

NADSON OLIVEIRA DE SOUZA

**MÉTODOS SINGLE-STEP E DESENVOLVIMENTO DE PAINEL CUSTOMIZADO
PARA AVALIAÇÃO GENÉTICA DE BOVINOS DA RAÇA BRAFORD E HEREFORD**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2019

Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa

T

S729m
2019 Souza, Nadson Oliveira de, 1988-
Métodos single-step e desenvolvimento de painel
customizado para avaliação genética de bovinos da raça Braford
e Hereford / Nadson Oliveira de Souza. – Viçosa, MG, 2019.
xii, 64 f. : il. (algumas color.) ; 29 cm.

Orientador: Marcos Deon Vilela de Resende.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Bovinos de corte. 2. Animais - Melhoramento genético.
3. Bovinos de corte - Seleção. I. Universidade Federal de Viçosa.
Departamento de Zootecnia. Programa de Pós-Graduação em
Genética e Melhoramento. II. Título.


CDD 22. ed. 636.213

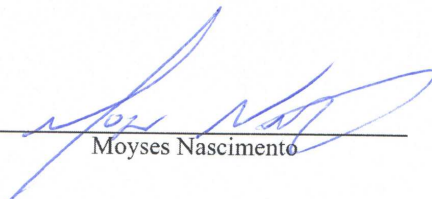
NADSON OLIVEIRA DE SOUZA


**MÉTODOS SINGLE-STEP E DESENVOLVIMENTO DE PAINEL
CUSTOMIZADO PARA AVALIAÇÃO GENÉTICA DE BOVINOS DA RAÇA
BRAFOR D E HEREFORD**


Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.


APROVADA: 26 de fevereiro de 2019.



Cláudio Napolís Costa


Moyes Nascimento


Camila Ferreira Azevedo


Fernando Flores Cardoso
(Coorientador)


Fabyano Fonseca e Silva
(Coorientador)


Marcos Deon Vilela de Resende
(Orientador)

“Educação não transforma o mundo. Educação muda as pessoas. Pessoas transformam o mundo.”

Paulo Freire

*Eu dedico esse trabalho para os meus pais
Olivio e Sirlei,*

“Tudo aquilo que sou, ou pretendo ser, devo a um anjo, minha mãe.”

Abraham Lincoln

AGRADECIMENTOS

A DEUS, pela vida, saúde e por todos os motivos que me levam a crer que sou a pessoa mais feliz e realizada desse mundo;

Aos meus pais, Sirlei Cedras Oliveira de Souza e Olívio Ferreira de Souza, exemplo de amor e caráter. Muito obrigado por toda educação e esforço para que hoje fosse possível. Pelo incentivo e compreensão que eu precisava percorrer meu caminho para alcançar todos os meus objetivos. Apesar de estarem longe, são eles o motivo de eu seguir na luta cotidiana;

Às minhas irmãs, Naiana Oliveira de Souza e Naiara Augusta Oliveira de Souza, pelo amor, carinho, pela torcida e estarem sempre presente ao longo de todo esse percurso;

Minha noiva e futura esposa, Gabriella Queiroz de Almeida, por sempre estar ao meu lado nos momentos bons e ruins me apoiando, ajudando e torcendo;

À Universidade Federal de Viçosa (UFV), especialmente ao programa de pós-graduação em genética e melhoramento pela oportunidade de realização do Doutorado;

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro;

Embrapa Pecuária Sul pelo apoio teórico e técnico e a Conexão Delta G por fornecer os dados e tornar esse trabalho possível;

Ao Dr. Marcos Deon Villela Resende, o qual tenho a honra em tê-lo como meu orientador, pela valiosa supervisão, ensinamento, paciência, amizade, oportunidades e pela ajuda quando eu mais precisei;

Ao Dr. Fernando Flores Cardoso, o qual tenho a honra em tê-lo como meu co-orientador, pela sua supervisão, ensinamento, oportunidades, ajudas e amizade;

Ao Dr. Fabyano Fonseca e Silva, o qual tenho a honra em tê-lo como meu co-orientador, pelo seu ensinamento, amizade, conselhos, oportunidades e ajudas;

À professora Dra. Camila Ferreira Azevedo, pela sua amizade, ensinamento e disponibilidade em participar da banca contribuindo para melhoria deste trabalho;

Ao professor Dr. Moisés Nascimento pela sua amizade, ensinamento e disponibilidade em participar da banca contribuindo para melhoria deste trabalho;

Ao Dr. Claudio Napolis Costa pelo ensinamento e disponibilidade em participar da banca contribuindo para melhoria deste trabalho;

Ao professor Dr. Robledo de Almeida Torres pela amizade e ensino;

À pesquisadora Dra. Bruna Pena Sollero, por sua ajuda, ensinamento e amizade;

Ao pesquisador Dr. Marcos Jun it Yokoo por sua ajuda, amizade e suporte;

À professora Dra. Daniela Andressa Lino Lourenco por me ter dado a oportunidade de trabalhar como bolsista de pesquisa na Universidade da Geórgia, ensinamento, dicas, conselhos e fornecer todo o apoio necessário para viver uma incrível experiência de vida no EUA e grupo;

Ao pesquisador Yutaka Masuda por toda amizade, dicas, paciência e ensinamento;

Ao pesquisador Ignacy Misztal e Shogo Tsuruta, por todas as conversas e ensinamento;

Ao prof. Dr. Leonardo L. Bhering por toda as conversas, ajudas, ensinamentos, oportunidades e amizade;

A todos os professores e colegas do Programa de Pós-Graduação em Genética e Melhoramento, em especial aos secretários Marco Túlio e Odilon Junior que sempre me ajudaram a resolver alguns problemas e sempre disponível e a todos que contribuíram de alguma forma para a execução deste trabalho e para minha formação acadêmica;

Aos professores, técnicos e funcionários do Departamento de Zootecnia – UFV;

Aos meus amigos da Biometria, Vinicius, Leonardo Peixoto, Lidiane, Sara, Janeo, Leonardo Correa, Larissa, Paulo, Andrei, Tassinari, Rodrigo, Igor, João, Rafael por bons momentos, piadas, agradável ambiente de trabalho e por sua amizade;

Aos meus amigos da Zootecnia Giovani, Edson, Hinayah, Felipe, Luiz Brito, Jeferson, Joashlley, Lais, Sirlene, Leonardo Glória, Lucas, Luciano, Matilde, Renata, Rogério, André Mauric, Delvan, Alessandra, Hugo, Geraldo, Yara, William, Barbara, Aline e Natalia por bons momentos, piadas, agradável ambiente de trabalho e por sua amizade;

Aos meus amigos de Bagé, Henry, Leonardo Lunardini, Ândrea, Gustavo, Gabriel, William, Rodrigo Azambuja, Rodrigo Costa, Bruno, Patrícia, Helena e Fábio por bons momentos, piadas, agradável ambiente de trabalho e por sua amizade;

Aos meus amigos do Estados Unidos – U.S. Breno, Ivan, André, Taylor, Yvette, Jorge, Ashley, Dennis, Pete, Prof. Rekaya, Andra, Susan, Profa. Fabiana, Prof. Fábio e Jefferson;

Ao professor Hao Cheng por ter me recebido por uma semana em seu laboratório na Universidade de Davis, Califórnia – U.S. e contribuir no desenvolvimento desta pesquisa;

Em especial ao Jeff Thompson e sua família que estavam presentes na minha vida durante o período em Athens-GA, U.S;

Para todos que direta ou indiretamente contribuíram para este trabalho.

