Agrárias da Universidade Federal da Grande Dourados (UFGD) no primeiro semestre de 2012, concluindo-o no segundo semestre de 2016. Em março de 2017, ingressou no Programa de pós-graduação em Zootecnia da UFGD, área de concentração em produção animal, concluindo-o em março de 2019. No segundo semestre de 2019, ingressou no curso de Doutorado no Programa de pós-graduação em Genética e Melhoramento Animal da Faculdade de Ciências Agrárias e Veterinárias – UNESP – Campus Jaboticabal, sob orientação do Prof. Dr. Humberto Tonhati e coorientação do Prof. Dr. Francisco Ribeiro de Araujo Neto. Sendo bolsista CAPES no período de agosto de 2019 a agosto de 2022.

AGRADECIMENTOS

Primeiramente agradeço a Deus, pelo dom da vida, por permitir e me capacitar para que eu chegasse até aqui, agradeço pela Graça recebida. Agradeço a Nossa Senhora que nos momentos de ansiedade sempre recorri ao seu amor de mãe que intercede ao Pai por nós.

Agradeço ao meu marido Nélio, pelo amor, zelo e companheirismo. Foi minha rede de apoio para que eu chegasse até aqui.

Agradeço aos meus pais, Marcia e Nelson por terem me educado e me preparado para a vida e por serem minha base. Sem a educação, o preparo e o incentivo deles com certeza eu não estaria aqui.

Agradeço aos amigos que eu fiz durante esse percurso: Karina, Rafael, Gabriela e Cher, com quem eu tive o privilégio de conviver parte do doutorado.

Agradeço ao professor Leonardo de Oliveira Seno, meu orientador de mestrado que me incentivou a fazer o doutorado.

Agradeço ao meu orientador Humberto Tonhati, pela oportunidade de ser sua orientada no doutorado.

Agradeço ao meu coorientador Francisco, sempre solícito, com certeza contribuiu muito para que eu chegasse até aqui, mesmo com aulas, sempre estive disposto a se reunir via google meet para me orientar nas análises e escritas, muitas vezes aos domingos e feriados. Agradeço imensamente.

Agradeço a CAPES que concedeu a bolsa de estudos durante 3 anos permitindo assim a minha permanência no doutorado

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

SUMÁRIO

RESUMO	III
ABSTRACT	IV
CAPÍTULO 1 – CONSIDERAÇÕES GERAIS.....	1
1.1 INTRODUÇÃO	1
1.2 OBJETIVO GERAL	2
1.3 REVISÃO DE LITERATURA	3
1.3.1 Importância da idade ao primeiro parto em búfalos	3
1.3.2 Interação genótipo-ambiente (IGA) via modelo de norma de reação (RNM).....	4
1.3.3 Aplicação da seleção genômica em búfalos	6
1.3.3.1 Princípios da seleção genômica	6
1.3.3.2 Seleção genômica no melhoramento genético de búfalos	8
1.3.4 RNM Genômica	10
1.4 REFERÊNCIAS	10
CAPÍTULO 2 – ANÁLISE GENÔMICA EM MODELOS DE NORMAS DE REAÇÃO PARA IDADE AO PRIMEIRO PARTO EM BÚFALAS MURRAH.....	16
2.1 INTRODUÇÃO	18
2.2 MATERIAL E MÉTODOS	19
2.2.1 Banco de dados e consistências	19
2.2.2 Estimação dos componentes de variância da norma de reação (RNM)	21
2.2.2.1 Gradiente ambiental (GA).....	21
2.2.2.2 Norma de reação tradicional (BLUP) e genômica (ssGBLUP)	22
2.2.3 Habilidade de predição dos métodos genômico e tradicional	23
2.2.4 Estudo de associação genômica ampla (GWAS) da interação genótipo-ambiente (IGA)	24
2.3 RESULTADOS	24
2.3.1 Componentes de variância e parâmetros genéticos da norma de reação (RNM)	24
2.3.2 Predição genômica das normas de reação (RNM)	26
2.3.3 Análise de associação genômica ampla (GWAS) para plasticidade (inclinação)	27
2.4 DISCUSSÃO	30
2.4.1 Componentes de variância obtido via norma de reação linear (RNM)	30
2.4.2 Predição genômica da norma de reação (RNM)	32
2.4.3 GWAS do parâmetro de inclinação do RNM	32
2.5 CONCLUSÃO.....	36
2.6 REFERÊNCIAS	37
CAPÍTULO 3 – COMPARAÇÃO DE METÓDOS SINGLE-STEP PARA A PREDIÇÃO GENÔMICA DA IDADE AO	

PRIMEIRO PARTO EM BÚFALOS LEITEIROS	42
3.1INTRODUÇÃO	44
3.2MATERIAL E MÉTODOS	45
3.2.1 Dados.....	45
3.2.2 Modelos PBLUP, ssGBLUP e WssGBLUP	46
3.2.3 Modelos de regressão Bayesiana em passo único (ssBR).....	48
3.2.4 Validação e análise de GEBVs.....	49
3.3RESULTADOS	50
3.4DISCUSSÃO	52
3.5CONCLUSÃO.....	56
3.6REFERÊNCIAS	56

APLICAÇÃO DE FERRAMENTAS GENÔMICAS NA AVALIAÇÃO GENÉTICA DA IDADE AO PRIMEIRO PARTO EM BÚFALAS DA RAÇA MURRAH

RESUMO - A idade ao primeiro parto (IPP) é uma das características reprodutivas mais importante em um rebanho leiteiro, pois está associada ao início da vida produtiva das fêmeas. No entanto, a IPP tem estimativa de herdabilidade baixa, com maior parte da variância fenotípica composta por fatores ambientais. O ambiente é um fator determinante para a expressão e funcionalidade dos genes, e a interação entre os genes e o ambiente são responsáveis pelas variações fenotípicas das características. Assim, é importante que o efeito da interação genótipo-ambiente (IGA) seja incluído em avaliações genéticas para identificar a sensibilidade dos animais em uma ampla gama de condições ambientais. Uma das formas de aumentar o desempenho reprodutivo das fêmeas, é por meio da seleção de animais robustos, e por meio da compreensão da base molecular subjacente à IGA da IPP. Além disso, a inclusão de informações genômicas pode contribuir para maiores ganhos genéticos de características de baixa herdabilidade como a IPP, pois fornecem avaliações genéticas mais acuradas e menos tendenciosas que os modelos tradicionais. Existem diferentes métodos que podem ser utilizados para a predição genômica, no entanto, estes podem diferir na capacidade preditiva dos valores genômicos. Dessa forma, os objetivos deste estudo foram: i) avaliar a IGA da IPP por meio do modelo de norma de reação (RNM); ii) avaliar a habilidade de predição do RNM obtidos via BLUP e ssGBLUP; iii) identificar genes relacionados a IGA da IPP e iv) avaliar a acurácia e a dispersão das predições genômicas para a IPP em búfalos Murrah utilizando PBLUP, ssGBLUP, WssGBLUP e ssBR (ssBA, ssBB π , ssBC π , ssBL e ssBRR). Foi observado um efeito importante da IGA na IPP, com incremento de acurácia nos valores genéticos preditos para o intercepto e inclinação da RNM quando foram incluídas informações genômicas. Foram identificadas regiões genômicas importantes decorrentes dos efeitos de IGA sobre a IPP. Após análise funcional, foram detectados cinco genes principais, envolvidos em dois processos biológicos ligados ao metabolismo de lipídios e a imunidade. Os métodos de predição indicaram que a utilização de informação genômica pode melhorar o ganho genético para a IPP, aumentando a acurácia e reduzindo a inflação/deflação das predições desta característica em comparação com o modelo tradicional baseado em pedigree. Além disso, entre todos os modelos genômicos de passo único estudados, o WssGBLUP e o ssBA foram os métodos mais vantajosos a serem utilizados na avaliação genômica da IPP de búfalos desta população.

Palavras-chave: alfabeto bayesiano, *Bubalus bubalis*, frequentista, interação genótipo-ambiente, reprodução, SNP

APPLICATION OF GENOMIC TOOLS IN THE GENETIC EVALUATION OF AGE AT FIRST CALVING IN BUFFALO BREED MURRAH

ABSTRACT – Age at first calving (AFC) is one of the most important reproductive traits in a dairy herd, since it is associated with the beginning of the productive life of females. However, AFC has a low heritability estimate, with most of the phenotypic variance composed of environmental factors. The environment is a determining factor for the expression and functionality of genes and the interaction between genes and environment are responsible for the phenotypic variations of traits. Thus, it is important that the effect of genotype-environment interaction (GEI) be included in genetic evaluations to identify the sensitivity of animals over a wide range of environmental conditions. One way to increase reproductive performance of females, is by selecting robust animals, and by understanding the molecular basis underlying the GEI of AFC. In addition, the inclusion of genomic information can contribute to greater genetic gains for traits of low heritability such as AFC, as they provide more accurate and less biased genetic evaluations than traditional models. There are different methods that can be used for genomic prediction. However, these may differ in the predictive ability of genomic values. Thus, the objectives of this study were: (i) evaluate the GEI of AFC via reaction norm model (RNM); (ii) evaluate the prediction ability of the RNM obtained via BLUP and ssGBLUP; (iii) identify genes related to GEI of AFC and iv) evaluate the accuracy and dispersion of genomic predictions for AFC in Murrah buffaloes using BLUP, ssGBLUP, WssGBLUP and ssBR (ssBA, ssBB π , ssBC π , ssBL and ssBRR). An important effect of GEI on AFC was observed, with increased accuracy in the predicted genetic values for the RNM intercept and slope when genomic information was included. Important genomic regions resulting from the effects of GEI on AFC were identified. After functional analysis, five major genes were detected, involved in two biological processes linked to lipid metabolism and immunity. The prediction methods indicated that the use of genomic information can improve the genetic gain for AFC by increasing the accuracy and reducing the inflation/deflation of the predictions of this trait compared to the traditional pedigree-based model. Furthermore, among all the single-step genomic models studied, WssGBLUP and ssBA were the most advantageous methods to be used in the genomic evaluation of AFC of buffaloes in this population.

Keywords: Bayesian alphabet, *Bubalus bubalis*, frequentist, genotype-environment interaction, breeding, SNP

CAPÍTULO 1 – Considerações gerais

1.1 Introdução

Os búfalos domésticos (*Bubalus bubalis bubalis*) tem um papel importante na produção leiteira, pois representam 15% do leite produzido mundialmente, sendo a segunda espécie com maior contribuição para a produção (Cesarani et al., 2021). No entanto, a produção leiteira é dependente da performance reprodutiva dos animais (Santos et al., 2020). Entre as características reprodutivas, a idade ao primeiro parto (IPP) é uma medida utilizada para avaliar a precocidade sexual das fêmeas bubalinas (Araujo Neto et al., 2020).

A precocidade sexual dos búfalos tem influência do genótipo, da nutrição, dos manejos reprodutivos, e de fatores endógenos (fisiológicos e hormonais) da espécie (Perera et al., 2011). Os búfalos são poliétricos estacionais, apresentando maior fertilidade quando há redução da duração do dia, como as estações de outono e inverno (D’Occhio et al., 2020). O estro das fêmeas não é evidente como em fêmeas bovinas, visto que não apresentam comportamento homossexual, há pequena ou nenhuma liberação de muco pela vulva e o cio geralmente é noturno (Vale e Ribeiro 2005). Estas características particulares dos búfalos levam a falhas na identificação do cio, e contribuem para o início tardio da atividade reprodutiva da espécie (Perera et al., 2011). Todos esses fatores, levam a definição errônea dada aos búfalos de animais de baixa fertilidade e de idade a puberdade tardia.

Apesar da importância da IPP para o início da atividade produtiva, esta não é incluída como critério de seleção, principalmente devido ao maior interesse na seleção de características produtivas (leite, gordura, proteína, entre outras) que são relacionadas ao rendimento do queijo *Mozzarella*, que é mais valorizado comercialmente (Fernandes et al., 2016).

Como há muitos fatores endógenos e exógenos que atuam sobre a IPP, a avaliação genética da IPP considerando a interação genótipo-ambiente (IGA) pode contribuir para a identificação de animais específicos para cada ambiente de produção, favorecendo dessa forma a precocidade sexual dos rebanhos bubalinos.

Apesar do efeito da IGA não ser incluído rotineiramente nas avaliações

genéticas, esta pode ser uma estratégia adotada para a seleção dos animais de forma mais acurada, pois considera a diversidade dos sistemas de produção (Freitas et al., 2019). Para isso, os valores genéticos da IPP podem se estimados considerando uma ampla gama de condições ambientais, por meio de modelos de norma de reação (RNM) via regressão aleatória (Mota et al., 2020a).

Além disso, com a inclusão de informações genotípicas nas avaliações genéticas, é possível obter maior confiabilidade nas estimativas dos valores genéticos de características de herdabilidade baixa, como a IPP. Na seleção genômica, diferentes metodologias têm sido aplicadas para a estimação dos valores genéticos genômicos, bem como estudos de comparações entre as metodologias (VanRaden, 2008; Aguilar et al., 2010; Wang et al., 2012; Fernandes et al., 2016). E tem como principal objetivo a obtenção de estimativas de maior acurácia para a seleção de animais jovens. Além da redução de custos com testes de progênie, a seleção genômica proporciona redução dos intervalos de geração e maiores taxas de ganho genético (Roos et al., 2011; Scott et al., 2021; Sousa et al., 2021).

A inclusão de dados genotípicos permite, além da estimação dos valores genéticos, a possibilidade de detectar regiões genômicas e genes por meio de estudos de associação genômica ampla (*genome-wide association study* - GWAS) e dessa forma, a identificação de mecanismos subjacentes que interferem em características importantes como a IPP. Dessa forma, metodologias como a IGA, predição genômica e estudos de GWAS apresentam-se como alternativas para avaliar características muito influenciadas por fatores ambientais como a IPP em búfalos da raça Murrah.

1.2 Objetivo geral

O objetivo deste estudo foi quantificar a IGA e identificar genes candidatos via norma de reação e GWAS para a IPP de búfalos. Além disso, avaliar diferentes métodos de predição genômica de passo único para identificar o modelo que proporciona maior acurácia para a estimação dos valores genéticos da IPP em búfalos.

1.3 Revisão de literatura

1.3.1 Importância da idade ao primeiro parto em búfalos

A idade a puberdade dos búfalos varia entre 18 a 46 meses e entre 250 a 400 kg do peso corporal. Portanto, há uma grande variação no início da vida reprodutiva da espécie, com grande parte dos rebanhos com idade a puberdade tardia (Warriach et al., 2015). A fertilidade dos búfalos é muito influenciada por fatores nutricionais e pela dificuldade de detecção do estro (estro silencioso), além dos búfalos serem poliéstricos estacionais de dias curtos, ou seja, concentram o período reprodutivo em estações com menor quantidade de luz durante o dia (outono, inverno) (Silva et al., 2021).

Protocolos de sincronização do estro e ovulação têm sido implementados em búfalos para aumentar a eficiência reprodutiva e os resultados obtidos para a taxa de prenhez variaram de 30 a 50%. Esses resultados são próximos a taxas de prenhez natural, e, portanto, a eficiência dessa biotécnica foi considerada baixa para búfalas (Warriach et al., 2015).

Apesar dos desafios para reduzir a IPP dos búfalos, a inclusão da IPP como critério de seleção pode contribuir para obtenção de fêmeas mais precoces e para redução de abates involuntários (Aspilcueta-Borquis et al., 2022).

Poucos estudos estimaram parâmetros genéticos para a IPP em búfalos e na maior parte dos estudos, as estimativas de herdabilidade foram baixas (Tabela 1). É um indicativo de que maior parte da variância fenotípica é atribuída a efeitos não genéticos (ambiente, manejo nutricional e reprodutivo, entre outros), e que a resposta a seleção da característica é lenta (Shao et al., 2021). No entanto, a IPP tem um valor econômico muito importante em uma fazenda, pois está relacionada ao início da vida reprodutiva e produtiva das fêmeas (Santos et al., 2020). Dessa forma, a IPP pode ser integrada como um critério de seleção para reduzir a idade da fêmea ao primeiro parto (Shao et al., 2021).

Amin et al. (2021) obtiveram correlações genéticas moderadas (0,30 – 0,37) entre IPP, intervalo de partos (IP), intervalo pós-parto até à prenhez (IPPG - dias abertos) e escore de parto (EP) em búfalas do Egito. Os resultados indicaram que as

fêmeas com maior IPP concebem e produzem bezerros de forma menos eficiente durante a vida reprodutiva, e, que primíparas com maior IPP apresentam maiores chances de partos distócicos, devido ao maior peso do bezerro. Em búfalas Murrah, Vilela et al. (2020) identificaram que além da seleção para a produção de leite aumentar a IPP e o intervalo de partos (IP), também afetou a qualidade do leite, diminuindo a porcentagem de gordura e proteína ao longo dos anos.

Tabela 1. Estimativas de herdabilidade (h^2) para idade ao primeiro parto (IPP) em diferentes raças de búfalos

Raças	h^2 da IPP	Bibliografia
Murrah	0,07	Seno et al. (2007)
	0,17	Camargo et al. (2015)
	0,28	Kumar et al. (2015)
	0,16	Barros et al. (2016)
	0,19	Freitas et al. (2019)
	0,16	Araujo Neto et al. (2020)
	0,12	Amim et al. (2021)
Búfalos da Colômbia	0,14	Agudelo-Gómez et al. (2015)
Nili-Ravi	0,36	Tamboli et al. (2021)
Búfalos do Egito	0,12	Salem et al. (2021)

1.3.2 Interação genótipo-ambiente (IGA) via modelo de norma de reação (RNM)

O Brasil tem uma grande variabilidade climática, onde predominam três tipos de clima (equatorial, tropical e temperado), sendo o clima tropical (clima quente com chuvas menos regulares) com domínio sobre maior parte do território nacional (Silveira Marinho et al., 2020). Além disso, o Brasil apresenta estações (sazonalidade) bem definidas de acordo com as regiões (norte, nordeste, sul, sudeste e centro-oeste), o que influencia sobre a disponibilidade e qualidade das forragens, fazendo com que fazendas adotem diferentes práticas de manejos nutricionais em cada período de cada ano. O efeito dessa diversidade, determinada pela sazonalidade, interfere sobre a eficiência reprodutiva dos animais (Pessoa et al., 2010). Dessa forma, tem-se verificado em estudos com bovinos e búfalos a importância de avaliar a interação genótipo-ambiente (IGA) em características reprodutivas (Chiaia et al., 2015;

Ambrosini et al., 2016; Santana Jr. et al., 2017; Araujo Neto et al., 2018; Freitas et al., 2019), uma vez que o fenótipo é determinado por fatores genéticos e ambientais e pela IGA.