BIOGRAFIA

NADSON OLIVEIRA DE SOUZA, filho de Olivio Ferreira de Souza e Sirlei Cedras Oliveira de Souza, nasceu em 27 de abril de 1988, na cidade de Jacobina, Bahia.

Em 2018, iniciou a graduação em Zootecnia na Universidade Federal do Vale do São Francisco (UNIVASF) na cidade de Petrolina - PE.

No agosto de 2010, transferiu-se para a Universidade Federal de Viçosa (UFV) em Viçosa-MG. Durante o curso de graduação na UFV se dedicou aos estudos e estágio na área de melhoramento animal, colando grau em fevereiro de 2013.

Em abril de 2013, iniciou o mestrado em Zootecnia na área de Melhoramento Animal pela Universidade Federal de Viçosa. Em fevereiro de 2015, submeteu-se aos exames finais de defesa de dissertação para obtenção do título de Magiste Scientiae em Zootecnia.

Em março de 2015, iniciou o doutorado em Genética e Melhoramento pela Universidade Federal de Viçosa, sendo que de maio de 2018 a outubro de 2018 foi pesquisador visitante na University of Georgia, nos Estados Unidos. Em fevereiro de 2019 submeteu-se a defesa.

SUMÁRIO

RESUMO	<i>xi</i>
ABSTRACT	<i>xiii</i>
INTRODUÇÃO GERAL	1
REFERÊNCIAS.....	5
CAPÍTULO 1.....	8
ACURÁCIA DE PREDIÇÃO PARA MÉTODOS SINGLE-STEP TESTANDO DIFERENTES PAINÉIS SNP PARA CARACTERÍSTICA DE RESISTÊNCIA AOS CARRAPATOS	8
Resumo.....	8
INTRODUÇÃO.....	9
MATERIAIS E MÉTODOS	10
RESULTADOS E DISCUSSÃO	17
CONCLUSÃO.....	29
AGRADECIMENTO.....	29
REFERÊNCIAS	29
CAPÍTULO 2.....	33
SELEÇÃO DE MARCADORES MOLECULARES UTILIZANDO ESTUDO DE ASSOCIAÇÃO GENÔMICA AMPLA BAYESIANA PARA CARACTERÍSTICA CONTAGEM DE CARRAPATOS EM BOVINOS DA RAÇA BRAFORD E HEREFORD	33
Resumo.....	33
INTRODUÇÃO.....	34
MATERIAL E MÉTODOS	35
RESULTADOS E DISCUSSÕES	41
CONCLUSÕES.....	59
AGRADECIMENTO.....	59
REFERÊNCIAS	59
CONCLUSÕES GERAIS.....	64

RESUMO

SOUZA, Nadson Oliveira, D.Sc., Universidade Federal de Viçosa, fevereiro de 2019. **Métodos single-step e desenvolvimento de painel customizado para avaliação genética de bovinos da raça Braford e Hereford.** Orientador: Marcos Deon Vilela de Resende. Coorientadores: Fernando Flores Cardoso e Fabyano Fonseca e Silva.

A utilização da informação genômica trouxe grandes oportunidades de aumento no ganho genético em gado de corte. Com a inclusão das informações dos marcadores SNP no processo de predição genômica, tornaram possível o desenvolvimento de métodos inovadores. O método single-step permite avaliar simultaneamente informações de fenótipos, pedigree e genômica, com a utilização de uma matriz combinada. O objetivo desse estudo foi comparar diferentes métodos single-step genômico em animais da raça Braford e Hereford para característica resistência aos carrapatos. Foram utilizados os métodos single-step GBLUP, single-step GBLUP ponderado (WssGBLUP) e o single-step de regressão bayesiana (SSBR). Os softwares da família BLUPF90 e o JWAS foram utilizados para prever os valores genéticos genômicos dos indivíduos para os métodos testados. O conjunto de dados foram divididos em três subconjuntos composto de animais Braford, Hereford e Braford_Hereford. Os resultados observados nesse estudo com o painel de 50K, sugerem uma diminuição nos valores da capacidade preditiva à medida que a raça Hereford é utilizada nas análises. Não houve diferença entre métodos single-step para o conjunto de dados Braford_Hereford. Para o conjunto de dados Braford, o método WssGBLUP foi superior e com o conjunto de dados Hereford, o WssGBLUP com duas e três iterações foram superiores. Com a informação molecular, estudos de associação genômica ampla (GWAS) permitem identificar loci ligados a determinada característica. O GWAS explora a existência de desequilíbrio de ligação entre SNPs e variantes causais, e assim permitindo selecionar SNPs mais informativos para customização de painéis de marcadores SNP de muito baixa densidade. O objetivo desse estudo foi identificar regiões genômicas e selecionar os SNPs mais representativos associados com a características resistência aos carrapatos. Adicionalmente, verificou a capacidade preditiva para os painéis customizados com base em animais da raça Braford e Hereford. Para isso, foram utilizados o método BayesB ($\pi = 0,99$) para a seleção dos marcadores altamente informativos para a característica e o

single-step de regressão bayesiana para prever os valores genéticos genômicos. A capacidade preditiva foi obtida pela correlação entre o fenótipo ajustado e o valor genético predito. Os softwares Gensel e o JWAS foram utilizados para a realização desse estudo. Os resultados observados nesse estudo sugerem a possibilidade de selecionar janelas superiores que explicam mais que 0,2% da variância genética para a característica. Os painéis customizados (tag SNP) para animais da raça Braford, Hereford e Braford_Hereford continham 35, 44 e 61 marcadores do tipo SNP que apresentaram maior efeito sobre a característica. O uso das tags SNP para animais representaria ao menos 67%, 72% e 28% da capacidade de predição obtidas pelo painel de 50K para animais Braford_Hereford, Braford e Hereford, respectivamente.

ABSTRACT

SOUZA, Nadson Oliveira, D.Sc., Federal University of Viçosa, February, 2019. **Single-step methods and custom panel development for genetic evaluation of Braford and Hereford beef cattle.** Advisor: Marcos Deon Vilela de Resende. Co-advisors: Fernando Flores Cardoso and Fabyano Fonseca e Silva.

The application of genomic information brought relevant growth opportunities of genetic gains in beef cattle. The inclusion of SNP markers information on the genomic prediction process, contributed to the development of innovative methods. The single-step method allows evaluate simultaneously the phenotypes, pedigree and genomic information, with the use of a combined matrix. The aim of this study was to compare different single-step genomic methods of animals of the breeds Bradford and Hereford for trait of ticks resistance. It were utilized the methods single-step GBLUP, single-step GBLUP weighted (WssGBLUP) and the single-step of bayesian regression (SSBR). The BLUPF90 family and JWAS software were used to predict the genetic genomic values of the individuals for the tested methods. The dataset was divided by three subsets composite of Bradford, Hereford and Bradford_Hereford animals. The results of this study with the 50K panel suggest a decrease on the predictive capacity values by the using of the Hereford breed. There was no difference among the single-step methods for the Braford_Hereford dataset. For the Braford dataset, the WssGBLUP method were top. With the molecular information, genomic wide association studies (GWAS) allow to identify loci linked to a determined trait. The GWAS explore the existence of linkage disequilibrium between SNPs and casual variants, allowing selecting more informative SNPs for customization of SNP markers panels of very low density. The aim of this study was to identify genomic regions and to select the most representative SNPs associated with the tick resistance trait. Additionally, it was verified the predictivity capacity for the customized panels based on Bradford and Hereford breeds. For this, it was utilized the BayesB method ($\pi = 0,99$) for the selection of highly informative markers for the trait and the Bayesian regression single-step to predict the genomic genetic values. The predictive capacity was obtained by the correlation between the adjusted phenotype and the genetic value predicted. The Gensel and JWAS softwares were utilized to the development of this study. The observed

results suggest the possibility to select the top windows that explain more than 0,2% of the genetic variance for the trait. The customized panels (tag SNP) for Braford, Hereford and Braford_Herdeford contained 35, 44 and 61 SNP markers, respectively, that present major effect on the trait. The use of tags SNP for animals represent at least 67%, 72% and 28% of the predictive capacity obtained by the 50K panel for Braford_Hereford, Braford and Hereford, respectively.

INTRODUÇÃO GERAL

A bovinocultura de corte brasileira apresenta papel de destaque no cenário mundial, uma vez que o Brasil é grande produtor e exportador de carne. Por consequência disso, a atividade contribui com grande parte do produto interno bruto (PIB) do país. O setor agropecuário brasileiro corresponde a 22,0% do PIB total, sendo que a pecuária é responsável por R\$ 400 bilhões, ou seja, 31,0% do PIB do agronegócio (ABIEC, 2017). Apesar da crise político-econômica que o país passou no ano de 2017, o PIB da pecuária cresceu 0,69% em relação ao ano anterior. O sucesso na produção é consequência de boa organização do setor, o que possibilita ações nas áreas de melhoramento, nutrição, reprodução e manejo sanitário. Outro ponto importante é a disponibilidade de recursos genéticos, que permite escolher raças que atenda as exigências do mercado.

Diante dos recursos genéticos existentes, as raças zebuínas (*Bos indicus*) ganham destaque por representar a maior porcentagem do rebanho nacional. A sua maior criação no país é devida as suas características como por exemplo adaptabilidade ao ambiente, ganho de peso e resistência a parasitas. Com o passar dos anos, o mercado passou a exigir animais mais eficientes em produção e com melhor qualidade de carcaça. Diante dessa necessidade, as raças taurinas (*Bos taurus*), como o Hereford, passaram a ganhar destaque devido ao seu potencial genético para maior qualidade de carcaça, que é definida pelo marmoreio. O marmoreio da carne representa a quantidade de gordura intramuscular, pois está intimamente relacionado com as características sensoriais da carne, possíveis de serem percebidas e apreciadas pelo consumidor (COSTA et al., 2002).

A raça Hereford se destaca por ser eficiente no ganho de peso e carne com bom marmoreio. A sua criação tem por objetivo a raça pura ou a formação do Braford, cruzamento entre Hereford x Zebu. A formação do Braford se deu pela necessidade de buscar animais que apresentasse melhor qualidade de carcaça e principalmente, animais com maior resistência a carrapatos. Isso é devido ao efeito de complementariedade entre duas raças diferentes, que permitiu unir características desejáveis e assim favorecer a sua expansão para outras regiões do Brasil. Estudos demonstraram que bovinos *Bos*

indicus são mais resistentes aos carrapatos do que os *Bos taurus* (Silva et al., 2010). A busca de animais mais resistentes aos carrapatos, é uma necessidade aparente no setor, devido aos problemas provocados ao sistema de produção.

A espécie de carrapatos *Rhipicephalus (Boophilus) microplus* é encontrado principalmente em países de clima tropical, como o Brasil. Esse parasita é responsável por transmitir doenças como a babesiose e anaplasmose, que juntas são comumente citadas como tristeza parasitária bovina. A tristeza parasitária bovina é um problema na pecuária bovina por provocar prejuízos econômicos, como aumento de mortalidade, redução na produção de carne, menor fertilidade e altos custos com tratamento. Esse parasita provoca perdas econômicas de milhões de dólares por ano no Brasil (Grisi, et al., 2002) e em outros países, como por exemplo a Austrália. O tratamento químico com acaricidas é um dos métodos mais utilizados no controle de carrapatos. No entanto, o uso indiscriminado de produtos químicos levou a seleção de parasitas mais resistentes a esses compostos químicos (Spagnol, Paranhos, Albuquerque 2010), o que conseqüentemente provoca a ocorrência de resíduos na carne. Com o objetivo de diminuir a aplicação de acaricidas, animais resistentes aos carrapatos vêm sendo utilizados na pecuária.

Os programas de melhoramento genético utilizam as avaliações genéticas com intuito de estudar e identificar animais resistentes aos carrapatos e assim promover ganhos genéticos pela seleção desses animais. Para identificação desses indivíduos superiores é ideal prever os valores genéticos (VG) de forma mais acurada. Desta forma, utilizar métodos genéticos estatísticos é necessário, principalmente as metodologias de modelos mistos, que incorporam informações de genealogia e de fenótipos do animal e de seus parentes para obtenção dos valores genéticos dos indivíduos via melhor preditor linear não viesado (Best Linear Unbiased Prediction – BLUP) (Henderson, 1975).

Com a utilização de marcadores de polimorfismo de um único nucleotídeo (Single Nucleotide Polymorphism – SNP) possibilitou o desenvolvimento de novos métodos para se estimar os efeitos destes marcadores. A seleção genômica (SG) permite prever valores genéticos genômicos (genomic estimate breeding values - GEBV) pelo uso da matriz de parentesco genômico (G) (VanRaden, 2008) ou pela estimação dos efeitos de

marcadores (Meuwissen et al., 2001). A seleção genômica baseia-se no fato de que os marcadores e o loci de características quantitativas (Quantitative Trait Loci – QTL) estejam em desequilíbrio de ligação (DL), isto é, associados, para garantir que os marcadores irão prever consistentemente os efeitos do QTL entre gerações e/ou populações (Hayes et al., 2009; Calus et al., 2010).