Diferentemente dos modelos de norma de reação (RNM), o modelo animal padrão não considera o efeito da IGA na estimação dos valores genéticos dos animais. Quando o efeito da IGA não é considerado, as estimativas dos valores genéticos estimados (VGE) podem ser tendenciosas e menor ganho genético pode ser obtido com a seleção (Streit et al., 2012). No RNM, o VGE é obtido considerando a IGA para diferentes níveis ambientais e contribui para que as respostas a seleção sejam mais satisfatórias, por meio da seleção de animais robustos (Santos et al., 2020).

A interação genótipo-ambiente (IGA), segundo Falconer (1952), é resultado da mudança dos desempenhos de um conjunto de genótipos quando estes são expostos a ambientes distintos. Dessa forma, um animal pode apresentar diferentes méritos genéticos quando avaliado em diferentes ambientes e as progênes dos touros avaliados podem apresentar grandes desvios da média da DEP (diferença esperada na progênie) ou da PTA (habilidade prevista de transmissão) de acordo com o ambiente em que as progênes serão produzidas (Santos et al., 2020). Portanto, conhecer os efeitos da IGA é uma alternativa para a seleção mais eficiente de indivíduos menos sensíveis a ambientes distintos (Falconer e Mackay, 1996) e dessa forma é possível obter maior resposta a seleção (Mulder e Bijma, 2005). Em outras palavras, os estudos de IGA descrevem qual é o comportamento da expressão dos genótipos dos animais quando diferentes ambientes são considerados (Santos et al 2020).

Uma das formas de avaliar o efeito da IGA é por meio da correlação genética da mesma característica (desempenho) em diferentes ambientes (Falconer, 1952), quando as correlações são inferiores a 0,80 há IGA (Robertson, 1959) e nesse caso cada fazenda deve contar um programa de avaliação genética específico ou deve-se selecionar touros que apresentem menor plasticidade fenotípica. Pois, quando a correlação é inferior a 0,80 é uma evidência de que há mudanças na atuação dos genes ou mudanças nos genes responsáveis pela variância genética aditiva da característica e, por este motivo caracteriza IGA (Chiaia et al., 2015).

Outra forma de avaliar a IGA, amplamente utilizada, é por meio de modelos de

norma de reação (RNM) (Chiai et al., 2015; Ambrosini et al., 2016; Oliveira et al., 2018; Araujo Neto et al., 2018; Freitas et al., 2019; Carvalheiro et al., 2019; Shi et al., 2021; Carvalho Filho et al., 2022). A RNM permite descrever linearmente o VGE dos animais em um gradiente ambiental contínuo, em que são obtidos VGE considerando a independência do ambiente (intercepto) e a dependência do ambiente (inclinação) (Shi et al., 2021). Diferentes modelos de norma de reação foram utilizados para identificar IGA em animais de produção, entre eles, modelos lineares e não lineares (Freitas et al., 2019; Carvalheiro et al., 2019; Shi et al., 2021; Carvalho Filho et al., 2022). Nos modelos de RNM, há diferentes pressuposições da variância residual, no RNM homogêneo, assume-se que a variância residual é homogênea e no RNM heterogêneo, assume-se que a variância residual é heterogênea. Nos modelos lineares, a regressão linear modela a norma de reação dos efeitos genéticos aditivos e residuais. E, nos modelos não lineares, a regressão quadrática polinomial modela a norma de reação dos efeitos genéticos aditivos e residuais (Carvalho Filho et al., 2022).

A avaliação do efeito da IGA sobre a IPP foi estudada com maior frequência na literatura e tem indicado a importância de considerar esse efeito em modelos de estimação de parâmetros genéticos para aumentar a eficiência reprodutiva dos animais considerando a variedade de ambientes e de sistemas de produção (Mota et al., 2020a). Em búfalas da raça Murrah, Freitas et al. (2019) avaliando IPP identificaram presença da IGA, pois houve diferenças nas estimativas de herdabilidade com a melhoria das condições ambientais, além disso, a correlação entre o intercepto e a inclinação indicaram reclassificação das estimativas dos valores genéticos dos animais.

1.3.3 Aplicação da seleção genômica em búfalos

1.3.3.1 Princípios da seleção genômica

A disponibilidade de painéis de alta densidade com custos mais acessíveis comparada a metodologia assistida por marcadores (microsatélites, RFLPs) tornou possível a genotipagem de um número expressivo de animais (Meuwissen et al., 2016;

Lourenço et al., 2020). Essas informações genotípicas (marcadores SNP) foram então implementadas em modelos estatísticos para predição do mérito genético dos animais e é denominada seleção genômica ou predição genômica (VanRaden, 2008).

A descoberta dos polimorfismos de nucleotídeo único (SNPs) como sendo os principais responsáveis pela variação da sequência do genoma foi fundamental para o desenvolvimento de metodologias para a implementação da genômica na avaliação genética de bovinos (Lourenço et al., 2020).

O BLUP genômico (GBLUP) foi a primeira metodologia desenvolvida para a incorporação de SNPs nas avaliações genéticas. Diferentemente do BLUP que utiliza o parentesco via pedigree para a predição dos VGEs, o GBLUP, inclui apenas o parentesco genômico (proporção de alelos compartilhados) (Lourenço et al., 2020). No entanto essa metodologia apresentava uma limitação, pois os valores genéticos genômicos eram obtidos apenas para animais genotipados (Christensen e Lund, 2010). Para contornar essa limitação, Aguilar et al. (2010) e Christensen e Lund, (2010) implementaram o método de estimação de valores genéticos genômicos por meio da combinação de fenótipos, genótipos e pedigree em uma única etapa (ssGBLUP). No ssGBLUP, o parentesco dos animais do pedigree é corrigido pelas informações genômicas disponíveis dos animais genotipados e as estimativas dos valores genéticos genômicos são mais acuradas (Gao et al., 2012). O ssGBLUP tornou-se o principal método utilizado para a estimação de valores genéticos devido a facilidade de integrar os dados fenotípicos, genotípicos e de pedigree em um único passo (Lourenço et al., 2020).

A seleção genômica assume que todos os marcadores podem estar ligados a genes que afetam características economicamente importantes e concentra-se na estimação dos seus efeitos (Meuwissen et al., 2016). A abordagem da seleção genômica, baseia-se na expectativa de que há QTLs (*Quantitative Trait Loci*) em desequilíbrio de ligação (DL) com um ou mais marcadores (Alexandre et al., 2021).

Uma das vantagens mais evidentes da seleção genômica é a possibilidade da seleção acurada de animais jovens, que ainda não podem ser avaliados em testes de progênie (Lopes et al., 2021). No entanto, a implementação da seleção genômica é dependente da acurácia das predições genômicas. A acurácia é influenciada por muitos fatores, entre eles, a população referência, o número de indivíduos

genotipados, o número de marcadores e a herdabilidade da característica, bem como o parentesco entre a população de referência a população de validação (Calus, 2010).

A implementação da genômica deve ser feita com muita cautela, pois as estimativas dos valores genômicos podem ser superestimados ou subestimados. Um dos métodos utilizados para avaliar a capacidade de predição dos modelos, é a validação cruzada obtida por meio da regressão linear (LR), essa metodologia compara as estimativas dos valores genéticos da população de referência (dados totais) com as estimativas dos valores genéticos obtidos da população de validação (dados parciais) (Legarra e Reverter, 2018). Esses procedimentos são uma série de estatísticas, que comparam as acurácias populacionais e os vieses das predições genômicas.

Diferentes métodos foram testados quanto a habilidade preditiva dos valores genéticos genômicos em termos de acurácia, viés e dispersão (Gao et al., 2012; Gautason et al., 2021) com a finalidade de comparar a capacidade preditiva dos modelos e reduzir a estimação tendenciosa dos valores genéticos de animais jovens e dessa forma maximizar o progresso genético (Legarra e Reverter, 2018).

Os principais benefícios da inclusão da seleção genômica são: aumento dos ganhos genéticos, redução do intervalo de gerações e aumento da acurácia dos VGEs (Christensen e Lund, 2010). Além disso, a incorporação da genômica para a estimação de parâmetros genéticos permitiu a implementação de estudos de GWAS para a identificação de marcadores e regiões genômicas associadas a características complexas (Li et al., 2018; Liu et al., 2018).

1.3.3.2 Seleção genômica no melhoramento genético de búfalos

O desenvolvimento do painel SNP de 90K para búfalos-de-rio (Axiom® Buffalo Genotyping Array 90K) permitiu estudos sobre a diversidade genética, estudos de GWAS e estudos de predição genômica em búfalos (Iamartino et al., 2017).

O primeiro estudo de GWAS utilizando o painel desenvolvido para búfalos, foi realizado em animais da raça Mediterrâneo (Iamartino et al., 2017), no entanto, os cromossomos e a posição de todos os SNPs foram baseados na sequência do genoma do *Bos taurus* (UMD 3.1), visto que, o genoma de referência do búfalo ainda

não estava disponível, além disso, foram utilizadas as informações de anotação gênica dos bovinos.

Posteriormente, o genoma de referência do búfalo (UOA_WB_1) tornou-se disponível (Low et al., 2019), com a última atualização disponibilizada pela National Dairy Development Board (NDDB) em novembro de 2021 (NDDB_SH_1). Desde a possibilidade da implementação do GWAS, foram desenvolvidos estudos para diferentes características em búfalos (Camargo et al., 2015; Li et al., 2018; Liu et al., 2018; Araujo Neto et al., 2018; Araujo Neto et al., 2020; Araujo Neto et al., 2022).

Muitos são os métodos disponíveis para a implementação genômica, entre os métodos single-step, estão o ssGBLUP, WssGBLUP (GBLUP ponderado de passo único) e o ssBR (regressão bayesiana de passo único).

O ssGBLUP permite o uso simultâneo de informações genômicas, pedigree e fenotípicas e pode reduzir possíveis vieses na estimação dos valores genéticos. O WssGBLUP, por sua vez, é um método iterativo, onde diferentes pesos são atribuídos aos SNPs (Wang et al., 2012).

Fernandes et al., (2016) propuseram o modelo de regressão bayesiana em passo único (ssBR), que é fundamentado nos efeitos dos marcadores (MEM) e no ajuste dos valores genéticos em termos dos efeitos de genótipos observados ou imputados. Dessa forma, os modelos ssBR combinam todos os dados disponíveis de animais genotipados e não genotipados e utilizam covariáveis de marcadores imputados para animais que não são genotipados. Esses modelos acomodam diferentes pressuposições *a priori* do alfabeto Bayesiano (BayesA, BayesB π , BayesC π , Bayes Lasso, BayesRR, entre outros), sendo está a principal diferença entre os métodos.

Dada a importância dos búfalos em países em desenvolvimento e sua alta contribuição para a produção de leite é essencial entender a arquitetura genética de características econômicas importantes e utilizar a seleção genômica para acelerar o melhoramento genético em programas de melhoramento genético de búfalos (Liu et al., 2017).

1.3.4 RNM Genômica

Com o advento da genômica e com a disponibilidade de informação genômica para diferentes espécies de animais de produção, foi possível a inclusão de informações genômicas nas análises de IGA. E tornou-se possível a detecção de regiões genômicas por meio de estudos de GWAS do modelo de norma de reação, considerando diferentes condições ambientais (Mota et al., 2020a).

Oliveira et al. (2018) avaliando a IGA para peso ao ano em bovinos Nelore descreveram maior acurácia utilizando o modelo ssGBLUP em relação ao modelo tradicional (BLUP). Também em Nelore, Mota et al. (2020b) obtiveram estimativas de parâmetros genéticos mais acuradas quando incluíram genômica na análise de RNM para identificar a IGA em características indicadoras de precocidade sexual. Avaliando a IGA por meio de GWAS para a IPP em bovinos Nelore, Mota, et al. (2020a) descreveram genes-chave relacionados a sinalização hormonal e ao metabolismo energético. Dessa forma, a inclusão de informações de marcadores moleculares nos modelos estatísticos para a estimação de VGE pode proporcionar maior acurácia e possibilitar a detecção das principais regiões genômicas que explicam a variação fenotípica entre ambientes da IPP.

1.4 Referências

Agudelo-Gómez DA, Savegnago RP, Buzanskas ME, Ferraudo AS, Munari DP, Cerón-Munoz MF (2015) Genetic principal components for reproductive and productive traits in dual-purpose buffaloes in Colombia. **Journal of Animal Science** 93:3801–3809. <https://pubmed.ncbi.nlm.nih.gov/26440159/>

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ (2010) Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. **Journal of Dairy Science** 93:743–752. <https://doi.org/10.3168/jds.2009-2730>

Alexandre PA, Li Y, Hine BC, Duff CJ, Ingham AB, Porto-Neto LR, Reverter A (2021) Bias, dispersion, and accuracy of genomic predictions for feedlot and carcass traits in Australian Angus steers. **Genetics Selection Evolution** 53. <https://doi.org/10.1186/s12711-021-00673-8>

Ambrosini DP, Malhado CHM, Filho RM, Cardoso FF, Carneiro PLS (2016) Genotype × environment interactions in reproductive traits of Nellore cattle in northeastern Brazil. **Tropical Animal Health and Production** 48:1401–1407. <https://doi.org/10.1007/s11250-016-1105-7>

Amin AMS, Abo-Ismael MK, Salem MMI (2021) Genetic parameters and genetic trends for reproductive traits in Egyptian buffalo. **Animal Reproduction Science** 231. <https://doi.org/10.1016/j.anireprosci.2021.106800>

Araujo Neto FR, Pegolo NT, Aspilcueta-Borquis RR, Pessoa MC, Bonifácio A, Lobo R B, Oliveira HN (2018) Study of the effect of genotype–environment interaction on age at first calving and production traits in Nellore cattle using multi-trait reaction norms and Bayesian inference. **Animal Science Journal** 89:939–945. <https://doi.org/10.1111/asj.12994>

Araujo Neto FR, Takada L, Santos DJA, Aspilcueta-Borquis RR, Cardoso DF, Nascimento AV, Leão KM, Oliveira HN, Tonhati H (2020) Identification of genomic regions related to age at first calving and first calving interval in water buffalo using single-step GBLUP. **Reproduction in Domestic Animals** 55:1565–1572. <https://doi.org/10.1111/rda.13811>

Araujo Neto FR, Santos JCG, Silva Arce CD, Borquis RRA, Santos DJA, Guimarães KC, Nascimento AV, Oliveira HN, Tonhati H (2022) Genomic study of the resilience of buffalo cows to a negative energy balance. **Journal of Applied Genetics** 63:379–388. <https://doi.org/10.1007/s13353-021-00680-x>

Aspilcueta-Borquis RR, Seno LO, Araujo Neto FR, Santos DJA, Hurtado-Lugo NA, Tonhati H (2022) Lifetime productivity: Genetic study of longevity and its associations with economically important traits in dairy buffaloes. **Livestock Science** 259. <https://doi.org/10.1016/j.livsci.2022.104900>

Barros CC, Aspilcueta-Borquis RR, Fraga AB, Tonhati H (2016) Genetic parameter estimates for production and reproduction traits in dairy buffaloes. **Revista Caatinga** 29: 216–221. <https://doi.org/10.1590/1983-21252016v29n125rc>

Calus MPL (2010) Genomic breeding value prediction: Methods and procedures. **Animal** 4:157–164. <https://doi.org/10.1017/S1751731109991352>

Camargo GMF, Aspilcueta-Borquis RR, Fortes MRS, Porto-Neto R, Cardoso DF, Santos DJA, Lehnert SA, Reverter A, Moore SS, Tonhati H (2015) Prospecting major genes in dairy buffaloes. **BMC Genomics** 16. <https://doi.org/10.1186/s12864-015-1986-2>

Carvalho R, Costilla R, Neves HHR, Albuquerque LG, Moore S, Hayes BJ (2019) Unraveling genetic sensitivity of beef cattle to environmental variation under tropical conditions. **Genetics Selection Evolution** 51. <https://doi.org/10.1186/s12711-019-0470-x>

Carvalho Filho I, Silva DA, Teixeira CS, Silva TL, Mota LFM, Albuquerque LG, Carneiro R (2022) Heteroscedastic Reaction Norm Models Improve the Assessment of Genotype by Environment Interaction for Growth, Reproductive, and Visual Score Traits in Nellore Cattle. **Animals** 12. <https://doi.org/10.3390/ani12192613>

Cesarani A, Biffani S, Garcia A, Lourenco D, Bertolini G, Neglia G, Misztal I, Macciotta NPP (2021) Genomic investigation of milk production in Italian buffalo. **Italian Journal of Animal Science** 20:539–547. <https://doi.org/10.1080/1828051X.2021.1902404>

Chiaia HLJ, Lemos MVA, Venturini GC, Aboujaoude C, Berton MP, Feitosa FB, Carneiro R, Albuquerque LG, Oliveira HN, Baldi F (2015) Genotype × environment interaction for age at first calving, scrotal circumference, and yearling weight in Nellore cattle using reaction norms in multitrait random regression models. **Journal of Animal Science** 93:1503–1510. <https://doi.org/10.2527/jas2014-8217>

Christensen OF, Lund MS (2010) Genomic prediction when some animals are not genotyped. **Genetics Selection Evolution** 42. <https://doi.org/10.1186/1297-9686-42-2>

D'Occhio MJ, Ghuman SS, Neglia G, Valle GD, Baruselli PS, Zicarelli L, Visintin JA, Sarkar M, Campanile G (2020) Exogenous and endogenous factors in seasonality of reproduction in buffalo: A review. **Theriogenology** 150:186-192. <https://doi.org/10.1016/j.theriogenology.2020.01.044>

Falconer DS (1952) The problem of environment and selection. **The American Naturalist** 86:293–298.

Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics. **Pearson**, Burnt Mill, England p. 480.