As informações dos marcadores moleculares além de contribuir para a obtenção dos valores genéticos mais acurados, podem auxiliar na identificação de SNPs associados as características de interesse e explorar as regiões do genoma em que esses marcadores estão localizados, buscando genes e/ou QTL relacionados à expressão fenotípica da característica (Utsunomiya et al., 2013). Este estudo é chamado de associação genômica ampla (genome-wide association study – GWAS).

A identificação dos marcadores mais significativos associados às características de interesse, pode levar ao delineamento de painéis de marcadores de baixa densidade explicando assim boa parte da variância genética dessas características, tornando-se soluções altamente desejáveis, com baixo custo e de extrema importância para a adoção da seleção genômica nos setores de produção animal (Van Eenennaam et al., 2014).

A utilização de painéis de marcadores SNPs, métodos foram desenvolvidos para que a informação genômica passasse a ser utilizado nas avaliações genéticas. Os métodos desenvolvidos até 2008, só levavam em consideração as informações fenotípicas de animais genotipados. Porém a partir de 2009, os métodos de Legarra et al. (2009) e Christensen e Lund (2010) permitiram o uso de todos os animais na avaliação genética, independente de possuírem genótipos e/ou fenótipos. Esse método que integra as informações genômicas, fenotípicas e de pedigree em um único passo é chamado BLUP genômico de único passo (single-step genomic BLUP – ssGBLUP).

O ssGBLUP combina o relacionamento baseado no parentesco tradicional (**A**) e na matriz de parentesco genômico (**G**), em uma única matriz chamada **H** (Misztal et al., 2009; Aguilar et al., 2010; Christensen & Mogens, 2010), utilizada para prever os valores genéticos genômicos. Segundo Wang et al. (2012), o ssGBLUP baseia-se no modelo infinitesimal, o qual assume igual variância para todos os efeitos dos SNPs, causando uma limitação no uso da metodologia, visto que esta não é a real situação para todas as características de interesse econômico. Estes autores propuseram o ssGBLUP

ponderado (WssGBLUP) que é o ssGBLUP, mas atribuindo pesos diferentes para os marcadores em um processo iterativo para atualizar as soluções dos SNPs. Entretanto, um desafio do ssGBLUP é o cálculo da matriz genômica inversa (\mathbf{G}^{-1}), que está em função do número de animais genotipados e exige alta demanda computacional para o cálculo da sua inversa.

Fernando et al. (2014) propuseram o método de regressão bayesiana de único passo (single-step Bayesian regression – SSBR), que não faz uso da \mathbf{G}^{-1} e da matriz de parentesco tradicional inversa dos animais genotipados (\mathbf{A}_{22}^{-1}). Além disso, o método permite assumir diferentes distribuições para efeito de marcadores, entre eles o BayesA e o BayesB. O BayesB assume, *a priori*, que o efeito dos marcadores SNPs é zero com probabilidade π , e apresentam uma distribuição normal com média zero e variância específica com probabilidade $(1 - \pi)$. A variância específica dos marcadores apresenta distribuição qui-quadrado invertida escalonada, com grau de liberdade e um parâmetro de escala, derivado de uma variância genética aditiva previamente assumida, sob pressuposição genética. Para BayesA, o valor de π é tratado como zero para que nenhum SNP tenha efeito nulo.

Os objetivos deste estudo foram comparar diferentes métodos single-step utilizando informações de animais da raça Braford e Hereford, identificar regiões genômica e marcadores SNP associados à resistência aos carrapatos e testar o desempenho preditivo do painel de baixa densidade baseado nos SNPs selecionados e comparar com o painel de SNP 50K.

REFERÊNCIAS

ABIEC. Relatório anual, perfil da pecuária no Brasil, 2017. <http://abiec.siteoficial.ws/images/upload/sumario-pt-010217.pdf> (Acessado em 04 de fevereiro de 2019).

Silva, A. M., Alencar, M. M., Regitano, L. C. A., Oliveira, M. C. S. Infestação natural de fêmeas bovinas de corte por ectoparasitas na Região Sudeste do Brasil. **Revista Brasileira de Zootecnia**, v. 39, p. 1477-1482, 2010.

Henderson, C.R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, v.31, p.423, 1975.

Fernando, R. L.; Dekkers, J. C. M.; Garrick, D. J. A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. **Genetics Selection Evolution**. 2014; 46:50.

Vanraden, P. M. Efficient methods to compute genomic predictions. **Journal of Dairy Science**, v. 91, p. 4414-4423, 2008.

Grisi, L. et al., Impacto econômico das principais ectoparasitoses em bovinos no Brasil. **A Hora Veterinária**, Por Alegre, v. 21, n. 125, p. 8-10, jan. 2002.

Spagnol, F. H., Paranhos, E. B., Albuquerque, G. R. Avaliação *in vitro* da ação sobre o *Rhipicephalus (Boophilus) microplus*, Canestrini, 1887 (Acari: ixodidae) de bovinos leiteiros no município de Itamaraju, Bahia, Brasil. **Ciência Animal Brasileira**. Goiânia, v. 11, n. 3, p. 731-736, 2010.

Meuwissen, T., Hayes, B. J., Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, p. 1819-1829, 2001.

Utsunomiya, Y. T., Do Carmo, A. S., Carvalheiro, R., Neves, H. H., Matos, M. C., Zavarez, L. B.; O'brien, A. M. P.; Sölkner, J.; Mcewan, J. C.; Cole, J. B.; Van Tassell, C. P.; Schenkel, F. S.; Da Silva, M. V. G. B.; Porto Neto, L. R.; Sonstegard, T. S.; José F Garcia, J. F. Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. **BMC genetics**, v. 14, n. 1, p. 52, 2013

Hayes, B.J.; Bowman, P.J.; Chamberlain, A.J.; Goddard, M. E. Invited review: Genomic selection in dairy cattle: Progress and challenges. **Journal of Dairy Science**, v. 92, n. 3, p. 433–443, 2009.

Calus, M. P. L. Genomic breeding value prediction: methods and procedures. **Animal**, v. 4, n. 2, p 157-164, 2010.

Costa, E.C.; Restle, J.; Brondani, I.L. et al. Composição física da carcaça, qualidade da carne e conteúdo de colesterol no músculo Longissimus dorsi de novilhos Red Angus superprecoces, terminados em confinamento e abatidos com diferentes pesos. **Revista Brasileira de Zootecnia**, v. 31, n. 1, p. 417-428, 2002 (supl.)

Van Eenennaam, A. L.; Weigel, K. A.; Young, A. E.; Cleveland, M. A.; Dekkers, J. C. Applied Animal Genomics: Results from the Field. *Annual Review of Animal Biosciences*, v. 2, n. 1, p. 105–139, 2014.