Fernandes PB, Marques KO, Araujo Neto FR, Oliveira DP, Hurtado-Lugo NA, Aspilcueta-Borquis RR, Tonhati H (2016) Genetic-Quantitative Study of the First-Service Pregnancy Probability of Murrah Heifers. **Reproduction in Domestic Animals** 51:428–434. <https://doi.org/10.1111/rda.12697>

Freitas GR, Hurtado-Lugo NA, Santos DJA, Aspilcueta Borquis RR, Pegolo NT, Tonhati H, Araújo Neto FR (2019) Genotype–environment interaction for age at first calving in buffaloes, using the reaction norm model. **Reproduction in Domestic Animals** 54:727–732. <https://doi.org/10.1111/rda.13414>

Gao H, Christensen OF, Madsen P, Nielsen US, Zhang Y, Lund MS, Su G (2012) Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. **Genetics Selection Evolution** 44. <https://doi.org/10.1186/1297-9686-44-8>

Gautason E, Sahana G, Su G, Benjamínsson BH, Jóhannesson G, Guldbbrandtsen B (2021) Short communication: investigation of the feasibility of genomic selection in Icelandic Cattle. **Journal of Animal Science** 99. <https://doi.org/10.1093/jas/skab139>

Iamartino D, et al. (2017) Design and validation of a 90K SNP genotyping assay for the water buffalo (*Bubalus bubalis*). **Plos One** 12. <https://doi.org/10.1371/journal.pone.0185220>

Kumar V, Chakravarty AK, Patil CS, Valsalan J, Mahajan A (2015) Estimate of genetic and non-genetic parameters for age at first calving in Murrah buffalo. **Indian Journal of Animal Sciences** 85:84–85. <https://www.researchgate.net/publication/279332133>

Legarra A, Reverter A (2018) Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. **Genetics Selection Evolution** 50. <https://doi.org/10.1186/s12711-018-0426-6>

Li J et al. (2018) Integrating RNA-seq and GWAS reveals novel genetic mutations for buffalo reproductive traits. **Animal Reproduction Science** 197:290-295. <https://doi.org/10.1016/j.anireprosci.2018.08.041>

Liu et al. (2018) Genome-wide association studies to identify quantitative trait loci affecting milk production traits in water buffalo. **Journal of Dairy Science** 101:433–444. <https://doi.org/10.3168/jds.2017-13246>

Lourenço D, Legarra A, Tsuruta S, Masuda Y, Aguilar I, Misztal I (2020) Single-Step Genomic Evaluations from Theory to Practice: Using SNP Chips and Sequence Data in BLUPF90. **Genes** 11. <https://doi.org/10.3390/genes11070790>

Lopes FB, Baldi F, Passafaro TL, Brunes LC, Costa MFO, Eifert EC, Narciso MG, Rosa GJM, Lobo RB, Magnabosco CU (2021) Genome-enabled prediction of meat and carcass traits using Bayesian regression, single-step genomic best linear unbiased prediction and blending methods in Nelore cattle. **Animal** 15. <https://doi.org/10.1016/j.animal.2020.100006>

Low WY et al. (2019) Chromosome-level assembly of the water buffalo genome surpasses human and goat genomes in sequence contiguity. **Nature Communications** 10. <https://doi.org/10.1038/s41467-018-08260-0>

Meuwissen T, Hayes B, Goddard M (2016) Genomic selection: A paradigm shift in animal breeding. **Animal Frontiers** 6:6–14. <https://doi.org/10.2527/af.2016-0002>

Mota LFM, Lopes FB, Fernandes Júnior GA, Rosa GJM, Magalhães AFB, Carneiro, R, Albuquerque LG (2020a) Genome-wide scan highlights the role of candidate genes on phenotypic plasticity for age at first calving in Nelore heifers. **Scientific Reports** 10. <https://doi.org/10.1038/s41598-020-63516-4>

Mota LFM, Fernandes GA, Herrera AC, Scalez DCB, Espigolan R, Magalhães AFB, Carneiro R, Baldi F, Albuquerque LG (2020b) Genomic reaction norm models exploiting genotype × environment interaction on sexual precocity indicator traits in Nelore cattle. **Animal Genetics** 51:210–223. <https://doi.org/10.1111/age.12902>

Mulder HA, Bijma P (2005) Effects of genotype X environment interaction on genetic gain in breeding programs. **Journal of Animal Science** 83:49–61. <https://doi.org/10.2527/2005.83149x>

Oliveira DP, Lourenco DAL et al. (2018) Reaction norm for yearling weight in beef cattle using single-step genomic evaluation. **Journal of Animal Science** 96:27–34. <https://doi.org/10.1093/jas/skx006>

Perera BMAO (2011) Reproductive cycles of buffalo. **Animal Reproduction Science** 124:194–199. <https://doi.org/10.1016/j.anireprosci.2010.08.022>

Pessoa GA, Navarro RB, Rauber LP, Brass KE, Silva CAM, Rubin MIB (2010) Influence of climatic conditions on in vitro production of bovine embryos. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia** 62:1381–1387. <https://doi.org/10.1590/S0102-09352010000600013>

Robertson A (1959) The sampling variance of genetic correlation coefficient. **Biometrics** 15:469–485. <https://doi.org/10.2307/2527750>

Roos APW, Schrooten C, Veerkamp RF, van Arendonk JAM (2011) Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls. **Journal of Dairy Science** 94:1559–1567. <https://doi.org/10.3168/jds.2010-3354>

Salem MMI, Nasr MAF, Amin AMS (2021) Principal component analysis of breeding values for birth weight milk and reproductive traits of the Egyptian buffalo. **Tropical Animal Health and Production** 53. <https://doi.org/10.1007/s11250-021-02625-2>

Santana Jr ML, Bignardi AB, Stefani G, el Faro L (2017) Genetic component of sensitivity to heat stress for nonreturn rate of Brazilian Holstein cattle. **Theriogenology** 98:101–107. <https://doi.org/10.1016/j.theriogenology.2017.04.052>

Santos JC, Malhado CHM, Carneiro PLS, Rezende MPG, Cobuci JA (2020) Genotype-environment interaction for age at first calving in Holstein cows in Brazil. **Veterinary and Animal Science** 9. <https://doi.org/10.1016/j.vas.2020.100098>

Scott BA, Haile-Mariam M, Cocks BG, Pryce JE (2021) How genomic selection has increased rates of genetic gain and inbreeding in the Australian national herd, genomic information nucleus, and bulls. **Journal of Dairy Science** 104:11832–11849. <https://doi.org/10.3168/jds.2021-20326>

Seno LO, Tonhati H, Cardoso VL, el Faro L, Sesana RC, Aspilcueta Borquis RR (2007) Genetic parameters for milk yield, age at first calving and interval between first and second calving in milk buffaloes. **Italian Journal of Animal Science** 6:397–400. <https://doi.org/10.4081/ijas.2007.s2.397>

Shao B et al. (2021) Genetic Features of Reproductive Traits in Bovine and Buffalo: Lessons From Bovine to Buffalo. **Frontiers in Genetics** 12. <https://doi.org/10.3389/fgene.2021.617128>

Shi R et al. (2021) Genotype-by-environment interaction in Holstein heifer fertility traits using single-step genomic reaction norm models. **BMC Genomics** 22. <https://doi.org/10.1186/s12864-021-07496-3>

Silva JCB et al (2021) Reproductive seasonality influences oocyte retrieval and embryonic competence but not uterine receptivity in buffaloes. **Theriogenology** 170:77–84. <https://doi.org/10.1016/j.theriogenology.2021.04.019>

Silveira Marinho KF, Andrade LMB, Spyrides MHC, Silva CMS, Oliveira CP, Bezerra BG, Mutti PR (2020) Climate profiles in Brazilian microregions. **Atmosphere** 11. <https://doi.org/10.3390/atmos11111217>

Sousa DR, Nascimento AV, Lôbo RNB (2021) Prediction of genomic breeding values of milk traits in Brazilian Saanen goats. **Journal of Animal Breeding and Genetics** 138:541–551. <https://doi.org/10.1111/jbg.12550>

Streit M, Reinhardt F, Thaller G, Bennewitz J (2012) Reaction norms and genotype-by-environment interaction in the German Holstein dairy cattle. **Journal of Animal Breeding and Genetics** 129:380–389. <https://doi.org/10.1111/j.1439-0388.2012.00999.x>

Tamboli P, Bharadwaj A, Chaurasiya A, Bangar YC, Jerome A (2021) Genetic Parameters for First Lactation and Lifetime Traits of Nili-Ravi Buffaloes. **Frontiers in Veterinary Science** 8. <https://doi.org/10.3389/fvets.2021.557468>

Vale W, Ribeiro H (2005) Características reprodutivas dos bubalinos: puberdade, ciclo estral, involução uterina e atividade ovariana no pós-parto. **Revista Brasileira de Reprodução Animal** 29:63-73. <http://cbra.org.br/pages/publicacoes/rbra/download/RE026.pdf>

VanRaden PM (2008) Efficient methods to compute genomic predictions. **Journal of Dairy Science** 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>

Vilela RNS, Sena TM, Aspilcueta-Borquis RR, Seno, LO, Araujo Neto FR, Scaletz DCB, Tonhati H (2020) Genetic correlations and trends for traits of economic importance in dairy buffalo. **Animal Production Science** 60:492–496. <https://doi.org/10.1071/AN19051>

Wang H, Misztal I, Aguilar I, Legarra A, Muir WM (2012) Genome-wide association mapping including phenotypes from relatives without genotypes. **Genetics Research** 94:73–83. <https://doi.org/10.1017/S0016672312000274>

Warriach HM, McGill DM, Bush RD, Wynn PC, Chohan KR (2015) A review of recent developments in buffalo reproduction - A review. Asian-Australasian **Journal of Animal Sciences** 28:451–455. <https://doi.org/10.5713/ajas.14.025>

CAPÍTULO 2 – Análise genômica em modelos de normas de reação para idade ao primeiro parto em búfalas Murrah

RESUMO - A idade ao primeiro parto (IPP) é uma medida de precocidade sexual associada ao início da vida produtiva das fêmeas de um rebanho leiteiro. Além disso, a menor IPP diminui o intervalo de geração e os abates involuntários das fêmeas. No entanto, a IPP tem estimativa de herdabilidade baixa, e é muito influenciada por fatores ambientais. Uma das formas de melhorar o desempenho reprodutivo das fêmeas, é selecionando animais robustos por meio da estimação dos valores genéticos (VG) utilizando modelos que incluam a interação genótipo-ambiente (IGA) com aplicação da norma de reação via regressão aleatória (RNM), e pela compreensão da base molecular subjacente à IGA da IPP. Assim, o objetivo deste estudo foi estimar os VG considerando a IGA via modelo de RNM e identificar genes candidatos a IGA da IPP de búfalas leiteiras por meio da associação genômica ampla (GWAS). Foram utilizados 1.795 dados de IPP de três rebanhos de búfalos da raça Murrah distribuídos no sudeste e nordeste do Brasil, um pedigree considerando 4 gerações que totalizaram 3.363 animais e informações genótípicas de 553 animais. Os gradientes ambientais foram formados a partir das soluções dos grupos de contemporâneos obtidos a partir da estimação dos VG da produção de leite (PL270). As estimativas de herdabilidade variaram de 0,15 a 0,39 ao longo do gradiente ambiental. O GWAS do parâmetro de inclinação do RNM identificou regiões genômicas importantes. A janela genômica que explicou maior proporção da variância genética da inclinação estava localizada no BBU1 e foi de 0,67%. Após análise funcional foram detectados cinco genes candidatos principais, envolvidos em dois processos biológicos. Os resultados sugerem que há IGA para a IPP em búfalos da raça Murrah, com reclassificação dos VG dos animais quando diferentes condições ambientais são consideradas. A inclusão de informações genômicas aumentou a acurácia dos valores genéticos do intercepto e da inclinação da norma de reação, indicando melhor ajuste dos dados. A análise de GWAS permitiu a identificação de genes importantes envolvidos nos efeitos da inclinação da norma de reação da IPP.

Palavras-chave: GWAS, herdabilidade, plasticidade, regressão aleatória, reprodução

CHAPTER 2 - Genomic analysis in reaction norm models for age at first calving in Murrah buffaloes

ABSTRACT - Age at first calving (AFC) is a measure of sexual precocity associated with the early productive life of females in a dairy herd. In addition, lower IPP decreases the generation interval and involuntary slaughters of females. However, AFC has low heritability estimate, and is highly influenced by environmental factors. One way to improve reproductive performance of females, is by selecting robust animals by estimating genetic values (EBV) using models that include the genotype-environment interaction (GEI), with application of the reaction norm via random regression (RNM), and by understanding the molecular basis underlying the GEI of AFC. Thus, the aim of this study was to estimate the EBVs considering GEI via the RNM model and to identify candidate genes for GEI of AFC of dairy buffaloes by broad genome-wide association (GWAS). We used 1,795 AFC data from three Murrah buffalo herds distributed in southeastern and northeastern Brazil, a pedigree considering 4 generations that totaled 3,363 animals and genotypic information of 553 animals. The environmental gradients were formed from the solutions of the groups of contemporaries obtained from the estimation of the EBV of milk production (PL270). Heritability estimates ranged from 0.15 to 0.39 along the environmental gradient. GWAS of the RNM slope parameter identified important genomic regions. The genomic window that explained the largest proportion of the genetic variance of the slope was located in BBU1 and was 0.67%. After functional analysis, five major candidate genes involved in two biological processes were detected. The results suggest that there is GEI for AFC in Murrah breed buffaloes, with reclassification of the animals' EBVs when different environmental conditions are considered. The inclusion of genomic information increased the accuracy of the genetic values of the intercept and slope of the reaction norm, indicating better fit of the data. GWAS analysis allowed the identification of important genes involved in the effects of the AFC reaction norm slope.

Keywords: GWAS, heritability, plasticity, random regression, breeding

2.1 Introdução

A idade ao primeiro parto (IPP), definida como número de dias ou meses da data de nascimento da novilha até a data do primeiro parto, é uma característica importante em rebanhos leiteiros, pois é uma medida de precocidade sexual e afeta diretamente a rentabilidade do sistema de produção, visto que, está relacionada com o início da vida reprodutiva e produtiva das fêmeas leiteiras e influencia diretamente os custos de substituição e, portanto, na vida útil das fêmeas (Santos et al., 2020). Dessa forma, em programas de melhoramento genético, há a preferência pela seleção de animais com menor IPP, pois, a menor IPP está associada com a diminuição do intervalo de geração, e consequentemente, ao maior progresso genético (Atashi et al., 2021).

Aspilcueta-Borquis et al. (2022) identificaram que a expectativa de vida produtiva das búfalas até 48 meses, definida como o número total de dias no leite registrado até 48 meses de idade, teve correlação de -0,11 com a IPP. Apesar da baixa correlação, houve evidência de que búfalas tem maior expectativa de permanecer no rebanho, quando tem a capacidade de produzir um bezerro precocemente. Em bovinos leiteiros, foi verificado que a longevidade das fêmeas no rebanho, também está correlacionada com a IPP, visto que um dos critérios utilizados para a permanência da vaca no rebanho é se a fêmea produziu ou não um bezerro quando exposta ao touro ou quando realizada a inseminação artificial. Consequentemente, a menor IPP reduz os abates involuntários das fêmeas (Sherwin et al., 2016).

No entanto, maior parte da variância fenotípica da IPP é oriunda da contribuição dos efeitos ambientais (nutrição, fatores climáticos, fatores sanitários, manejos reprodutivos, entre outros) (Santos et al., 2020), o que ocasiona estimativas de herdabilidade baixa. No Brasil foram descritas estimativas de herdabilidade variando de 0,16 a 0,19 em rebanhos de búfalos, portanto a IPP é uma característica de ganho genético lento (Araujo Neto et al., 2020; Camargo et al., 2015; Freitas et al., 2019). Grande parte dessa baixa variância genética aditiva da característica é devido a seleção intensa para características produtivas, enquanto há menor interesse sobre as características reprodutivas tanto em bovinos quanto em búfalos (Berry et al., 2014; Cesarani et al., 2020; Jamuna e Chakravarty et al., 2016). Maiores estimativas de

herdabilidade tem sido relatadas para a produção de leite de búfalas, as quais variaram de 0,22, a 0,28, e, portanto, é a mais utilizada como critério de seleção, devido a relação entre produção de leite e o rendimento de queijo *Mozzarella*, que tem sido o principal objetivo de seleção (Agudelo-Gómez et al., 2015; Aspilcueta-Borquis et al., 2010; Malhado et al., 2013).

Uma das formas de melhorar o desempenho reprodutivo das fêmeas, é por meio da seleção de animais robustos, que respondem bem a qualquer tipo de ambiente. Modelos que incluem a interação genótipo-ambiente (IGA) podem avaliar o efeito de diferentes circunstâncias ambientais sobre o desempenho dos animais, e um método utilizado amplamente para determinar a sensibilidade dos animais de produção em diferentes ambientes é a norma de reação (RNM) via regressão aleatória (Mota et al., 2020).

Com o advento da genômica, maiores acurácias nas estimativas dos valores genéticos (VG) tem sido observada com a aplicação das normas de reação genômicas e tem permitido a identificação de regiões genômicas associadas aos efeitos da IGA (Carvalho et al., 2019; Oliveira et al., 2018).

A raça Murrah tem papel importante na produção leiteira, bem como socioeconômico no mundo todo. Dessa forma, o objetivo deste estudo foi avaliar a IGA para a IPP por meio do modelo de norma de reação linear, avaliar a habilidade de predição com a inclusão de dados genotípicos e identificar regiões e genes candidatos a IGA da IPP de búfalas leiteiras da raça Murrah.

2.2 Material e métodos

2.2.1 Banco de dados e consistências

Os dados fenotípicos foram provenientes de 2.314 registros de idade ao primeiro parto (IPP) de búfalas da raça Murrah nascidas entre 1995 e 2017. Os dados representavam três rebanhos localizados nos estados do Rio Grande do Norte (RN), Ceará (CE) e São Paulo (SP) – Brasil, que estavam conectados devido ao compartilhamento de material genético entre fazendas, o que foi confirmado por meio do teste de conectabilidade do pacote AMC (Roso e Schenkel, 2006).

Para este estudo, a aprovação do Comitê de Cuidados e Uso de Animais não foi necessária, uma vez que os dados foram obtidos a partir de uma base de dados previamente disponível.

A IPP foi calculada como a diferença em meses entre a data do primeiro registro de parto e a data de nascimento das novilhas. Foram mantidos apenas registros de IPP com mais ou menos três desvios padrão da média dos grupos de contemporâneos (GCs) e GCs que continham no mínimo sete animais e foram formados pela concatenação de animais nascidos na mesma fazenda, com mesmo ano e estação de nascimento.