Misztal, I.; Legarra, A.; Aguilar, I. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. **Journal of Dairy Science**, v. 92, p. 4648–4655, 2009.

Aguilar, I.; Misztal, I.; Johnson, D. L.; Legarra, A.; Tsuruta, S.; Lawlor, T. J. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. **Journal of Dairy Science**, v. 93, p. 743–752, 2010.

Legarra, A.; Aguilar, I.; Misztal, I. A relationship matrix including full pedigree and genomic information. **Journal of Dairy Science**, v. 92, n. 9, p. 4656-4663, 2009.

Christensen, O. F.; Lund, M. S. Genomic prediction when some animals are not genotyped. **Genetics Selection Evolution**, v. 42, n. 1, p. 2, 2010.

Wang, H.; Misztal, I.; Aguilar, I.; Legarra, A.; Muir, W. Genome-wide association mapping including phenotypes from relative without genotypes. **Genetic Research**, v. 94, p. 73–83, 2012.

CAPÍTULO 1

ACURÁCIA DE PREDIÇÃO PARA MÉTODOS SINGLE-STEP TESTANDO DIFERENTES PAINÉIS SNP PARA CARACTERÍSTICA DE RESISTÊNCIA AOS CARRAPATOS

Resumo: O método genômico single-step BLUP (ssGBLUP) faz uso da matriz de relacionamento calculada a partir das informações de pedigree e marcadores. O método regressão bayesiana single step (SSBR) é uma extensão do ssGBLUP com abordagem bayesiana para estimação do efeito de marca, permitindo atribuir diferentes distribuições aos marcadores SNPs. O objetivo dessa pesquisa é comparar diferentes métodos single-step para resistência aos carrapatos em bovinos de corte da raça Braford (BO) e Hereford (HH). Três conjuntos de dados foram utilizados, sendo 9.414, 4.460 e 13.874 registros de contagem de carrapatos para BO, HH e BO_HH, respectivamente. Um total de 4.529 indivíduos foram genotipados. Modelo de repetibilidade foi utilizado para estimar os componentes de variância e parâmetros genéticos. Os métodos ssGBLUP, ssGBLUP ponderado (WssGBLUP) e SSBR foram comparados usando validação cruzada. A herdabilidade e repetibilidade para BO, HH e BO_HH foram 0,24 (0,30), 0,07 (0,21) e 0,18 (0,27), respectivamente. Os métodos que utilizam informações simultânea de pedigree, fenótipo e marcadores foram superiores ao BLUP baseado no pedigree, exceto para o conjunto de dados HH, em que o SSBR foi inferior e semelhante. Com o painel de 50K, os métodos single-step não foram diferentes, exceto para animais Hereford em que o SSBR foi inferior. O WssGBLUP apresentou um leve ganho nos valores de capacidade de predição entre os outros métodos single-step para os dois tipos de agrupamento. Utilizar informação genômica melhora os valores de capacidade de predição em relação ao método baseado somente em pedigree. A depender da população e raça utilizada, a capacidade de predição pode diminuir à medida que adiciona uma raça diferente em um banco de dados.

Palavras-chave: bovino de corte, melhoramento animal, seleção genômica, SSBR, ssGBLUP

INTRODUÇÃO

Os programas de melhoramento genético utilizam as avaliações genéticas para avaliar e identificar animais geneticamente superiores para características de interesse e assim promover ganhos genéticos pela seleção. O melhoramento animal tradicionalmente tem sido realizado com base na metodologia de modelos mistos, incorporando informações de genealogia e de fenótipo dos animais e de seus parentes, para obtenção dos valores genéticos dos indivíduos através do método BLUP (Best Linear Unbiased Prediction; Henderson, 1975).

Com o surgimento e a constante utilização da informação genômica no melhoramento animal, os painéis de diferentes densidades de marcadores passaram a ser utilizados nas avaliações genéticas, baseado em polimorfismo de nucleotídeo único (em inglês, single nucleotide polymorphism – SNP). A seleção genômica substitui a matriz de parentesco baseado no relacionamento médio, matriz **A**, pela matriz de parentesco realizada, matriz **G**. Com a seleção genômica, os animais selecionados passam a ter seu valor genético genômico predito pela estimação dos efeitos de marcadores SNPs, permitindo que animais sem fenótipo possam ter predição genômica sem ainda ter informação fenotípica, como proposto por Meuwissen et al. (2001).

Legarra et al. (2009) e Christensen e Lund (2010) propuseram a metodologia de melhor preditor linear não viesado genômico de único passo (em inglês, single-step genomic best linear unbiased prediction – ssGBLUP), que utiliza toda a informação de fenótipos, pedigree e genômica, simultaneamente, em uma avaliação genética. O ssGBLUP faz uso da matriz **H** que combina a matriz de relacionamento entre animais baseada no pedigree (**A**) e a diferença entre a matriz de relacionamento entre os animais baseado no pedigree e genômica (Legarra et al. 2009, Aguilar, et al. 2010). O ssGBLUP assume que todos os efeitos de SNPs tem variância comum. Este método comparado ao BLUP pedigree (PBLUP) e o método baseado somente em animais fenotipados e genotipados (BLUP genômico - GBLUP), apresentou ter acurácia similar ou superior (Cardoso et al., 2015). No entanto, o ssGBLUP não permite definir diferentes efeitos para todos os SNPs, visto que esta não é a real situação para todas as características de interesse econômico. Diante disso, Wang et al. (2012) propuseram o ssGBLUP

ponderado (em inglês, Weight ssGBLUP – WssGBLUP), em que diferentes pesos podem ser atribuídos para cada marcadores em um processo iterativo para atualizar as soluções dos SNPs. Com quantidade maior de animais dentro de uma avaliação genética, o método ssGBLUP e o WssGBLUP passou a ser ineficiente computacionalmente, devido a necessidade de obter as inversas das matrizes **G** e a **A₂₂**.

Fernando et al. (2014) propuseram o método de regressão bayesiana de único passo (em inglês, single-step Bayesian Regression - SSBR), que não precisa calcular a inversa da matriz **G** e a **A₂₂**. Além disso, este método permite assumir diferentes distribuições para efeito de marcador. O método BayesB é realizado com base em um modelo mistura com variância específica para cada loco, com uma fração de probabilidade π para efeitos zeros (Meuwissen et al. 2001) e quando $\pi = 0$, BayesB é semelhante ao BayesA. A presente pesquisa tem o objetivo comparar diferentes métodos single-step utilizando informações de animais da raça Braford e Hereford para resistência aos carrapatos.

MATERIAIS E MÉTODOS

A aprovação do Comitê de Cuidados e Uso de Animais não foi obtida para este estudo porque o conjunto de dados foi gerado a partir de vários estudos anteriores.

Dados Fenotípicos e Genotípicos

Os registros de contagem de carrapatos utilizados nas análises de diferentes rebanhos do programa de melhoramento da conexão Delta G (Rio Grande do Sul, Brasil). A conexão Delta G é uma associação que reúne um seletor grupo de agroempresas brasileiras, tendo como principal objetivo gerar e utilizar tecnologia de ponta para aumentar a rentabilidade da pecuária de corte.

Os registros fenotípicos foram obtidos a partir de 4.011 Braford e 1.856 Hereford. Dos 5.867 animais, 607 animais possuem somente 1 registro, 2.513 animais possuem 2 registros e os demais possuem 3 registros, totalizando 13.874 registros fenotípicos. Os

animais mantidos no arquivo de dados tinham entre 326 e 729 dias de idade no momento da contagem.

Os grupos contemporâneos (GC) foram formados pela combinação dos efeitos da fazenda, sexo, ano de nascimento, grupo de gerenciamento e data de contagem, com uma quantidade de 136 GC para Hereford e 171 para Braford. A característica avaliada nesse estudo foi contagem de carrapatos transformada por $\log_{10}(x_i + 1.0)$, onde x_i é a quantidade de carrapatos contabilizados na lateral esquerda do animal.

Como a população é composta por duas raças (Braford e Hereford), foram gerados três bancos de dados para realização deste estudo. O primeiro banco de dados continha apenas animais da raça Braford (BO) com 9.414 fenótipos, o segundo conjunto de dados continha apenas animais da raça Hereford (HH) com 4.460 registros e o terceiro conjunto de dados continha animais das raças Braford e Hereford (BO_HH) com 13.874 registros. O agrupamento da raça Braford e Hereford no terceiro conjunto de dados se deve a fato do objetivo comum de melhoramento pelo Programa de Melhoramento da Conexão Delta G e devido a similaridade genética das duas raças, já que os Brafords são compostos com uma contribuição de 62,5% da raça Hereford (Piccoli et al., 2014; Biegelmeier et al., 2016).

Um total de 4.529 indivíduos das duas raças foram genotipados em painéis de diferentes densidades. Sendo que 230 reprodutores foram genotipados no painel de alta densidade (BovineHD – Illumina® bread chip – 777.962 SNPs) e 162, 3.728, 287, 122 animais genotipados nos painéis de SNPs de 150K, 50K, 30K e 27K, respectivamente. Diferentes critérios de controle de qualidade foram utilizados para SNPs e amostras. Os critérios adotados para exclusão de SNPs foram alelo menos frequentes (em inglês, Minor Allele Frequency - MAF) (< 3%), taxa de genotipagem (Call rate - CR) (< 98%) e Equilíbrio de Hardy-Weinberg ($P > 10^{-7}$). Para amostra, os critérios utilizados para a avaliação da exclusão foram CR (< 90%), desvio de heterozigose acima de três desvios padrões, genótipos idênticos (> 99,5%) e dados duplicados. Todos os painéis foram filtrados para selecionar apenas SNPs em comum com o painel de 50K. Após a edição, um total de 39.922 marcadores SNP e 4.493 animais permaneceram para as análises. Dados perdidos em genótipos foram imputados segundo a metodologia *Sliding Windows* pelo programa FImpute (Sargolzaei et al., 2011).

Modelos estatísticos

BLUP baseado no pedigree (PBLUP)

Nesta análise, o seguinte modelo de repetibilidade baseado no pedigree foi utilizado para obter o valor genético para contagem de carrapatos:

$$y = X\beta + Za + Wpe + e \quad [1]$$

em que, y é o vetor de observações fenotípicas transformados para os animais com contagem; X , Z e W são matrizes de incidência para os efeitos fixos, genético e de ambiente permanente, respectivamente; β são as médias geral, efeitos de grupo contemporâneo (fixo), proporção racial (covariável), heterozigosidade (covariável), perda por recombinação (covariável), idade no dia da contagem linear e quadrático (covariável); a é o efeito genético aditivo, assumindo que $a \sim N(0, A\sigma_a^2)$, em que A é a matriz de numeradores dos coeficientes de parentesco e σ_a^2 é a variância genética aditiva; pe é o efeito de ambiente permanente, assumindo que $pe \sim N(0, I\sigma_{pe}^2)$, em que I é a matriz identidade e σ_{pe}^2 é a variância de ambiente permanente; e e é o termo residual, assumindo que $e \sim N(0, I\sigma_e^2)$, em que σ_e^2 é a variância residual.

Método single-step GBLUP (ssGBLUP) e o ssGBLUP ponderado (WssGBLUP)

Nesta análise, o seguinte modelo de repetibilidade também foi utilizado para obter o valor genético genômico para contagem de carrapatos:

$$y = X\beta + Za + Wpe + e$$

onde y é o vetor de observações fenotípicas transformados para ambos os animais genotipados e não genotipados; X , Z e W são matrizes de incidência para os efeitos fixos, genético e de ambiente permanente, respectivamente; β são a média geral, efeitos de grupo contemporâneo (fixo), proporção racial (covariável), heterozigosidade (covariável), perda por recombinação (covariável), idade no dia da contagem linear e quadrático (covariável); a é o efeito genético aditivo, assumindo que $a \sim N(0, H\sigma_a^2)$, em que H é a matriz de parentesco baseada tanto na informação genômica quanto na

informação de pedigree e o σ_a^2 é a variância genética aditiva; pe é o efeito de ambiente permanente, assumindo que $pe \sim N(0, I\sigma_{pe}^2)$, em que I é a matriz identidade e σ_{pe}^2 é a variância de ambiente permanente; e e é o termo residual, assumindo que $e \sim N(0, I\sigma_e^2)$, em que σ_e^2 é a variância residual. A matriz H possui a seguinte inversa (Aguilar et al., 2010):

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

Em que, A é a matriz do numerador do coeficiente de parentesco baseado no pedigree para todos os animais; A_{22} é a matriz do numerador do coeficiente de parentesco baseada no pedigree somente para os animais genotipados; e G é a matriz de parentesco genômica para os animais genotipados. Para construção da matriz G foi utilizado a metodologia proposta por VanRaden (2008) que pode ser escrito da seguinte forma:

$$G = \frac{ZZ'}{2 \sum p_i(1 - p_i)}$$

em que Z é a matriz de marcadores com dimensão que corresponde ao número de animais (n) e de marcadores (m); p_i representa a frequência do alelo no locus i .

Para WssGBLUP as soluções dos efeitos dos SNPs (\hat{u}) foram obtidos a partir dos valores genéticos, seguindo o seguinte algoritmo (Wang et al., 2012 – cenário S1):

- 1- Na primeira iteração $D_{(t)} = I$; $G_{(t)} = Z D_{(t)} Z' \lambda$, em que t é o número da iteração; D é a matriz diagonal de peso dos marcadores SNPs; λ é uma razão de variâncias ou uma constante de normalização (VanRaden, 2008)

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

em que σ_u^2 é a variância genética capturada por cada SNP, σ_a^2 é a variância genética aditiva e M é o número de marcadores SNPs.