A produção de leite acumulada aos 270 dias (PL270) é o principal critério de seleção adotado nos três rebanhos, dessa forma, é uma característica que expressa o nível tecnológico das fazendas e a torna adequada para a avaliação da sensibilidade à variação ambiental. Portanto, as soluções dos GCs da avaliação genética da PL270 foram utilizadas como gradiente ambiental na RNM para avaliar a IGA da IPP. A PL270 foi obtida a partir das produções de leite mensuradas nos dias dos controles leiteiros, foram consideradas as produções de leite a partir do quinto dia do início da lactação de búfalas com pelo menos quatro controles e duração de lactação mínima de 120 dias. Os GCs foram formados a partir da concatenação de animais nascidos na mesma fazenda e com mesmo ano e estação de parto, e permaneceram apenas GCs que continham no mínimo sete animais.

As estações de nascimento e de parto utilizadas para a formação dos GCs para a IPP e PL270 foram definidas de acordo com a região onde estão localizadas as fazendas e foram agrupadas em duas classes, estação seca (classe 1) e estação chuvosa (classe 2). Para o estado do RN, os meses de outubro a março foram considerados da classe 1 e os meses de abril a setembro pertencentes a classe 2. Para o estado do CE os meses de junho a janeiro foram pertencentes a classe 1 e os meses de fevereiro a maio pertencentes a classe 2. E, para o estado de SP os meses de abril a setembro foram pertencentes a classe 1 e outubro a março a classe 2.

O pedigree da IPP e da PL270 foram referentes a quatro gerações. A quantidade de informações fenotípicas e de pedigree, bem com as estatísticas descritivas do banco de dados após as consistências foram descritas na Tabela 1.

Na Tabela 2, foram descritas a estrutura genealógica e genotípica do banco de dados.

Tabela 1. Estatística descritiva da idade ao primeiro parto (IPP) e da produção de leite acumulada aos 270 dias (PL270)

Características	Número de informações	Média	DP	Número de GCs	Pedigree
PL270 (kg)	2.026	1.742,73	574,43	82	3.103
IPP (meses)	1.795	37,01	5,00	76	3.363

GCs: Grupos de contemporâneos

Tabela 2. Estrutura genealógica e estrutura dos marcadores utilizados para a realização das análises de normas de reação para idade ao primeiro parto em búfalas.

Parâmetros	Número de observações
Número de animais no pedigree	3.363
Número de pais	224
Número de mães	1.751
Número de animais genotipados	553
Número de animais com fenótipo e genótipo	439
Número de marcadores SNP	44.574

Um total de 553 animais (539 fêmeas e 14 machos) genotipados com o Axiom Buffalo Genotyping Array 90 K (Iamartino et al., 2017) estavam disponíveis para a análise de associação genômica ampla em passo único (ssGWAS). Foram utilizados somente SNPs presentes em cromossomos autossômicos e permaneceram apenas marcadores com call rate maior que 95%, MAF maior que 3% e p-valor para o teste de equilíbrio Hardy Weinberg menor que 10^{-5} . Todas as amostras apresentaram call rate maior que 90% e, portanto, todas permaneceram para a análise. Além disso, marcadores sem posição conhecida foram removidos. Após o controle da qualidade, 553 amostras e 44.574 marcadores permaneceram para as análises posteriores.

2.2.2 Estimação dos componentes de variância da norma de reação (RNM)

2.2.2.1 Gradiente ambiental (GA)

Previamente a análise de RNM para a IPP, foram obtidas as soluções dos GCs da PL270 por meio do modelo animal tradicional (BLUP) utilizando as informações de pedigree (matriz A). As soluções foram obtidas por meio do modelo animal. Em notação matricial o modelo utilizado foi:

$$y = X\beta + Z\alpha + \varepsilon,$$

em que, y é o vetor dos registros fenotípicos da PL270. β é o vetor dos efeitos fixos (CG e idade da vaca linear e quadrática como covariáveis), α é o vetor dos efeitos aditivos e ε o vetor dos resíduos; X e Z , representam nesta ordem, matrizes de incidência que associam β e α ao vetor de observações. As seguintes pressuposições foram assumidas:

$$E(y) = X\beta E(\alpha) = 0 \quad E(\varepsilon) = 0$$

$$Var(\alpha) = A\sigma_{\alpha}^2 \quad Var(\varepsilon) = I_n\sigma_{\varepsilon}^2$$

em que σ_{α}^2 e σ_{ε}^2 representam as variâncias genética aditiva e residual; A e I_n são as matrizes de parentesco aditiva e identidade, respectivamente. Posteriormente as soluções dos GCs da PL270 foram padronizadas para ter uma média 0 e variância 1.

2.2.2.2 Norma de reação tradicional (BLUP) e genômica (ssGBLUP)

Após definição do GA, a análise para obter os parâmetros genéticos da IPP em todos os GA foi realizada por meio do modelo de norma de reação via regressão aleatória linear. Em notação matricial, o modelo utilizado foi:

$$y = X\beta + \Phi\alpha + \varepsilon$$

em que y , é o vetor de observações da IPP, β é o vetor de efeitos fixos (GC), α é o vetor dos efeitos genéticos aditivos dos coeficientes de regressão aleatória e ε é o vetor de efeitos residuais aleatórios; X é a matriz de incidência que relaciona as

observações aos efeitos fixos e; Φ é a matriz de incidência com coeficientes de regressão relacionados aos efeitos genéticos aditivos. As estimativas de componentes de (co)variância da RNM foram obtidas pela probabilidade máxima restrita utilizando os programas REMLF90 e AIREMLF90 (Miszta et al., 2014).

2.2.3 Habilidade de predição dos métodos genômico e tradicional

Para avaliar a habilidade de predição entre as estimativas dos valores genéticos e genômicos das normas de reação (BLUP e ssGBLUP), o método LR (linear regression) proposto por Legarra e Reverter (2018) foi utilizado. Esta metodologia baseia-se na aplicação de estatísticas que comparam os resultados obtidos por meio da utilização de um banco de dados completo (todas as informações fenotípicas disponíveis) e, um banco de dados parcial formado pela eliminação dos fenótipos de determinados animais (chamados de indivíduos-foco). Utilizando os valores genéticos/genômicos dos indivíduos-foco estimados em análises com o banco completos ($\hat{u}_w$) e parcial ($\hat{u}_p$), foram então calculadas as estatísticas relacionadas com a habilidade de predição.

Para isso, o método de amostragem consistiu em eliminar um indivíduo genotipado (animais-foco) do banco de dados fenotípico e realizar a análise com N-1 animais. Este processo foi realizado 439 vezes (número de animais com fenótipo e genótipo), sendo obtido ao final um arquivo contendo $\hat{u}_w$ e $\hat{u}_p$, utilizado para o cálculo das seguintes estatísticas:

Acurácia	Dispersão	Correlação
$r_{A\hat{A}} = \sqrt{\frac{cov(\hat{u}_w, \hat{u}_p)}{(1 - \bar{F})\sigma_g^2}}$	$b_{w,p} = \frac{cov(\hat{u}_w, \hat{u}_p)}{var(\hat{u}_p)}$	$\rho_{p,w} = \frac{cov(\hat{u}_p, \hat{u}_w)}{\sqrt{var(\hat{u}_w)var(\hat{u}_p)}}$

onde $\bar{F}$ é o coeficiente de endogamia médio dos indivíduos-foco, e σ_g^2 é a variância genética aditiva. Entretanto, é importante ressaltar que a estatística de acurácia utilizada neste estudo, consiste em uma modificação da forma proposta por Legarra e Reverter (2018), uma vez que: a) considera a variância genética estimada, pois as

fazendas utilizam como critério de seleção a produção e qualidade do leite e; b) como indicado no próprio artigo dos autores, o denominador da fórmula pode ser reduzido ao assumir a coancestria com valor próximo a da endogamia.

2.2.4 Estudo de associação genômica ampla (GWAS) da interação genótipo-ambiente (IGA)

O estudo de associação genômica (GWAS) foi realizado para o parâmetro da inclinação (b_1) da RNM utilizando a proporção da variância explicada por janelas genômicas com 1 Mb. Foram consideradas apenas janelas genômicas que explicaram pelo menos 0,3% da variância genética aditiva e obtidos os genes candidatos posicionais inseridos dentro dessas janelas considerando a montagem do genoma bubalino UOA_WB_1 (Low et al., 2019) disponível no NCBI (<https://www.ncbi.nlm.nih.gov/genome/gdv>). Posteriormente, dispendo da lista de genes, foi realizada a análise funcional, empregando o pacote ClueGo (Bindea et al., 2009) disponível para o programa Cytoscape (Shannon et al., 2003). Para esta finalidade foram considerados a base de ontologia de genes (GO) para processos de sistema imune e biológicos; kappa score superior a 0,5; especificidade de network global; sem mínimo de tamanho de cluster; foram considerados significativos valores de p corrigidos pelo teste de *Bonferroni step down* abaixo do nível de significância de 0,05.

2.3 Resultados

2.3.1 Componentes de variância e parâmetros genéticos da norma de reação (RNM)

As variâncias (erro padrão) da inclinação, do intercepto e residual foram de 3,241 (0,739), 0,377 (0,506) e 8,857(0,609), respectivamente. A proporção entre as variâncias da inclinação (b_1) e do intercepto (b_0) foi de 0,12, e a correlação entre o intercepto e a inclinação foi positiva e de alta magnitude (0,74).

As amplitudes das soluções dos CG da PL270 obtidos utilizando o modelo

animal padrão, variaram de $-1,5$ a $1,2$, e representaram ambientes de baixo a alto nível de produção, respectivamente (Figura 1A). As estimativas de herdabilidade variaram de $0,15$ a $0,39$ do menor ao maior nível de produtividade, indicando que há aumento das estimativas de herdabilidade para a IPP quando as condições ambientais são melhores (Figura 1A). Além disso, esses resultados indicam que há IGA para a IPP de búfalas leiteiras da raça Murrah.

As estimativas de correlação genética da IPP entre os diferentes gradientes ambientais (GA) variaram de $0,75$ a $0,99$ (Figura 1B), com estimativas mais baixas entre ambientes extremos, indicando IGA.

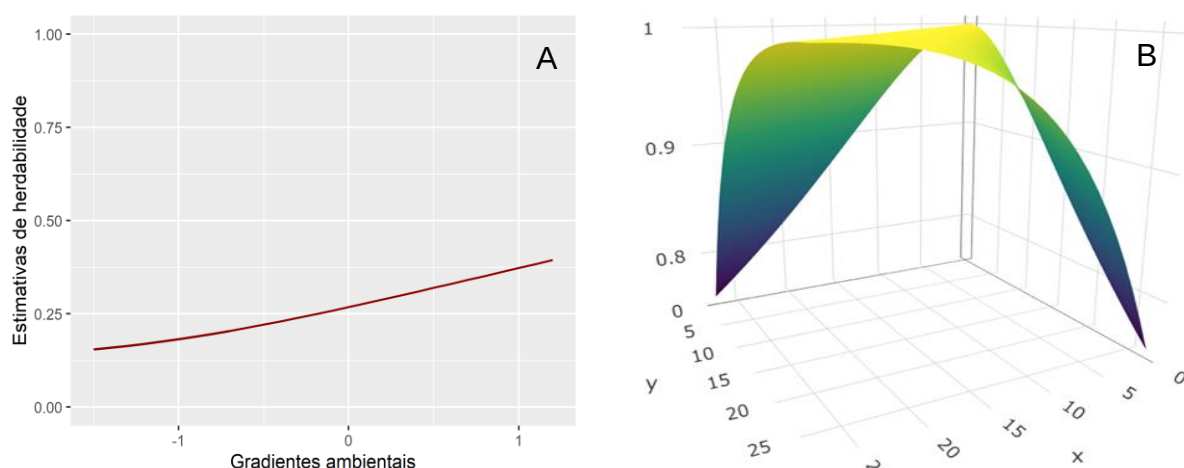


Figura 1. Estimativas de herdabilidade ao longo do gradiente ambiental (A) e estimativas de correlação genética entre os gradientes ambientais (B) obtidas via modelo norma de reação linear para idade ao primeiro parto em búfalas da raça Murrah

As normas de reação dos touros que tinham no mínimo 20 filhas, apresentaram um padrão semelhante nos valores genéticos (VG) estimados via PBLUP (EBV) e via ssGBLUP (GEBV) com poucas diferenças na reclassificação dos animais de um modelo para o outro (Figura 2).

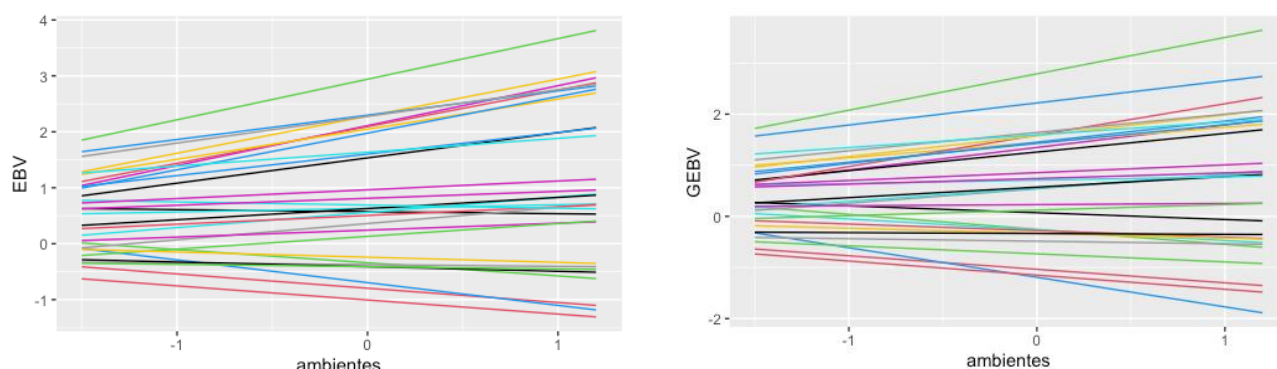


Figura 2. Normas de reação genéticas (EBV) e genômicas (GEBV) obtidos pelo modelo de norma de reação linear para idade ao primeiro parto, considerando touros com no mínimo 20 filhas com registros

2.3.2 Predição genômica das normas de reação (RNM)

O RNM ssGBLUP apresentou incrementos em acurácia (r_{AA}) na proporção de 5,21% e 4,63% para os valores genéticos do intercepto e da inclinação, respectivamente (Tabela 3).

Com relação a estatística de dispersão ($\hat{D}$), o método PBLUP superestimou os valores genéticos do intercepto e da inclinação, enquanto o método ssGBLUP foi caracterizado pela subestimação dos valores genômicos do intercepto e da inclinação. Tendo em vista que o valor ideal deste parâmetro é a unidade, o método ssGBLUP reduziu a dispersão dos valores genéticos e proporcionou melhor ajuste para o intercepto e para a inclinação do RNM. Maior incremento fenotípico foi observado para o intercepto (0,67%) em relação a inclinação (0,26%) quando há a inclusão de informações genotípicas (ssGBLUP).

Tabela 3. Acurácia (r_{AA}), dispersão ($\hat{D}$) e incremento fenotípico (Inc_{PHE}) dos valores genéticos (PBLUP) e genômicos (ssGBLUP) obtidos pelo modelo de norma de reação da idade ao primeiro parto, considerando os dados completos e parciais

Parâmetro	PBLUP			ssGBLUP		
	r_{AA}	$\hat{D}$	Inc_{PHE}	r_{AA}	$\hat{D}$	Inc_{PHE}
b_0	0,422	1,038	0,742	0,444	0,974	0,747
b_1	0,367	1,042	0,762	0,384	0,978	0,764

A acurácia dos valores genéticos e genômicos ao longo do gradiente ambiental aumentou quando houve a inclusão de marcadores SNP no modelo em relação ao método baseado em pedigree (PBLUP) (Figura 3).

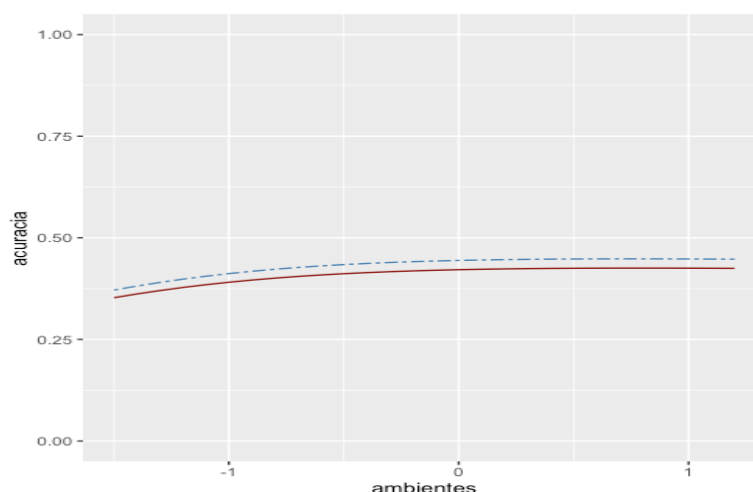


Figura 3. Acurácia dos valores genéticos (linha vermelha) e genômicos (azul) ao longo gradiente ambiental para idade ao primeiro parto em búfalas da raça Murrah.

2.3.3 Análise de associação genômica ampla (GWAS) para plasticidade (inclinação)

Sete regiões genômicas foram detectadas considerando a porcentagem da variância genética explicada pelas janelas genômicas do parâmetro de inclinação do RNM, com valor de limiar superior a 0,3% para a IPP. Essas regiões genômicas estavam distribuídas nos cromossomos 1, 8, 11, 16 e 18 (Figura 4) e totalizaram 2,92% da variância genética total. A janela genômica que explicou a maior porcentagem de variância (0,67% para a inclinação) foi a janela localizada no BBU1 (137.434.116 a 138.426.182 pb). Considerando o critério estatístico p -valor, não houve SNPs significativos.