- 2- Calcula o valor genético de todos os animais utilizando o ssGBLUP;
- 3- Calcula $\hat{u}_{(t)} = \lambda D_{(t)} Z' G_{(t)}^{-1} \hat{a}_g$; em que \hat{a}_g é o valor genético estimado dos animais genotipados;
- 4- Calcula os pesos $d_{i(t+1)} = \hat{u}_{i(t)}^2 2p_i(1 - p_i)$ para cada SNP (Zhang et al., 2010);
- 5- Normaliza os pesos para a variância genética aditiva permanecer constante

$$D_{(t+1)} = \frac{tr(D_{(o)})}{tr(D_{(t+1)})} D_{(t+1)};$$
- 6- Calcula $G_{(t+1)} = ZD_{(t+1)}Z' \lambda$;
- 7- Finalização da iteração ou volta para o passo 3 para iniciar nova iteração.

Assim, a matriz G foi recalculada para a predição dos efeitos dos SNPs a cada iteração, enquanto o \hat{a}_g foi obtido uma única vez, não sendo alterado durante as iterações. Foram testadas seis iterações para o ssGBLUP ponderado (WssGBLUP). O número máximo de iterações escolhido para o WssGBLUP do presente estudo foram 4, 3 e 3 para o conjunto de dados Braford, Hereford e Braford_Hereford, de acordo com a Figura 1. A escolha do número máximo de iteração foi com base no maior valor da capacidade preditiva obtido (Figura 1).

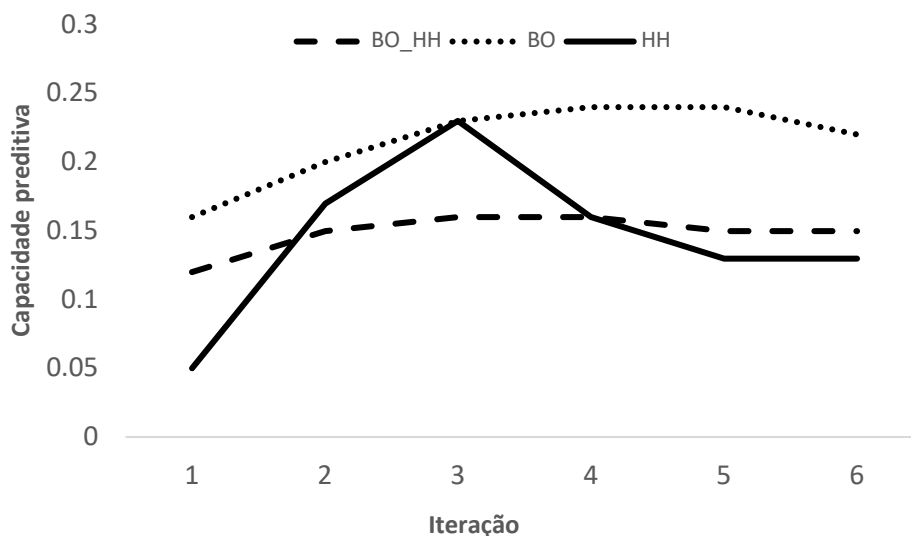


Figura 1. Capacidade preditiva para o método WssGBLUP para as seis primeiras iterações para os conjuntos de dados Braford, Hereford e Braford_Hereford.

Método single-step regressão bayesiana (SSBR)

Todos os valores genéticos genômico (GEBV) também foram obtidos pela regressão bayesiana de único passo (Fernando et al., 2014), usando *a priori* BayesB para efeitos de marcadores com $\pi = 0,99$, para cada ciclo de Monte Carlos via Cadeia de Markov (MCMC), com cadeia de 100.000 amostras e 10.000 de burn-in. Essas variâncias locus-específicas apresentam distribuição qui-quadrado invertida escalonada, com graus de liberdade e com um parâmetro de escala derivado de uma variância genética aditiva previamente assumida, sob pressuposição genética. Os indivíduos foram separados em dois grupos, designado com subscrito “1” para animais não genotipados e “2” para animais genotipados. O modelo de regressão bayesiana de único passo pode ser representado como:

$$\mathbf{y} = \mathbf{X}^* \boldsymbol{\beta} + \mathbf{Z}^* \mathbf{a} + \mathbf{W} \mathbf{pe} + \mathbf{U} \boldsymbol{\varepsilon} + \mathbf{e}$$

onde \mathbf{y} é o vetor de fenótipo para ambos indivíduos genotipados e não genotipados; $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$, $\mathbf{X}^* = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$, $\mathbf{Z}^* = \begin{bmatrix} \mathbf{Z}_1 & \hat{\mathbf{M}}_1 \\ \mathbf{Z}_2 & \mathbf{M}_2 \end{bmatrix}$, $\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix}$, $\mathbf{U} = \begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{0} \end{bmatrix}$, $\mathbf{e} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix}$. O $\boldsymbol{\beta}$ é a média geral e vetor de efeitos fixos especificados anteriormente, \mathbf{a} é o vetor aleatória de efeitos de substituição dos marcadores, com $a | \sigma_a^2 \sim N(0, I \sigma_a^2)$; \mathbf{pe} é o vetor aleatório de ambiente permanente, com $\mathbf{pe} | \sigma_{pe}^2 \sim N(0, I \sigma_{pe}^2)$; $\boldsymbol{\varepsilon}$ é o resíduo de imputação, *a priori* para $\boldsymbol{\varepsilon}$ é $\boldsymbol{\varepsilon} | \sigma_a^2 \sim N(0, (A_{11} - A_{12} A_{22}^{-1} A_{12}) \sigma_a^2)$ com $(\sigma_a^2 | v_a, S_a^2) \sim v_a S_a^2 X_{v_a}^2$; \mathbf{Z}_1 e \mathbf{Z}_2 são matrizes de incidência relativo ao valor genético dos indivíduos não genotipados e genotipados para os seus fenótipos e; \mathbf{e}_1 e \mathbf{e}_2 são vetores para o termo residual aleatório para indivíduos não genotipados e genotipados, respectivamente, *a priori* para e_i é $e_i | \sigma_e^2 \sim iid N(0, \sigma_e^2)$ com $(\sigma_e^2 | v_e, S_e^2) \sim v_e S_e^2 X_{v_e}^2$. A matriz $\hat{\mathbf{M}}_1$ representa a matriz de covariáveis SNP imputados para animais não genotipados que são derivados de parentes genotipados e que foi obtida da seguinte forma: $\mathbf{A}^{11} \hat{\mathbf{M}}_1 = -\mathbf{A}^{12} \mathbf{M}_2$ em que \mathbf{M}_2 é matriz de covariáveis SNP para os animais genotipados. $\hat{\mathbf{g}} = \begin{bmatrix} \hat{\mathbf{M}}_1 \\ \mathbf{M}_2 \end{bmatrix} \hat{\mathbf{a}} + \begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{0} \end{bmatrix} \hat{\boldsymbol{\varepsilon}}$.