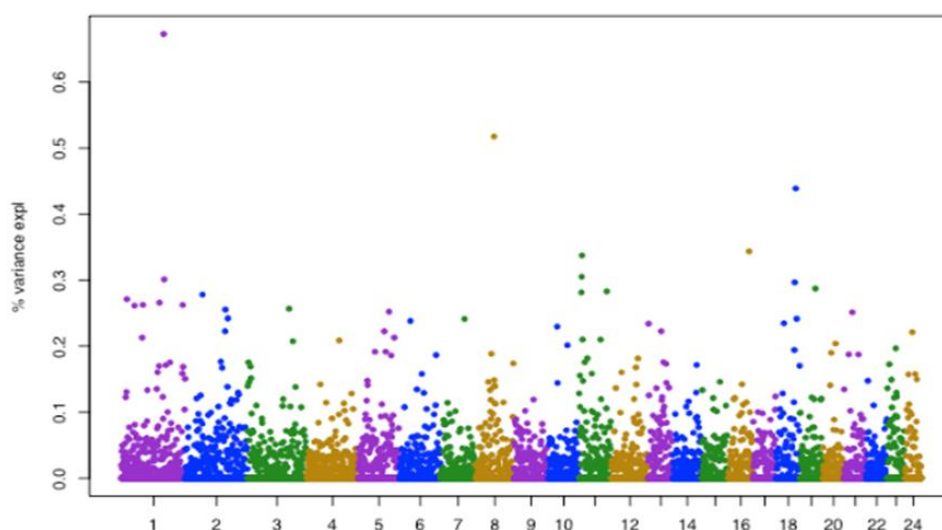


Figura 4. Manhattan-plot considerando a porcentagem da variância explicada por janelas exclusivas de 1 Mb para o parâmetro de inclinação (b_1) do modelo de norma de reação (RNM) para idade ao primeiro parto (IPP) em búfalas da raça Murrah.

O número de SNPs contidos nas janelas genômicas significativas variaram de 15 a 28 e totalizaram 88 genes candidatos (Tabela 4). Posteriormente, foram identificados apenas genes que participavam de processos biológicos importantes demonstrados pela análise funcional, dos quais, 5 genes foram detectados (KCNN4, NECTIN2, APOE, APOC2 e APOC4). Esses genes estavam divididos em duas rotas metabólicas: regulação das células T e partículas lipoprotéicas de baixa densidade.

Tabela 4. Número de SNPs (N_{snp}), cromossomo, posição inicial (PI), posição final (PF), porcentagem da variância genética explicada por marcadores moleculares da inclinação da norma de reação RNM (%Variância) e genes encontrados em janelas de 1Mb para IPP em búfalas Murrah

N_{snp}	Cromossomo	PI	PF	%Variância	Genes (UOA_WB_1)
25	1	137434116	138426182	0,67288	NAALADL2; LOC102416103; LOC112584561
17	8	58728756	59717614	0,51750	THAP5; GPR141; ELMO1; LOC102391616; LOC102390377; LOC112586493; DNAJB9; LOC112586492
28	18	51798888	52796278	0,43879	LOC102410032; SMG9; LOC102392841; ZNF285; NECTIN2; LOC102392564; PPP1R37; IRGQ; LOC112580167; ZNF45; LOC112580123; ZNF180; LOC112580321; LOC112580323; GEMIN7; LOC112580318; PLAUR; LYPD5; KCNN4; LOC102391408; LOC112580319; LOC112580320; PVR; LOC112580475; LOC112577731; CBLC; TRNAC-GCA; LOC102399059; CLASRP; APOE; BCAM; TOMM40; RELB; ZNF576; IRGC; ZNF428; CADM4; LOC102411349; LOC102414175; SRRM5; ZNF283; LOC102413321; ZNF404; ZNF235; CEACAM19; IGSF23; CEACAM20; ZNF112; CEACAM16; LOC112580322; BCL3; APOC4; ZNF296; APOC2; CLPTM1;
18	16	67875000	68813391	0,34340	ELMOD1; LOC112579386; LOC112579751; LOC102390409; GUCY1A2; CWF19L2; ALKBH8
24	11	9496548	10490769	0,33730	TRNAS-GGA; LOC112587770; LOC102403944; LOC112587769; TRNAK-CUU; TRNAK-UUU; STON2; LOC112587772; SEL1L
15	11	8399315	9288646	0,30511	LOC112588001
20	1	138477820	139476110	0,30109	NLGN1; SPATA16; LOC102389641; NCEH1; ECT2

2.4 Discussão

2.4.1 Componentes de variância obtido via norma de reação linear (RNM)

Correlações genéticas altas e positivas entre o intercepto (b_0) e a inclinação (b_1) indicam que os animais com maior valor VGE no intercepto também tendem a ter maior VGE na inclinação, e que seus desempenhos melhoram quando há menor restrição ambiental (Mota et al., 2020). Os resultados, indicaram que búfalos com melhores VGEs para a IPP respondem de forma mais eficiente quando são produzidos em condições ambientais melhores, e, portanto, os fenótipos são considerados plásticos.

Ambrosini et al. (2016) utilizando o RNM heterocedástico para a estimação da IGA para IPP em Nelore, obtiveram correlação de 0,90 entre o b_0 e b_1 , indicando que animais com maior VGE tem maior capacidade de expressar o potencial genético em ambientes menos restritivos. Resultados semelhantes foram observados por Freitas et al. (2019) para a IPP em búfalas, onde relataram correlação de 0,56 entre b_0 e b_1 . Correlações genéticas entre b_0 e b_1 inferiores (0,245) foram relatadas para IPP de novilhas Nelore (Araujo Neto et al., 2018).

O aumento das estimativas de herdabilidade em gradientes ambientais com maiores níveis de produção pode ser resultado das melhores condições ambientais fornecidas aos animais, que resultam em maiores estimativas da variância genética aditiva devido a maior capacidade de expressar o potencial genético em ambientes menos restritos (Santos et al., 2020; Lemos et al., 2015). As estimativas de herdabilidade foram semelhantes as relatadas na literatura (com uma variação de 0,12 a 0,48) para a IPP em búfalos utilizando modelos tradicionais sem a inclusão de IGA (Agudelo-Gómez et al., 2015; Amin et al., 2021; Araujo Neto et al., 2020; Camargo et al., 2015; Tamboli et al., 2022).

As estimativas de herdabilidade para a IPP foram semelhantes as descritas por Freitas et al. (2019) que utilizaram as médias da produção de leite (PL) dos GCs para a definição do GA, as quais variaram de 0,13 a 0,39 com maiores estimativas de herdabilidade em ambientes favoráveis. Santos et al. (2020) avaliando o efeito da IGA para a IPP em novilhas da raça Holandesa utilizando o RNM homocedástico,

obtiveram estimativas de herdabilidade que variaram de 0,04 a 0,28, com maiores estimativas de herdabilidade em ambientes desafiadores e menores estimativas nos melhores ambientes, entretanto, o critério utilizado para a definição do GA foi a solução dos GCs da IPP.

Chiaia et al. (2015) utilizando a média do ganho de peso pós-desmame dos GC como GA, obtiveram estimativas de herdabilidade variando de 0,09 a 0,50 para IPP em Nelore, com maiores estimativas obtidas nos melhores ambientes. Ambrosini et al. (2016) obtiveram estimativas de herdabilidade para IPP variando de 0,06 a 0,12 em Nelore, com maiores herdabilidades em gradientes ambientais melhores e concluíram que a seleção de touros com maiores VGEs pode ser feita para gradientes ambientais melhores. Lemos et al. (2015) obtiveram estimativas de herdabilidade para a IPP em Nelore variando de 0,05 a 0,47, utilizando as médias do ganho de peso pós-desmame como GA, com maiores estimativas de herdabilidade em gradientes ambientais melhores e concluíram que maior resposta à seleção pode ser obtida em ambientes com manejos e condições nutricionais superiores.

Segundo Oliveira et al. (2018) essa variabilidade das estimativas de herdabilidade estão relacionadas a característica utilizada para a definição do GA, ao banco de dados e a metodologia utilizada. O aumento da estimativa de herdabilidade à medida que houve aumento dos níveis tecnológicos, mensurados indiretamente pelo GA, indica que os animais com maiores valores genéticos médios respondem a seleção quando são produzidos em condições ambientais menos desafiadoras, ou seja, a expressão do potencial genético dos animais é dependente do ambiente.

As correlações genéticas entre GA extremos foram $\geq 0,75$ (Figura 1B). Correlações genéticas inferiores a 0,80 entre os gradientes ambientais, indicam que a variância genética aditiva é influenciada pelo ambiente e que há reclassificação dos animais nos diferentes gradientes ambientais devido a IGA (Araujo Neto et al., 2018; Freitas et al., 2019; Mota et al., 2020). Além disso, sugerem que diferentes grupos de genes são expressos (Pégolo et al., 2009). Em bovinos Nelore, Mota et al. (2020) descreveram correlações genéticas entre os GA extremos de 0,30 e -0,12 para prenhez precoce de novilhas e perímetro escrotal, respectivamente. Freitas et al. (2019) obtiveram correlação genética de 0,39 entre gradientes ambientais extremos ao avaliar a IGA da IPP em búfalas.

2.4.2 Predição genômica da norma de reação (RNM)

Em um programa de melhoramento genético, modelos de norma de reação podem ser utilizados para a estimação de valores genéticos/genômicos específicos para determinado ambiente ou para utilizar EBV/GEBV do parâmetro b_1 para selecionar animais robustos, por exemplo. A maioria dos estudos de normas de reação que estimaram a habilidade de predição para a obtenção do GEBV ou EBV ao longo do gradiente, descreveram poucos resultados para a variável latente b_1 . Outro importante ponto a ser considerado, é que os estudos utilizaram diferentes estatísticas para o cálculo da acurácia.

O incremento de acurácia verificado (Tabela 3) para o parâmetro de inclinação, foi semelhante ao descrito por Araujo Neto et al. (2022), onde, avaliando a resiliência de búfalas leiteiras a alterações de peso no início da lactação, verificaram aumento na acurácia (PEV) média de 5% para o parâmetro b_1 em um modelo de regressão aleatória utilizando o método ssGBLUP.

Para o incremento de acurácia ao longo do GA, os valores foram semelhantes aos descrito por Oliveira et al. (2018), que trabalhando com modelos de normas de reação para peso em bovinos da raça Nelore, verificaram que a utilização do método ssGBLUP possibilitou um incremento médio de acurácia na ordem de 7,9%, comparada a abordagem sem utilização de informação genômica. Entretanto, maiores incrementos foram descritos por Mota et al. (2020), que verificou aumento de acurácia da ordem de 14,61 a 21,18% para perímetro escrotal e de 12,74 à 22,56% para prenhez precoce em novilhas, ao serem incluídas informações genômicas nos modelos de normas de reação.

2.4.3 GWAS do parâmetro de inclinação do RNM

A baixa porcentagem da variância genética aditiva total explicada pelas janelas genômicas, corroboram para definição de característica poligênica para a IPP. Janelas com baixa variância genética aditiva total explicada para a inclinação, também foi relatada em normas de reação para peso ao ano em bovinos Nelore, os quais relataram valores de no máximo 0,82% (Oliveira et al., 2018).

Os resultados da análise funcional, identificaram processos biológicos que sugerem que a IPP é influenciada pela imunidade e pelo metabolismo de lipídios. A janela genômica localizada no BBU18 (51.798.888 – 52.796.278) foi a região genômica com maior quantidade de genes identificados e a única região onde estavam localizados os genes que participaram dos processos biológicos da análise funcional. Dubon et al. (2021) estudando a IPP em bovinos da raça Nelore, também encontraram regiões genômicas importantes no BTA18.

Os genes candidatos APOE, APOC2 e APOC4 localizados no processo biológico fazem parte do grupo de genes das apolipoproteínas. Segundo Veshkini et al. (2022) as apolipoproteínas em geral, são fundamentais para a homeostase do colesterol e, também atuam sobre a função imune, com papéis anti-inflamatórios e pro-inflamatórios em vacas leiteiras.

Os processos biológicos identificados indicaram que há relação entre o metabolismo lipídico e a imunidade com a IPP de búfalas. Segundo Melo-Sterza et al. (2021) a concentração e o perfil lipídico do fluído folicular (FF) influencia na viabilidade e na qualidade dos oócitos de vacas leiteiras, além disso, foi verificado que o estresse térmico atua sobre o metabolismo energético, alterando as concentrações de glicose, IGF1 (fator de crescimento semelhante à insulina tipo 1), ureia, colesterol e NEFA (ácidos graxos não esterificados) do soro e do FF do folículo dominante influenciando na qualidade de oócitos e embriões. Portanto, como a qualidade dos oócitos também é dependente do perfil lipídico do FF, fatores nutricionais da dieta das fêmeas podem alterar a composição lipídica do FF e influenciar na qualidade dos oócitos.

Zachut et al. (2016) avaliando o proteoma do FF de vacas Holandesas israelenses classificadas como menos férteis, identificaram maior expressão de três apolipoproteínas, sendo elas: APOF, APOD e APOC4 e menor expressão da APOC2. E concluíram que as apolipoproteínas atuam como biomarcadores do metabolismo lipídico do FF de vacas que apresentaram menor fertilidade. Portanto, fatores ambientais indiretos como nutrição e estresse térmico parecem influenciar na concentração de lipídios do FF e consequentemente, na qualidade dos oócitos de novilhas bubalinas devido a detecção das apolipoproteínas na IGA da IPP.

Em um estudo do proteoma do FF de folículos de búfalas, a APOC-A1 foi identificada entre as proteínas abundantes, segundo os autores, essa apolipoproteína

já foi identificada em éguas, cadelas, matrizes suínas e búfalas. Além disso, das 363 proteínas identificadas no FF, 32% e 23% pertenciam aos processos biológicos de resposta ao estresse e do sistema imune, respectivamente (Fu et al., 2016). Portanto, os genes APOC2, APOC4, APOE, NECTIN2 e KCNN4 relacionados as rotas de regulação de células T e de metabolismo de lipídios parecem desempenhar um papel importante na proteção do folículo de novilhas Murrah. Marques et al. (2022) caracterizando o proteoma do FF de búfalas, identificaram varias isoformas de apolipoproteínas, no entanto, não foi encontrada nenhuma relação entre essas proteínas e o desenvolvimento folicular.

Ramadan et al. (2001) identificaram que células do sistema imune, como macrófagos, linfócitos, neutrófilos e eosinófilos foram significativamente maiores no corpo lúteo (CL) maduro em relação ao corpo hemorrágico ou no CL em regressão, e concluíram que o sistema imune auxilia nas funções ovarianas (ovulação e luteólise) de búfalas. Segundo Shirasuna et al. (2012), o sistema imunológico é um dos responsáveis pela regulação da função ovariana, principalmente porque as células imunes compõem o CL durante todo o ciclo estral das fêmeas e podem afetar a função endócrina, o desenvolvimento, a manutenção e a regressão do CL. Foi observado que os leucócitos (macrófagos e linfócitos T) aumentaram significativamente no CL em luteólise de vacas, humanos, coelhos, suínos, éguas e camundongos.

Em bovinos, o sistema imunológico é influenciado negativamente pelo estresse térmico. Os linfócitos T (células T) desempenham um papel importante no equilíbrio da função imune, atuando principalmente durante a exposição a agentes estressores (Smith e Vale, 2006). O cortisol é um dos fatores relacionados a inibição da expressão de genes ligados a ativação dos linfócitos T quando os bovinos passam por estresse térmico e comprometem a função dos linfócitos reduzindo a imunidade celular (Lacetera et al., 2006).

Segundo Bagath et al. (2019) as adaptações fisiológicas e celulares que acontecem durante o estresse térmico estão relacionados a redução do desempenho reprodutivo dos animais. Dessa forma, os efeitos da IGA na IPP de novilhas Murrah parece ter correlação com fatores imunológicos visto que os genes das rotas metabólicas também estão relacionados processos biológicos do sistema imune.

Em termos GO, o gene APOE está relacionado ao transporte de lipídios,

processos de metabolismo dos lipídios e homeostase de triglicerídeos. Em vacas leiteiras em transição (periparto), a apolipoproteína APOE está altamente associada ao VLDL (lipoproteína de densidade muito baixa - *Very Low Density Lipoprotein*) e ao HDL (lipoproteína de alta densidade - *High Density Lipoprotein*) e os níveis elevados dessa apolipoproteína acarreta o desenvolvimento de doenças relacionadas ao metabolismo de lipídios, como por exemplo, a esteatose hepática ou fígado gorduroso devido ao comprometimento do metabolismo do colesterol e consequente acúmulo de triglicerídeos nos hepatócitos (Veshkini et al., 2022). Segundo Bernabucci et al. (2004), alterações na síntese de VLDL no fígado de vacas leiteiras de transição estão relacionadas a sinais clínicos de esteatose hepática, e que há três proteínas envolvidas na síntese e secreção de VLDL no fígado, sendo elas: ApoB100, APOE e MTP (proteína de transferência de triglicerídeos). Sendo a APOE mais relacionada com a produção de VLDL pelos hepatócitos e a secreção de VLDL pelo fígado.

O gene APOC2 em termos GO está associado ao transporte de lipídios, ao processo metabólico lipídico e a regulação positiva do processo catabólico de triglicerídeos. O gene APOC4 foi detectado no plasma das vacas suplementadas com ácidos graxos essenciais e ácido linoleico conjugado após o parto, seu papel, no entanto, não está bem definido em bovinos, mas em humanos, está associado ao metabolismo de triglicerídeos e de VLDL (Veshkini et al., 2022). Kumar et al. (2021) conduziram um estudo do proteoma do FF pós-parto de búfalas Murrah e a análise da rede de interação proteína-proteína revelou que as apolipoproteínas APOA1, APOA2, APOA4 e APOC3 interagiram entre si e com outras proteínas, e estão entre as proteínas que desempenham um papel fisiológico e endócrino importante no ovário, participando da ovulação, regressão do corpo lúteo, esteroidogênese e apoptose celular no folículo.

O gene APOE foi localizado próximo a um SNP ligado a eficiência reprodutiva de éguas Pura Raza Espanõl, juntamente com o gene NECTIN2, foram relacionados a processos biológicos de desenvolvimento e maturação de gametas de machos e fêmeas (Laseca et al., 2022). O gene NECTIN2 faz parte da superfamília das imunoglobulinas que são moléculas de adesão celular e é expresso nas células de Sertoli. Camundongos com alterações no gene NECTIN2 tiveram produção normal de espermatozoides, mas tiveram problemas na formação da cabeça e da peça

intermediária dos espermatozoides o que prejudicou a penetração dos espermatozoides nos óvulos tornando os machos inférteis (Mueller et al., 2003). Segundo Mueller et al. (2003), camundongos deficientes de NECTIN2 tiveram a junção Sertoli-Espermátide formadas de forma inadequada. Homens com mutações em regiões codificadoras dos genes NECTIN2 e NECTIN3 também apresentaram anormalidades anatômicas nos espermatozoides, porém a produção de espermatozoides não foi prejudicada (Bronson et al., 2017). Os termos GO relacionados com o gene NECTIN2 incluem regulação positiva da resposta imune mediada por imunoglobulinas e organização do citoesqueleto.

O gene KCNN4 (*Potassium Calcium-Activated Channel Subfamily N Member 4*) é o membro 4 da subfamília N do canal de potássio ativado por cálcio (Chen et al., 2022), denominado também de KCa3.1. Os termos GO relacionados com o KCNN4 incluem processos do sistema imunológico, transporte de íons de potássio e cálcio e homeostase do volume celular. A proteína 4 do canal intracelular de cloreto (CLIC4), foi identificada no FF de búfalas e foi relacionada a processos celulares, como a formação de canais iônicos e remodelamento das membranas intracelulares e foi sugerida como uma proteína que desempenha um papel no transporte de íons nos folículos (Fu et al., 2016). Dessa forma, as proteínas relacionadas com os canais de potássio, como a KCNN4, também podem estar desempenhando um papel chave no transporte de potássio para os folículos.

2.5 Conclusão

Os resultados dos componentes de variância e estimativas de herdabilidade, bem com as correlações genéticas entre gradientes ambientais extremos indicaram que há efeito da IGA na IPP em búfalos leiteiros. A inclusão de informações genômicas aumentou a acurácia dos valores genéticos do intercepto e da inclinação da norma de reação, indicando melhor ajuste dos dados quando a metodologia ssGBLUP foi utilizada. O modelo de associação genômica ampla permitiu a identificação de genes importantes envolvidos nos efeitos da inclinação da norma de reação da IPP, onde cinco genes principais foram envolvidos em dois processos biológicos ligados ao metabolismo de lipídios e imunidade.

2.6 Referências

Agudelo-Gómez DA, Savegnago RP, Buzanskas ME, Ferraudo AS, Prado Munari D, Céron-Muñoz (2015) Genetic principal components for reproductive and productive traits in dual-purpose buffaloes in Colombia. **Journal of Animal Science** 93:3801–3809. <https://doi.org/10.2527/jas2015-8940>

Ambrosini DP, Malhado CHM, Filho RM, Cardoso FF, Carneiro PLS (2016) Genotype × environment interactions in reproductive traits of Nellore cattle in northeastern Brazil. **Tropical Animal Health and Production** 48:1401–1407. <https://doi.org/10.1007/s11250-016-1105-7>

Amin AMS, Abo-Ismael MK, Salem MMI (2021) Genetic parameters and genetic trends for reproductive traits in Egyptian buffalo. **Animal Reproduction Science** 231. <https://doi.org/10.1016/j.anireprosci.2021.106800>

Araujo Neto FR, Pegolo NT, Aspilcueta-Borquis RR, Pessoa MC, Bonifácio A, Lobo RB, Oliveira HN (2018) Study of the effect of genotype–environment interaction on age at first calving and production traits in Nellore cattle using multi-trait reaction norms and Bayesian inference. **Animal Science Journal** 89:939–945. <https://doi.org/10.1111/asj.12994>

Araujo Neto FR, Takada L, Santos DJA, Aspilcueta-Borquis RR, Cardoso DF, Nascimento AV, Leão KM, Oliveira HN, Tonhati H (2020) Identification of genomic regions related to age at first calving and first calving interval in water buffalo using single-step GBLUP. **Reproduction in Domestic Animals** 55:1565–1572. <https://doi.org/10.1111/rda.13811>

Araujo Neto FR, Santos JCG, Silva Arce CD, Aspilcueta-Borquis RR, Santos DJA, Guimarães KC, Nascimento, AV, Oliveira HN, Humberto T (2022) Genomic study of the resilience of buffalo cows to a negative energy balance. **Journal of Applied Genetics** 63:379–388. <https://doi.org/10.1007/s13353-021-00680-x>

Aspilcueta-Borquis RR, Araujo Neto FR, Baldi F, Bignardi AB, Albuquerque LG, Tonhati H (2010) Genetic parameters for buffalo milk yield and milk quality traits using Bayesian inference. **Journal of Dairy Science** 93:2195–2201. <https://doi.org/10.3168/jds.2009-2621>

Aspilcueta-Borquis RR, Seno LO, Araujo Neto FR, Santos DJA, Hurtado-Lugo NA, Tonhati H (2022) Lifetime productivity: Genetic study of longevity and its associations with economically important traits in dairy buffaloes. **Livestock Science** 259. <https://doi.org/10.1016/j.livsci.2022.104900>

Atashi H, Asaadi A, Hostens M (2021) Association between age at first calving and lactation performance, lactation curve, calving interval, calf birth weight, and dystocia in Holstein dairy cows. **Plos One**, 16. <https://doi.org/10.1371/journal.pone.0244825>

Bagath M, Krishnan G, Devaraj C, Rashamol VP, Pragna P, Lees AM, Sejian V (2019) The impact of heat stress on the immune system in dairy cattle: A review. **Research in Veterinary Science** 126:94-102. <https://doi.org/10.1016/j.rvsc.2019.08.011>

Bernabucci U, Ronchi B, Basiricò L, Pirazzi D, Rueca F, Lacetera N, Nardone A (2004) Abundance of mRNA of apolipoprotein B100, apolipoprotein E, and microsomal triglyceride transfer protein in liver from periparturient dairy cows. **Journal of Dairy Science** 87:2881–2888. [https://doi.org/10.3168/jds.S0022-0302\(04\)73418-9](https://doi.org/10.3168/jds.S0022-0302(04)73418-9)

Berry DP, Wall E, Pryce JE (2014) Genetics and genomics of reproductive performance in dairy and beef cattle. **Animal** 8:105–121. <https://doi.org/10.1017/S1751731114000743>

Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J (2009) ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. **Bioinformatics** 25:1091-1093. <https://doi.org/10.1093/bioinformatics/btp101>

Bronson R, Mikhailik A, Schwedes J, Gnatenko D, Hatchwell E (2017) Detection of candidate nectin gene mutations in infertile men with severe teratospermia. **Journal of Assisted Reproduction and Genetics** 34:1295–1302. <https://doi.org/10.1007/s10815-017-0985-4>

Camargo GMF, Aspilcueta-Borquis RR, Fortes MRS, Porto-Neto R, Cardoso DF, Santos DJA, Lehnert SA, Reverter A, Moore SS, Tonhati, H (2015) Prospecting major genes in dairy buffaloes. **BMC Genomics** 16. <https://doi.org/10.1186/s12864-015-1986-2>

Carvalho R, Costilla R, Neves HHR, Albuquerque LG, Moore S, Hayes BJ (2019) Unraveling genetic sensitivity of beef cattle to environmental variation under tropical conditions. **Genetics Selection Evolution** 51. <https://doi.org/10.1186/s12711-019-0470-x>

Cesarani A, Gaspa G, Masuda Y, Degano L, Vicario D, Lourenco DAL, Macciotta NPP (2020) Variance components using genomic information for 2 functional traits in Italian Simmental cattle: Calving interval and lactation persistency. **Journal of Dairy Science** 103:5227–5233. <https://doi.org/10.3168/jds.2019-17421>

Chen S, Su X, Mo Z (2022) KCNN4 is a Potential Biomarker for Predicting Cancer Prognosis and an Essential Molecule that Remodels Various Components in the Tumor Microenvironment: A Pan-Cancer Study. **Frontiers in Molecular Biosciences** 3. <https://doi.org/10.3389/fmolb.2022.812815>

Chiaia HLJ, Lemos MVA, Venturini GC, Aboujaoude C, Berton MP, Feitosa FB, Carneiro R, Albuquerque LG, Oliveira HN, Baldi F (2015) Genotype x environment interaction for age at first calving, scrotal circumference, and yearling weight in Nellore cattle using reaction norms in multitrait random regression models. **Journal of Animal Science** 93:1503–1510. <https://doi.org/10.2527/jas.2014-8217>

Dubon MAC, Pedrosa VB, Feitosa FLB, Costa RB, de Camargo GMF, Silva MR, Pinto LFB (2021) Identification of novel candidate genes for age at first calving in Nellore cows using a SNP chip specifically developed for *Bos taurus indicus* cattle. **Theriogenology** 173:156-162. <https://doi.org/10.1016/j.theriogenology.2021.08.011>

Freitas GR, Hurtado-Lugo NA, Santos DJA, Aspilcueta Borquis RR, Pegolo NT, Tonhati H, Araújo Neto FR (2019) Genotype–environment interaction for age at first calving in buffaloes, using the reaction norm model. **Reproduction in Domestic Animals** 54:727–732. <https://doi.org/10.1111/rda.13414>

Fu Q, Huang Y, Wang Z, Chen F, Huang D, Lu Y, Liang X, Zhang M (2016) Proteome Profile and Quantitative Proteomic Analysis of Buffalo (*Bubalus bubalis*) Follicular Fluid during Follicle Development. **International Journal of Molecular Sciences** 17. <https://doi.org/10.3390/ijms17050618>

Iamartino D et al. (2017) Design and validation of a 90K SNP genotyping assay for the water buffalo (*Bubalus bubalis*). **Plos One** 12. <https://doi.org/10.1371/journal.pone.0185220>

Jamuna V, Chakravarty AK (2016) Evaluation of fertility in relation to milk production and productivity of Murrah buffaloes. **Animal Reproduction Science** 171:72–80. <https://doi.org/10.1016/j.anireprosci.2016.06.001>

Kumar S, Balhara AK, Buragohain L, Kumar R, Sharma RK, Phulia SK, Mohanty AK, Singh I (2021) Identification of novel proteomics markers involved in ovarian endocrinology of riverine buffalo (*Bubalus bubalis*). **Biological Rhythm Research** 52:1448-1460. <https://doi.org/10.1080/09291016.2019.1658061>

Laseca N, Demyda-Peyrás S, Valera M, Ramón M, Escribano B, Perdomo-González DI, Molina A (2022) A genome-wide association study of mare fertility in the Pura Raza Español horse. **Animal** 16. <https://doi.org/10.1016/j.animal.2022.100476>

Lacetera N, Bernabucci U, Scalia D, Basiricò L, Morera P, Nardone A (2006) Heat stress elicits different responses in peripheral blood mononuclear cells from Brown Swiss and Holstein Cows. **Journal of Dairy Science** 89:4606–4612. [https://doi.org/10.3168/jds.s0022-0302\(06\)72510-3](https://doi.org/10.3168/jds.s0022-0302(06)72510-3)

Legarra A, Reverter A (2018) Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. **Genetics Selection Evolution** 50. <https://doi.org/10.1186/s12711-018-0426-6>

Lemos MVA, Chiaia HLJ, Berton MP, Feitosa FLB, Aboujaoude C, Venturini GC, Oliveira HN, Albuquerque LG, Baldi F (2015) Reaction norms for the study of genotype-environment interaction for growth and indicator traits of sexual precocity in Nellore cattle. **Genetics and Molecular Research** 14:7151-7162. <https://doi.org/10.4238/2015.June.29.9>

Low WY, Tearle R, Bickhart DM et al. (2019) Chromosome-level assembly of the water buffalo genome surpasses human and goat genomes in sequence contiguity. **Nature communications** 10. <https://doi.org/10.1038/s41467-018-08260-0>

Malhado CHM, Malhado ACM, Ramos AA, Carneiro PLS, Souza JC, Pala A (2013) Genetic parameters for milk yield, lactation length and calving intervals of Murrah buffaloes from Brazil. **Revista Brasileira de Zootecnia** 42:565–569. <https://doi.org/10.1590/S1516-35982013000800005>

Marques NFS, Codognoto VM, Souza FF, Scott C, Janini LCZ, Brochine S, Tironi SMT, Camargo LS, Alvarez MVN, Oba E (2022) Proteomics of follicular fluid from buffaloes (*Bubalus bubalis*): Unraveling the secrets of follicular development. **Livestock Science** 260. <https://doi.org/10.1016/j.livsci.2022.104947>

Melo-Sterza FA, Poehland R (2021) Lipid Metabolism in Bovine Oocytes and Early Embryos under In Vivo, In Vitro, and Stress Conditions. **International Journal of Molecular Science** 26. <https://doi.org/10.3390/ijms22073421>

Misztal I, Tsuruta S, Lourenço D, Masuda Y, Aguilar I, Legarra A, Vitezica Z (2014) BLUPF90 family of programs. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all8.pdf

Mota LFM, Fernandes Jr GA, Herrera AC, Scaletz DCB, Espigolan R, Magalhães AFB, Carneiro R, Baldi F, Albuquerque LG (2020) Genomic reaction norm models exploiting genotype × environment interaction on sexual precocity indicator traits in Nellore cattle. **Animal Genetics** 51:210–223. <https://doi.org/10.1111/age.12902>

Mueller S, Rosenquist TA, Takai Y, Bronson RA, Wimmer E (2003) Loss of Nectin-2 at sertoli-spermatid junctions leads to male infertility and correlates with severe spermatozoan head and midpiece malformation, impaired binding to the zona pellucida, and oocyte penetration. **Biology of Reproduction** 69:1330–1340. <https://doi.org/10.1095/biolreprod.102.014670>

Oliveira DP et al. (2018) Reaction norm for yearling weight in beef cattle using single-step genomic evaluation. **Journal of Animal Science** 96:27–34. <https://doi.org/10.1093/jas/skx006>

Pégolo NT, Oliveira HN, Albuquerque LG, Bezerra LA, Lôbo RB (2009) Genotype by environment interaction for 450-day weight of Nelore cattle analyzed by reaction norm models. **Genetics and Molecular Biology** 32:281-287. <https://doi.org/10.1590/S1415-47572009005000027>

Ramadan AA, Selim SA, Hassan HH, Wahba MA (2001) Immune regulation of ovarian function in buffaloes (*Bubalus bubalus*). **Theriogenology** 55:661-669. [https://doi.org/10.1016/S0093-691X\(01\)00434-4](https://doi.org/10.1016/S0093-691X(01)00434-4)

Roso VM, Schenkel FS (2006) AMC – A computer program to assess the degree of connectedness among contemporary groups. 8th World Congress on Genetics Applied to Livestock Production. <https://www.researchgate.net/publication/285687222>

Santos JC, Malhado CHM, Carneiro PLS, Rezende MPG, Cobuci JA (2020) Genotype-environment interaction for age at first calving in Holstein cows in Brazil. *Veterinary and Animal Science* 9 <https://doi.org/10.1016/j.vas.2020.100098>

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome Research** 13:2498-504. <https://doi.org/10.1101/gr.1239303>

Sherwin VE, Hudson CD, Henderson A, Green MJ (2016) The association between age at first calving and survival of first lactation heifers within dairy herds. **Animal** 10:1877–1882. <https://doi.org/10.1017/S1751731116000689>

Shirasuna K, Nitta A, Sineenard J, Shimizu T, Bollwein H, Miyamoto A (2012) Vascular and immune regulation of corpus luteum development, maintenance, and regression in the cow. **Domestic Animal Endocrinology** 43:198-211. <https://doi.org/10.1016/j.domaniend.2012.03.007>

Smith SM, Vale WW (2006) The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress. **Dialogues in Clinical Neuroscience** 8:383–395. <https://doi.org/10.31887/DCNS.2006.8.4/ssmith>

Tamboli P, Bharadwaj A, Chaurasiya A, Bangar YC, Jerome A (2022) Association between age at first calving, first lactation traits and lifetime productivity in Murrah buffaloes. **Animal Bioscience** 35:1151–1161. <https://doi.org/10.5713/ab.21.0182>

Veshkini A, Hammon HM, Vogel L, Viala D, Delosi re M, Tr scher A, D jean S, Ceciliani F, Sauerwein H, Bonnet M (2022) Plasma proteomics reveals crosstalk between lipid metabolism and immunity in dairy cows receiving essential fatty acids and conjugated linoleic acid. **Scientific Reports** 12. <https://doi.org/10.1038/s41598-022-09437-w>

Zachut M, Sood P, Levin Y, Moallem U (2016) Proteomic analysis of preovulatory follicular fluid reveals differentially abundant proteins in less fertile dairy cows. **Journal of Proteomics** 139:122-129. <https://doi.org/10.1016/j.jprot.2016.03.027>

CAPÍTULO 3 – Comparação de métodos single-step para a predição genômica da idade ao primeiro parto em búfalos leiteiros

RESUMO – A idade ao primeiro parto é uma característica importante que deve ser considerada nos programas de melhoramento genético de búfalos leiteiros, assim como o uso de novas metodologias para a sua utilização. Este estudo teve como objetivo comparar a capacidade preditiva de diferentes métodos de passo único na estimação da idade ao primeiro parto de búfalos Murrah. Foram utilizadas informações fenotípicas de 2.247 animais, registros genealógicos de 3.363 animais, e marcadores SNP para 553 animais. Foram analisados os seguintes métodos: predição baseada em pedigree (PBLUP), GBLUP de passo único (ssGBLUP), GBLUP de passo único ponderado (WssGBLUP), e métodos de regressão Bayesianos de passo único (ssBA, ssBB π , ssBC π , ssBL e ssBRR). Para comparar as metodologias, a acurácia e dispersão dos EBVs/GEBVs foram estimadas utilizando o método LR. As estimativas de acurácia para os animais genotipados variaram entre 0,30 (ssBRR) e 0,39 (WssGBLUP). Em relação à dispersão da predição do GEBV, os métodos ssBC π (0,99) e WssGBLUP (1,00) obtiveram os melhores resultados com os animais genotipados. Os resultados indicaram que a utilização de informação genômica pode melhorar o ganho genético para a idade ao primeiro parto, aumentando a acurácia e reduzindo a inflação/deflação das predições desta característica em comparação com o modelo tradicional baseado em pedigree. Além disso, entre todos os modelos genômicos de passo único estudados, o WssGBLUP e o ssBA foram os métodos mais vantajosos a serem utilizados na avaliação genômica da idade ao primeiro parto de búfalos desta população.

Palavras-chave: BLUPs, inferência bayesiana, seleção genômica, SNPs

CHAPTER 3 - Comparison of single-step methods for genomic prediction of age at first calving in dairy buffaloes

ABSTRACT - The age at first calving (AFC) is an important trait to be considered in genetic improvement programs for dairy buffaloes, justifying the application of new technologies for its use. Thus, this work was designed with the aim of comparing the predictive ability of different single-step methods using age at first calving information in Murrah buffaloes. Phenotypic information from 2,247 animals, genealogical records from 3,363 animals, and SNP markers for 553 animals were used. The following methods were analyzed: pedigree-based prediction (PBLUP), single-step GBLUP (ssGBLUP), weighted single-step GBLUP (WssGBLUP), and single-step Bayesian regression methods (ssBA, ssBB π , ssBC π , ssBL and ssBRR). To compare the methodologies, the selection accuracy and dispersion of EBVs/GBEVs were estimated using the LR method. Accuracy estimates for the genotyped animals ranged from 0.30 (single-step Bayes Ridge Regression) to 0.39 (WssGBLUP). Regarding the dispersion of the GEBV prediction, ssBC π (0.99) and WssGBLUP (1.00) obtained the best results with the genotyped animals. The results indicate that the use of genomic information can improve the genetic gain for age at first calving by increasing the accuracy and reducing inflation/deflation of predictions of this trait compared to the traditional pedigree-based model. In addition, among all genomic single-step models studied, WssGBLUP and ssBA were the most advantageous methods to be used in the genomic evaluation of age at first calving of buffaloes from this population

Keywords: BLUPs, Bayesian Inference, genomic selection, SNPs

3.1 Introdução

A idade ao primeiro parto (IPP) é uma importante característica indicadora da eficiência reprodutiva, sendo, portanto, necessária sua inclusão em programas de melhoramento genético para a espécie bubalina (Aspilcueta-Borquis et al., 2022). Sabe-se, contudo, que o progresso genético obtido pela seleção direta da IPP pode ser muito lento devido à baixa magnitude das estimativas de herdabilidade desta característica em búfalos (Camargo et al., 2015). A avaliação genômica é a melhor estratégia para a predição dos valores genéticos da IPP, uma vez que proporciona aumento de acurácia em relação a avaliação tradicional (Calus et al., 2010). Além disso, Camargo et al. (2015) e Araujo Neto et al. (2020a) relataram regiões genômicas associadas a esta característica em estudos de GWAS com esta população, evidenciando a possibilidade de inclusão de informações genômicas em programas de melhoramento da espécie.

O método GBLUP (ssGBLUP) de passo único tornou-se a principal metodologia utilizada para avaliações genômicas em bovinos leiteiros e búfalos (Araujo Neto et al., 2020a; Araujo Neto et al., 2020b; Cesarani et al., 2021; Lazáro et al., 2021), dada a facilidade de integrar as matrizes de relacionamentos, baseadas na informação de pedigree (A) e genômica (G - *genetic relationship matrix* - GRM), sem grandes alterações nas equações do modelo misto (Misztal et al., 2020). Além disso, esta metodologia permite atribuir pesos diferentes para cada marcador, originando o método GBLUP de passo único ponderado - WssGBLUP (Zhang et al., 2016). Posteriormente, Fernando et al. (2014) propuseram metodologias de passo único combinadas com métodos de regressão Bayesianos (ssBR), como opção ao ssGBLUP. No entanto, poucos estudos foram desenvolvidos comparando os resultados das diferentes metodologias de passo único (Lee et al. 2017; Gao et al. 2018; Zhou et al., 2018). Assim, o objetivo deste estudo foi comparar a capacidade preditiva de diferentes métodos de passo único na estimação de valores genômicos para a IPP em búfalos da raça Murrah.

3.2 Material e métodos

3.2.1 Dados

Para a realização deste estudo, foram utilizados os registros de idade ao primeiro parto (IPP) de 2.247 fêmeas da raça Murrah, nascidas entre 1995 e 2016, pertencentes a 4 fazendas de produção de búfalos. Os grupos de contemporâneos (GC) foram formados considerando animais nascidos na mesma fazenda, ano e estação de nascimento. As estações de nascimento foram definidas como seca e chuvosa. Animais com registros abaixo ou acima de três desvios padrão da média do CG e animais pertencentes a GC com menos de cinco indivíduos foram removidos. O arquivo de pedigree totalizou 3.363 animais.

Um total de 553 animais (539 fêmeas e 14 touros) foram genotipados com o painel Axiom Buffalo Genotyping Array 90 K (Iamartino et al., 2017). Apenas os SNPs presentes nos cromossomos autossômicos, bem como aqueles com call rate maior que 95%, MAF maior que 3% e nível de significância para o teste de equilíbrio de Hardy-Weinberg de 10^{-5} e, amostras com call rate maior que 90% foram mantidos na análise. A base de dados (fenotípica e genotípica) após a etapa de consistência foi descrita na Tabela 1.

Tabela 1. Descrição do conjunto de dados de idade ao primeiro parto obtidos após a consistência

Descrição	Número de registros	
	Conjunto de dados completo	Conjunto de dados parcial
Nº de animais com fenótipos	2.247	1.596
Nº de animais com genótipos e fenótipos	509	420
Nº de grupos de contemporâneos	101	73
Média	37,34	36,97
Desvio-padrão	5,52	5,07

3.2.2 Modelos PBLUP, ssGBLUP e WssGBLUP

O conjunto de metodologias utilizado para a estimação dos valores genéticos/genômicos baseia-se na utilização de uma matriz de parentesco entre indivíduos em equações de modelos mistos, sendo as diferenças constituídas pelo tipo de informação utilizada (pedigree ou genotípica) e a forma como estas matrizes são construídas. O modelo BLUP (PBLUP) baseado em pedigree pode ser descrito como:

$$y = X\beta + Z\alpha + e$$

em que, y é o vetor dos registros fenotípicos (IPP); β é o vetor de efeitos fixos (GC); X é a matriz de incidência que associa β a y ; α é o vetor dos efeitos aditivos; Z é a matriz de incidência que associa α a y e e é o vetor dos efeitos aleatórios. As seguintes suposições foram feitas:

$$E(y) = X\beta \quad E(\alpha) = 0 \quad E(e) = 0$$

$$Var(\alpha) = A\sigma_\alpha^2 \quad Var(e) = I_n\sigma_e^2$$

em que σ_α^2 e σ_e^2 representam as variâncias genéticas aditivas e residuais, respectivamente; A e I_n são os relacionamentos baseados em pedigree e uma matriz identidade, respectivamente.

O método ssGBLUP é uma extensão do modelo PBLUP, no qual as matrizes de parentesco de pedigree (A) e o parentesco genômico (G) são combinadas em uma única matriz, denominada H , descrita por Aguilar et al. (2010). Assim, o seu inverso pode ser obtido como:

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

em que A^{-1} e G^{-1} são matrizes inversas A e G , respectivamente, e $-A_{22}^{-1}$ é a inversa

de A somente de animais genotipados. A matriz G foi obtida de acordo com VanRaden (2008) como:

$$\mathbf{G} = \frac{\mathbf{ZDZ}'}{2 \sum p_i(1 - p_i)},$$

onde p_i é frequência do alelo menor (MAF) do SNP i ; $\mathbf{Z}$ é a matriz relacionando os genótipos de cada locus centrada por frequências alélicas e $\mathbf{D}$ é a matriz diagonal de pesos para as variâncias dos SNPs, (d_i = peso do SNP i). O modelo ssGBLUP é definido como $\mathbf{D} = \mathbf{I}$, de modo que o peso de todos os SNPs seja igual a 1.

No método WssGBLUP (ssGBLUP ponderado; Wang et al., 2012) são atribuídos aos marcadores diferentes pesos por meio de um processo iterativo com as seguintes etapas:

1. Definição da matriz diagonal de variância ou ponderação do SNP como uma identidade, $\mathbf{D}^{(t)} = \mathbf{I}$, em que $t = 1$;
2. A matriz de relacionamento genômico é configurada para t como $\mathbf{G}^{(t)} = \mathbf{ZD}^{(t)}\mathbf{Z}' / (2 \sum p_i(1 - p_i))$;
3. Os valores genéticos genômicos (GEBV ou $\hat{\alpha}$) para t são obtidos;
4. Os GEBV $\hat{\alpha}^{(t)}$ são convertidos para efeitos de SNP $\hat{u}$ como $\hat{u}^t = \mathbf{D}^{(t)}\mathbf{Z}'(\mathbf{G}^{(t)})^{-1}\hat{\alpha}^{(t)}$;
5. O peso do i th SNP (o i th elemento de $\mathbf{D}$ ou d_i) é calculado;
6. Os pesos dos SNPs são normalizados em $\mathbf{D}^{(t+1)}$ para ter variâncias genéticas constantes nos efeitos dos SNPs;
7. Três iterações a partir da etapa 2

Para os métodos acima mencionados, o programa BLUPF90++ (Misztal et al., 2014) foi utilizado para a estimativa dos componentes de variância e para a etapa de validação. A estimativa dos efeitos dos marcadores e o cálculo da ponderação (para o método WssGBLUP) foram realizados com o programa POSTGSF90 (Misztal et al., 2014). Na etapa de ponderação dos marcadores, os pesos foram definidos por uma função não linear proposta por VanRaden et al. (2009).

3.2.3 Modelos de regressão Bayesianas em passo único (ssBR)

Nos modelos Bayesianos em passo único, o modelo generalizado para a combinação das informações de animais genotipados e não-genotipados pode ser descrito como indicado por Fernando et al. (2014):

$$\begin{bmatrix} y_n \\ y_g \end{bmatrix} = \begin{bmatrix} X_n \\ X_g \end{bmatrix} \beta + \begin{bmatrix} Z_n & 0 \\ 0 & Z_g \end{bmatrix} \begin{bmatrix} M_n \alpha + \epsilon \\ M_g \alpha \end{bmatrix} + e$$

onde os vetores e matrizes para animais não genotipados são denotados com o subscrito n e para os animais com genótipos com o subscrito g . Assim, y_n e y_g representam os vetores de observações fenotípicas; β é o vetor de efeitos sistemáticos; X_n e X_g matrizes de incidência para os efeitos fixos; Z_n e Z_g representam matrizes de incidência associados aos valores genômicos dos animais; α , ϵ e e , representam nesta ordem, os vetores de efeitos de marcadores, resíduos da imputação e vetor de resíduos do modelo; M_g é a matriz de marcadores para os animais genotipados e M_n é a matriz de marcadores imputados para os animais não genotipados. A matriz M_n pode ser obtida como:

$$M_n = A_{ng} A_{22}^{-1} M_g$$

em que A_{ng} representa a matriz de parentesco com base no pedigree entre os animais genotipados e não genotipados e; A_{22}^{-1} representa a inversa da matriz de parentesco com base no pedigree entre animais genotipados.

Os seguintes métodos ssBR foram utilizados: BayesA (ssBA- Meuwissen et al., 2001), BayesB π (ssBB π - Meuwissen et al., 2001), BayesC π (ssBC π - Habier et al., 2011), Bayes Lasso (ssBL- Yi and Xu, 2008) e Bayes Ridge Regression (ssBRR- Campos et al., 2013), onde as pressuposições de cada um são mostradas na Tabela 2. O pacote hibayes (Yin et al., 2022), disponível para o programa estatístico R (R core Team, 2021), foi utilizado para executar as análises. Um total de 350,000 amostras e um período de burn-in de 150,000 amostras foram gerados, a convergência foi avaliada utilizando análises gráficas no programa estatístico R (R core Team, 2021).

Tabela 2. Descrição das pressuposições dos diferentes métodos de regressão bayesiana de passo único, empregados para análise genômica da idade ao primeiro parto em búfalas leiteiras

Modelos	Pressuposições
ssBA	$\alpha_j N)$ $\sigma_{\alpha j}^2 \chi^{-2}(v, S)$
ssBB π	$\alpha_j (1 - \pi)N(0, \sigma_{\alpha}^2) + \pi\delta_0$ $\sigma_{\alpha j}^2 \chi^{-2}(v, S)$
ssBC π	$\alpha_j (1 - \pi)N(0, \sigma_{\alpha}^2) + \pi\delta_0$ $\alpha_j N)$
ssBL	$\alpha_j N)$ $\sigma_{\alpha j}^2 Expon(\lambda^2/2)$
ssBRR	$\alpha_j N)$

ssBA: Single step Bayes A; ssBB π : Single step Bayes B π ; ssBC π : Single step Bayes C π ; ssBL: Single step Bayes Lasso; ssBRR: Single step Bayes Ridge Regression

3.2.4 Validação e análise de GEBVs

Utilizando as soluções dos SNPs, obtidas com os diferentes métodos, foi calculada a distância euclidiana entre os métodos, a fim de verificar a similaridade na estimativa dos efeitos dos marcadores e, subsequentemente, foi realizada uma análise de agrupamento, utilizando o método hierárquico. Para esta análise, foram utilizadas as funções *dist* e *hclust*, presente no programa R.

Para avaliar a capacidade preditiva dos métodos, foram utilizadas as estatísticas de acurácia e dispersão propostas por Legarra e Reverter (2018). Esta metodologia baseia-se na predição do GEBV utilizando um conjunto de dados completo ($\hat{u}_w$) e um conjunto de dados parcial ($\hat{u}_p$). Sendo o $\hat{u}_p$ formado pela omissão de registros fenotípicos de alguns animais (animais-foco). Neste trabalho, o conjunto de dados de validação (banco parcial) foi composto por animais nascidos após o ano de 2010. Em seguida, a acurácia e a dispersão das predições foram avaliados como:

$$r_{AA} = \sqrt{\frac{\text{Acurácia}}{\text{Dispersão}}} = \sqrt{\frac{\text{cov}(\hat{u}_w, \hat{u}_p)}{(1 - \bar{F})\sigma_g^2}} \quad b_{w,p} = \frac{\text{cov}(\hat{u}_w, \hat{u}_p)}{\text{var}(\hat{u}_p)}$$

onde $\bar{F}$ é o coeficiente de endogamia médio dos animais-foco, e σ_g^2 é a variância genética aditiva. É importante ressaltar que a estatística de acurácia utilizada, consiste em uma modificação da forma proposta por Legarra e Reverter (2018), uma vez que: a) considera a variância genética estimada, pois as fazendas utilizam como critérios de seleção a produção e qualidade do leite e; b) como indicado no próprio artigo dos autores, o denominador da fórmula pode ser reduzido ao assumir a coancestria com valor próximo a da endogamia.

3.3 Resultados

De acordo com o dendrograma obtido pela análise de cluster dos efeitos dos marcadores estimados utilizando diferentes modelos de passo único, o modelo ssBA mostrou maior semelhança com o WssGBLUP e o ssGBLUP (que estavam ainda mais próximos no primeiro cluster) do que com os outros modelos ssBR (Figura 1).

Comparando os manhattan-plots das soluções dos SNPs dos modelos de passo único, o efeito de *shrinkage* foi mais evidente, com sinais claros de grande efeito em algumas regiões genômicas, quando os modelos WssGBLUP, ssBC π e ssBB π foram utilizados. O modelo WssGBLUP, apesar da sua semelhança com o ssGBLUP na análise de agrupamento, mostrou um padrão ainda mais distinto na estimativa dos efeitos dos SNPs, com a indicação de marcadores com efeitos de destaque em muitos cromossomos.

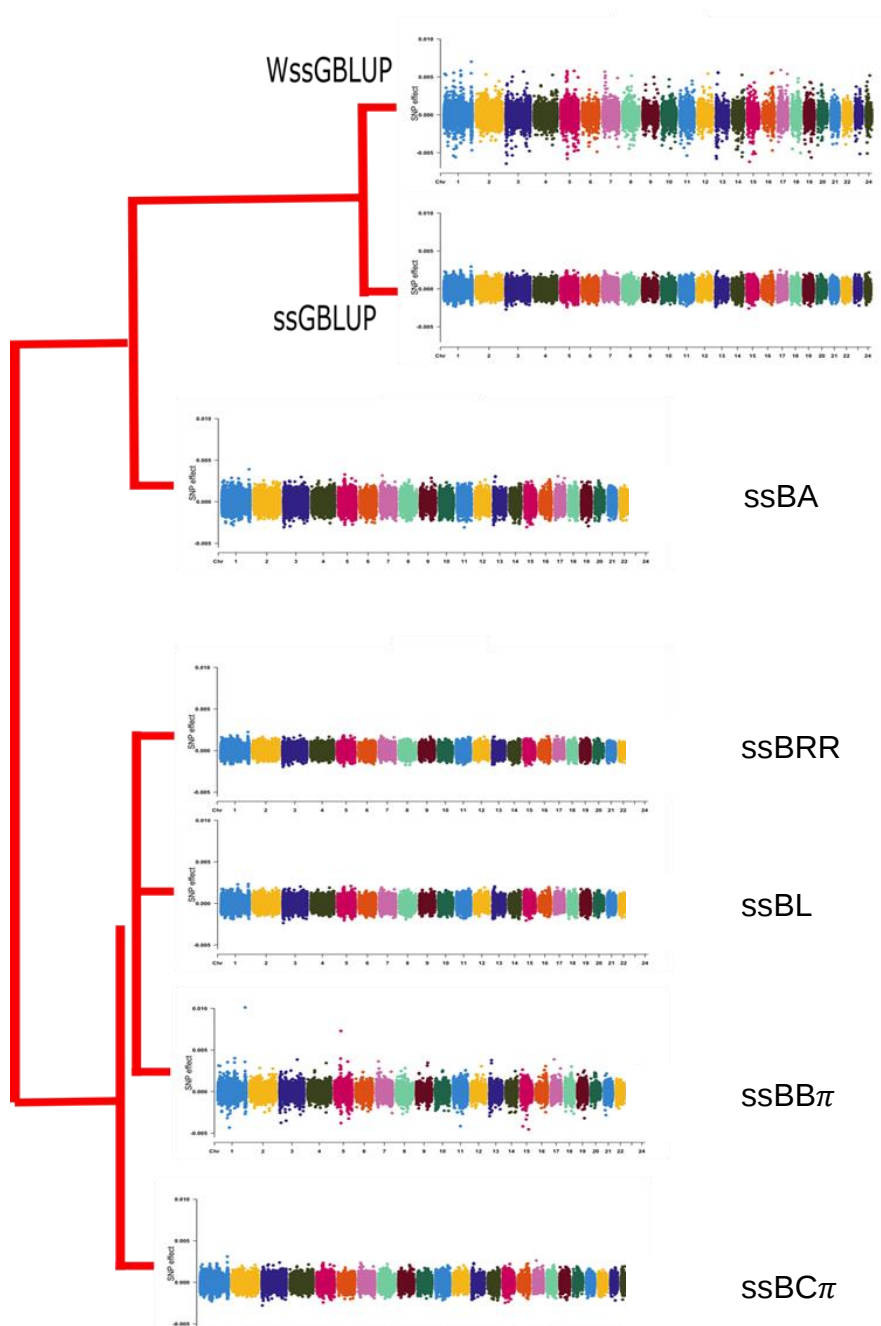


Figura 1. Estimativas dos efeitos dos marcadores SNP (manhattan plot) para os diferentes métodos empregados*, e análise de agrupamento baseado na similaridade entre os efeitos dos marcadores

*ssBA: Single step Bayes A; ssBB π : Single step Bayes B π ; ssBC π : Single step Bayes C π ; ssBL: Single step Bayes Lasso; ssBRR: Single step Bayes Ridge Regression

As estimativas de acurácia para os animais genotipados variavam entre 0,30 (ssBRR) a 0,39 (WssGBLUP) (Tabela 3). Foram observadas estimativas de baixa acurácia para animais não genotipados, com valores que variaram entre 0,29

(ssGBLUP) a 0,33 (ssBA). O modelo PBLUP apresentou acurácias de 0,30 e 0,29 para animais genotipados e não genotipados, respectivamente. Quanto à dispersão da predição do GEBV (Tabela 3), os modelos ssBC π (0,99) e WssGBLUP (1,00) obtiveram os melhores resultados com os animais genotipados, enquanto que para os animais não genotipados apenas o WssGBLUP (1,02) apresentou o melhor valor, embora também fosse maior que 1. O modelo PBLUP teve os piores resultados, com valores distantes 35% do valor de referência (1,0).

No conjunto de dados completo, a correlação de Pearson entre GEBVs e EBVs (preditos via PBLUP) variou entre 0,92 (ssBRR) e 0,93 (ssGBLUP), enquanto que para animais com informação de genótipo e fenótipo, a correlação entre GEBVs e EBVs variou entre 0,71 (ssBRR) e 0,77 (WssGBLUP). As correlações entre todos os GEBVs preditos utilizando os modelos genômicos de passo único foram superiores a 0,95.

Tabela 3. Estimativas de acurácia e dispersão dos animais com genótipo e sem genótipo obtidas por meio dos modelos de passo único e PBLUP

	Acurácia		Dispersão	
	Com genótipo	Sem genótipo	Com genótipo	Sem genótipo
ssBA	0,37	0,33	0,87	0,94
ssBB π	0,32	0,31	0,85	0,89
ssBC π	0,32	0,31	0,99	0,94
ssBL	0,31	0,31	1,02	0,94
ssBRR	0,30	0,31	0,96	0,90
ssGBLUP	0,33	0,29	0,85	0,97
WssGBLUP	0,39	0,31	1,00	1,02
PBLUP	0,30	0,29	0,65	0,69

ssBA: Single step Bayes A; ssBB π : Single step Bayes B π ; ssBC π : Single step Bayes C π ; ssBL: Single step Bayes Lasso; ssBRR: Single step Bayes Ridge Regression

3.4 Discussão

Existem poucos relatos na literatura sobre métodos de comparação e predição genômica para características reprodutivas, e ainda não é possível encontrar estudos que compararam a capacidade preditiva entre métodos de passo único para IPP. Além disso, várias medidas de acurácia foram relatadas por diferentes autores, o que torna ainda mais difícil a comparação entre seus valores. Por exemplo, Costa et al. (2019),

trabalhando apenas com animais da raça Nelore genotipados e considerando a correlação entre GEBV e IPP ajustada ($r_{GEBV,Y*}$) como medida de capacidade de predição, encontraram valores estimados com magnitudes semelhantes aos encontrados em nosso estudo com BayesC π (0,31) e Bayes-Lasso (0,31). Enquanto Toghiani et al. (2017) relataram valores menores utilizando BayesA (0,148) e BayesC π (0,15), estudando a mesma raça e usando o mesmo tipo de acurácia. Em ambos os estudos citados, as estimativas de acurácia não foram apresentadas para o modelo PBLUP, o que não nos permite comparar com nossos resultados. Possivelmente, para fins de comparação, verificar os incrementos de acurácia em relação ao PBLUP é a melhor abordagem para discutir nossos resultados. Assim, toda a literatura discutida sobre capacidade de predição foi feita com base na diferença de desempenho entre os métodos utilizados neste estudo.

Considerando a porcentagem do incremento de acurácia em relação ao PBLUP nos animais genotipados, observamos três grupos de modelos: incrementos menores que 5% - ssBL e ssBRR; incrementos próximos de 10% - ssBB π , ssBC π e ssGBLUP e; incrementos maiores que 25% - ssBA e WssGBLUP. Incrementos de acurácia pelo uso de marcadores SNP com magnitudes semelhantes às que encontramos para IPP têm sido descritos na literatura para características produtivas em búfalos. Para a produção de leite, gordura e proteína, maiores incrementos de acurácia ($r_{GEBV,Y*}$) são apresentados por Herrera et al., (2021) (13,04% a 76,47%) e Cesarani et al., (2021) (15,28 % para 33,33%) comparando ssBA e WssGBLUP com PBLUP.

Aspilcueta-Borquis et al. (2015), estudando a mesma população deste estudo, relataram um aumento na acurácia média (baseada na PEV) dos EBVs variando de 8,52 a 12,05% para várias características produtivas, quando as predições ssGBLUP e PBLUP foram comparadas. Este aumento na acurácia, observada com a utilização da informação genômica deve-se à captura adicional de variações de amostragem mendeliana e de parentesco entre animais (Cesarani et al., 2020; Christensen et al., 2012). A semelhança na capacidade da predição entre os diferentes métodos genômicos de passo único, contudo, poderia estar relacionada com o reduzido número de animais avaliados, o pequeno subconjunto de animais genotipados e a natureza poligênica da IPP (Calus, 2010; Campos et al., 2013).

Para animais não genotipados, como esperado, o aumento da acurácia através

da utilização dos métodos de passo único foi inferior ao incremento obtido para animais genotipados, destacando-se os modelos com magnitudes semelhantes e o modelo ssBA como o método de maior contribuição. Gao et al. (2018) estudando bovinos leiteiros vermelhos da Finlândia também não encontraram diferenças entre ssGBLUP e ssBayesC. Como discutido por Gao et al. (2018), com a hipótese de normalidade multivariada, a imputação explícita utilizada pelo método ssBR e a imputação implícita aplicada pelos modelos ssGBLUP e WssGBLUP são equivalentes quando a característica assume uma arquitetura poligênica, como a IPP. Isto pode explicar a acurácia semelhante para animais não genotipados obtida pelos modelos tradicionais de passo único e ssBR. Apesar do baixo incremento de acurácia observado para animais não genotipados, nas análises de passo único as relações entre os animais da população de referência e os animais da população de validação afetam a acurácia para os indivíduos não genotipados, uma vez que o GEBV destes indivíduos depende tanto dos efeitos marcadores como dos resíduos de imputação (Zhou et al., 2018).

Na análise de agrupamento hierárquico dos efeitos dos marcadores, dois grupos ficaram evidentes: um com abordagens Bayesianas com *shrinkage* diferencial; e outro com modelos baseados em GRM (WssGBLUP e ssGBLUP) e Bayesianos sem *shrinkage* diferencial (ssBA). A semelhança entre os modelos baseados em ssBA e GRM observada neste estudo (tanto pela análise de agrupamento como pela magnitude da acurácia estimada) pode ser explicada pela natureza poligênica da característica e pela robustez do método ssBA para diferentes arquiteturas genéticas. Zhou et al. (2018), estudando os efeitos do número QTLs na predição da acurácia com dados simulados, verificaram que o método ssBA era robusto e realizado de forma semelhante ao ssGBLUP em cenários com um grande número de QTLs.

O critério de dispersão está relacionado ao grau de inflação ou deflação das predições, e foi medido no conjunto de validação como a inclinação da regressão do GEBV obtido com o conjunto de dados completo sobre o GEBV obtido a partir do conjunto de dados de validação. Os resultados mostraram que a inclusão de informação genômica reduz a dispersão da predição em comparação com as avaliações tradicionais (PBLUP), independentemente do modelo de passo único considerado. A mesma superioridade na estatística através da utilização de

informação genômica foi relatada por Lázaro et al., (2021) estudando várias características relacionadas ao leite por meio de modelos de regressão aleatória na mesma população de búfalos, enquanto outros autores, tais como Gao et al., (2018) também analisando características do leite em vacas leiteiras, não relatam nenhum benefício.

Além disso, os GEBVs preditos com WssGBLUP não se dispersaram em relação ao que era esperado para o subconjunto de animais genotipados, e também apresentaram pouca deflação para a predição de animais não genotipados. A baixa dispersão dos GEBVs preditos com o WssGBLUP em comparação com as predições dos outros modelos genômicos deve-se provavelmente ao autoajuste dos pesos dos SNPs ao conjunto de dados (etapas de ponderação), que regulam a escala do efeito aditivo ao conjunto de dados completo anteriormente (*a priori*).

Em geral, alguns pontos devem ser considerados em relação aos métodos realizados neste estudo. Para os modelos de regressão bayesiana comuns, os efeitos do marcador podem não representar a variância genética aditiva total, e o uso de um efeito poligênico residual extra permite aumentar a acurácia e melhorar a dispersão dos GEBVs. Assim, combinar as duas informações, marcadores genéticos e pedigree, em uma análise de passo único é a melhor forma prática de aplicar os métodos Bayesianos para avaliação genética até agora. Além disso, os modelos sem seleção de variáveis, mas que ajustam a variância heterogênea para efeitos de marcadores tiveram o melhor desempenho para predição da IPP, além de aumentar a acurácia do GEBV para animais não genotipados. Esses achados são consistentes com a arquitetura poligênica da IPP, manifestada por sua baixa herdabilidade.

Por fim, vale ressaltar que modelos genômicos multi-características de passo único são importantes perspectivas para aplicação da metodologia WssGBLUP e ssBR em avaliações genéticas, devido à possibilidade de aumentar a acurácia dos GEBVs quando duas ou mais características são altamente correlacionadas (Silva et al., 2014). Atualmente, o WssGBLUP é aplicado apenas em análises unicaracterística ou em modelos com efeitos genéticos não correlacionados, devido à dificuldade de definir pesos únicos para construir a matriz H. Por outro lado, o método ssBR sob modelos multi-características já demonstrou aumentar a acurácia da predição. Portanto, outras investigações multi-características usando o método ssBA (o melhor

entre os modelos ssBR em nossa análise) devem ser realizadas.

3.5 Conclusão

Os resultados indicam que o uso de informações genômicas podem aumentar o ganho genético para idade ao primeiro parto, com maior incremento de acurácia e reduzindo a inflação/deflação das predições dessa característica em comparação com o modelo tradicional baseado em pedigree. Os modelos de passo único WssGBLUP e ssBA foram os métodos mais vantajosos para serem utilizados na avaliação genômica da idade ao primeiro parto de búfalas dessa população.

3.6 Referências

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ (2010) Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. **Journal of Dairy Science** 93:743–752. <https://doi.org/10.3168/jds.2009-2730>

Araujo Neto FR, Takada L, Santos DJA, Aspilcueta-Borquis RR, Cardoso DF, Nascimento AV, Leão KM, Oliveira HN, Tonhati H (2020a) Identification of genomic regions related to age at first calving and first calving interval in water buffalo using single-step GBLUP. **Reproduction in Domestic Animals** 55:1565–1572. <https://doi.org/10.1111/rda.13811>

Araujo Neto FR, Santos DJA, Fernandes Júnior GA, Aspilcueta-Borquis RR, Nascimento AV, de Oliveira Seno L, Tonhati H, de Oliveira HN. (2020b) Genome-wide association studies for growth traits in buffaloes using the single step genomic BLUP. **Journal of Applied Genetics** 61:113-115. <https://doi.org/10.1007/s13353-019-00528-5>

Aspilcueta-Borquis RR, Araujo Neto FR, Santos DJ, Hurtado-Lugo NA, Silva JA, Tonhati H (2015) Multiple-trait genomic evaluation for milk yield and milk quality traits using genomic and phenotypic data in buffalo in Brazil. **Genetics and Molecular Research** 14:18009-18017. <https://doi.org/10.4238/2015.December.22.27>

Aspilcueta-Borquis RR, Seno LO, Araujo Neto FR, Santos DJA, Hurtado-Lugo NA, Tonhati H (2022) Lifetime productivity: Genetic study of longevity and its associations with economically important traits in dairy buffaloes. **Livestock Science** 259. <https://doi.org/10.1016/j.livsci.2022.104900>

Calus MPL, Meuwissen THE, De Roos APW, Veerkamp RF (2008) Accuracy of genomic selection using different methods to define haplotypes. **Genetics** 178:553-561. <https://doi.org/10.1534/genetics.107.080838>

Calus MPL (2010) Genomic breeding value prediction: Methods and procedures. **Animal** 4:157–164. <https://doi.org/10.1017/S1751731109991352>

Camargo GMF, Aspilcueta-Borquis RR, Fortes MRS, Porto-Neto R, Cardoso DF, Santos DJA, Lehnert SA, Reverter A, Moore SS, Tonhati H (2015) Prospecting major genes in dairy buffaloes. **BMC Genomics** 16:872. <https://doi.org/10.1186/s12864-015-1986-2>

Campos G, Hickey JM, Pong-Wong R, Daetwyler HD, Calus MPL (2013) Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. **Genetics** 193:327–345. <https://doi.org/10.1534/genetics.112.143313>

Cesarani A, Garcia A, Hidalgo J, Degano L, Vicario D, Macciotta NPP, Lourenco D (2020) Genomic information allows for more accurate breeding values for milkability in dual-purpose Italian Simmental cattle. **Journal of Dairy Science** 104:5719–5727. <https://doi.org/10.3168/jds.2020-19838>

Cesarani A, Biffani S, Garcia A, Lourenco D, Bertolini G, Neglia G, Misztal I, Macciotta NPP (2021) Genomic investigation of milk production in Italian buffalo. **Italian Journal of Animal Science** 20:539–547. <https://doi.org/10.1080/1828051X.2021.1902404>

Christensen OF, Madsen P, Nielsen B, Ostersen T, Su G (2012) Single-step methods for genomic evaluation in pigs. **Animal** 6:1565–1571. <https://doi.org/10.1017/S1751731112000742>

Costa RB, Irano N, Diaz IDPS, Takada L, Hermisdorff IC, Carvalheiro R, Baldi F, Oliveira HN, Tonhati H, Albuquerque LG (2019) Prediction of genomic breeding values for reproductive traits in Nellore heifers. **Theriogenology** 125:12-17. <https://doi.org/10.1016/j.theriogenology.2018.10.014>

Fernando RL, Dekkers JC, Garrick DJ (2014) A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. **Genetics Selection Evolution** 46. <https://doi.org/10.1186/1297-9686-46-50>

Gao H, Koivula M, Jensen J, Strandén I, Madsen P, Pitkänen T, Aamand GP, Mäntysaari EA (2018) Short communication: Genomic prediction using different single-step methods in the Finnish red dairy cattle population. **Journal of Dairy Science** 101:10082-10088. <https://doi.org/10.3168/jds.2018-14913>

Habier D, Fernando RL, Kizilkaya K, Garrick DJ (2011) “Extension of the bayesian alphabet for genomic selection.” **BMC Bioinformatics** 12. <https://doi.org/10.1186/1471-2105-12-186>

Herrera JRV, Flores EB, Duijvesteijn N, Moghaddar N and van der Werf JH (2021) Accuracy of Genomic Prediction for Milk Production Traits in Philippine Dairy Buffaloes. **Frontiers in Genetics** 12. <https://doi.org/10.3389/fgene.2021.682576>

Iamartino D, Nicolazzi EL et al (2017) Design and validation of a 90K SNP genotyping assay for the water buffalo. **Plos One** 12. <https://doi.org/10.1371/journal.pone.0185220>

Lázaro SF, Tonhati H, Oliveira HR, Silva AA, Nascimento AV, Santos DJA, Stefani G, Brito LF (2021) Genomic studies of milk-related traits in water buffalo (*Bubalus bubalis*) based on single-step genomic best linear unbiased prediction and random regression models. **Journal of Dairy Science** 104:5768-5793. <https://doi.org/10.3168/jds.2020-19534>

Lee J, Cheng H, Garrick D, Golden B, Dekkers J, Park K, Lee D, Fernando R (2017) Comparison of alternative approaches to single-trait genomic prediction using genotyped and non-genotyped Hanwoo beef cattle. **Genetics Selection Evolution** 49:2. <https://doi.org/10.1186/s12711-016-0279-9>

Legarra A, Reverter A (2018) Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. **Genetics Selection Evolution** 50. <https://doi.org/10.1186/s12711-018-0426-6>

Meuwissen THE, Hayes BJ, Goddard ME (2001). "Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps." **Genetics**, 157:1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>

Misztal I, Tsuruta S, Lourenço D, Masuda Y, Aguilar I, Legarra A, Vitezica Z (2014) BLUPF90 family of programs. Disponível em: <https://nce.ads.uga.edu/wiki/doku.php>

Misztal I, Lourenco D, Legarra A (2020) Current status of genomic evaluation. **Journal of Animal Science** 98. <https://doi.org/10.1093/jas/skaa101>

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <https://www.R-project.org/>.

Silva MVB, Santos DJA, Boison AS, Utsunomiya ATH, Carmo AS, Sonstegard TS, Cole JB, Van Tassel CP (2014) The development of genomics applied to dairy breeding. **Livestock Science** 166:66-75. <https://doi.org/10.1016/j.livsci.2014.05.017>

Toghiani S, Hay E, Sumreddee P, Geary TW, Rekaya R, Roberts AJ (2017) Genomic prediction of continuous and binary fertility traits of females in a composite beef cattle breed. **Journal of Animal Science** 95:4787-4795. <https://doi.org/10.2527/jas2017.1944>

VanRaden PM (2008) Efficient methods to compute genomic predictions. **Journal of Dairy Science** 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>

VanRaden PM, van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS (2009) Invited review: Reliability of genomic predictions for North American Holstein bulls. **Journal of Dairy Science** 92:16–24. <https://doi.org/10.3168/jds.2008-1514>

Wang H, Misztal I, Aguilar I, Legarra A, Muir W (2012) Genome-wide association mapping including phenotypes from relatives without genotypes. **Genetics Research** 94:73–83. <https://doi.org/10.1017/S0016672312000274>

Yi N, Xu S (2008) “Bayesian LASSO for Quantitative Trait Loci Mapping.” **Genetics** 179:1045–1055. <https://doi.org/10.1534/genetics.107.085589>

Yin L, Zhang H, Li X, Zhao S, Liu X (2022) hibayes: An R Package to Fit Individual-Level, Summary-Level and Single-Step Bayesian Regression Models for Genomic Prediction and Genome-Wide Association Studies, **bioRxiv**. <https://doi.org/10.1101/2022.02.12.480230>

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** 7. <https://doi.org/10.3389/fgene.2016.00151>

Zhou L, Mrode R, Zhang S, Zhang Q, Li B, Liu JF (2018) Factors affecting GEBV accuracy with single-step Bayesian models. **Heredity** 120:100-109. <https://doi.org/10.1038/s41437-017-0010-9>