Herdabilidade, repetibilidade e acurácia

A herdabilidade (h^2) e a repetibilidade (r^2) foram calculadas da seguinte forma:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}; \quad r = \frac{\sigma_a^2 + \sigma_{pe}^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}$$

em que σ_a^2 é a variância genética aditiva; σ_{pe}^2 é a variância de ambiente permanente e σ_e^2 é a variância residual.

Para verificar a eficácia do uso da informação genômica na predição do valor genético dos animais genotipados sem registros fenotípicos, a acurácia (Acc) foi obtida pelos modelos PBLUP e o ssGBLUP da seguinte forma:

$$Acc = \sqrt{1 - \frac{PEV}{\sigma_a^2}}$$

Validação cruzada e acurácia da predição

Dos 4.040 indivíduos genotipados, foram formados dois subgrupos considerando somente animais com fenótipos e genótipos. Para validação cruzada 2.978 animais Braford e 1.062 animais Hereford, foram divididos em grupos por duas estratégias usando o R (R Core Team, 2013). No primeiro cenário, o método de agrupamento K-médias, que utiliza a matriz de distância calculada por alelos idênticos por estado seguindo a proposta de Saatchi et al. (2011), foi usada para dividir em grupos os animais dos três conjuntos de dados (BO, HH, BO_HH). Já no segundo cenário, o método de agrupamento aleatório foi utilizado para os três arquivos de dados, preservando o mesmo número de grupos definido pelo K-médias. Os mesmos grupos formados para validação cruzada foi aplicado para os métodos PBLUP, ssGBLUP, WssGBLUP e o SSRB.

A capacidade preditiva, dentro de cada grupo da validação cruzada, foi calculada como a correlação entre os fenótipos transformados corrigido para os efeitos fixo (a^*) e

o valores genéticos preditos (\hat{a}) para cada conjunto de dados. O viés de predição foi avaliado pela regressão simples do fenótipo transformados corrigido e o valor genético predito.

Para os métodos PBLUP, ssGBLUP e WssGBLUP foi utilizado o software da família BLUPF90 (Miszta, 2000, 2008) e para o método SSRB foi utilizado JWAS: pacote Julia para análise de seleção genômica (Cheng et al., 2014).

RESULTADOS E DISCUSSÃO

Em programas de avaliação genética é de fundamental importância sabermos se a população apresenta variabilidade genética aditiva para a característica e uma forma de conhecermos é estimando a herdabilidade. Quando a avaliação envolve a utilização de fenótipos com medidas repetidas, o parâmetro de repetibilidade também se torna importante dentro de uma avaliação. Valores altos para estimativa da repetibilidade indicam que é possível prever o valor real do indivíduo com um número relativamente pequeno de medições (Cornacchia et al, 1995) e que pouco ganho em acurácia haverá com o aumento do número de medidas repetidas (Falconer, 1987).

As estimativas de componentes de variância, herdabilidade e repetibilidade para contagem de carrapatos obtido pelo modelo de repetibilidade [1] usando o pedigree para o modelo PBLUP via análise bayesiana, estão apresentados na Tabela 1, 2 e 3. No presente estudo, os resultados de herdabilidade mostram que existe variabilidade genética para a característica resistência aos carrapatos, no entanto, observa-se que ao incluir a raça Hereford no conjunto de dados de Braford, a herdabilidade e repetibilidade diminuíram. Esse comportamento é explicado pelo baixo valor de variância genética (0,006) que a raça Hereford apresenta, além disso, a herdabilidade estimada também foi de baixa magnitude (0,07).

Tabela 1: Média a posteriori (PMédia), desvio padrão a posteriori (PSD), e os percentis (%) de 2,5 e 97,5 para os componentes de variância e parâmetros genéticos da característica para o conjunto de dados contendo apenas animais da raça Braford.

Parâmetros	Braford			
	PMédia	PSD	2.5%	97.5%
Variância genética aditiva (σ_a^2)	0,023	0,003	0,002	0,029
Variância de ambiente permanente (σ_{pe}^2)	0,008	0,003	0,002	0,014
Variância residual (σ_e^2)	0,072	0,001	0,069	0,075
Herdabilidade (h^2)	0,23	0,03	0,01	0,09
Repetibilidade (r^2)	0,30	0,02	0,17	0,25

Tabela 2: Média a posteriori (PMédia), desvio padrão a posteriori (PSD), e os percentis (%) de 2,5 e 97,5 para os componentes de variância e parâmetros genéticos da característica para conjunto de dados contendo apenas animais da raça Hereford.

Parâmetros	Hereford			
	PMean	PSD	2.5%	97.5%
Variância genética aditiva (σ_a^2)	0,006	0,002	0,001	0,008
Variância de ambiente permanente (σ_{pe}^2)	0,013	0,002	0,010	0,020
Variância residual (σ_e^2)	0,069	0,002	0,066	0,074
Herdabilidade (h^2)	0,07	0,02	0,01	0,09
Repetibilidade (r^2)	0,21	0,02	0,17	0,25

Os valores médios para a característica contagem de carrapatos foram 1,33, 1,49 e 1,38 para o conjunto de dados Braford, Hereford e Braford_Hereford. A herdabilidade e repetibilidade, utilizando fenótipos somente da raça Braford, foram 0,23 e 0,30, respectivamente (Tabela 1), e para a raça Hereford foram 0,07 e 0,21, respectivamente (Tabela 2). Para o conjunto Braford_Hereford, as estimativas de herdabilidade e repetibilidade foram 0,18 e 0,27, respectivamente (Tabela 3). Para a mesma característica e utilizando somente o conjunto de dados Braford_Hereford (mesma

população do presente estudo), Cardoso et al. (2015) encontraram valores similares para os parâmetros genéticos avaliados na presente pesquisa

Independente do conjunto de dados avaliados, a variância residual apresentou valores de 0,07. A contribuição genética sobre a característica para animais HH é muito baixa, ou seja, os efeitos não genéticos influenciam a característica. Animais Braford apresentaram maior valor de variância genética e menor valor de ambiente permanente em relação ao HH, conseqüentemente, podemos observar que ao realizar o cruzamento das duas raças (Hereford e Zebu) o efeito genético sobre a característica aumenta.

Tabela 3: Média a posteriori (PMédia), desvio padrão a posteriori (PSD), e os percentis (%) de 2,5 e 97,5 para os componentes de variância e parâmetros genéticos da característica para o conjunto de dados contendo animais Braford_Hereford.

Parâmetros	Braford_Hereford			
	PMean	PSD	2.5%	97.5%
Variância genética aditiva (σ_a^2)	0,018	0,003	0,012	0,023
Variância de ambiente permanente (σ_{pe}^2)	0,009	0,003	0,005	0,015
Variância residual (σ_e^2)	0,072	0,001	0,069	0,074
Herdabilidade (h^2)	0,18	0,03	0,12	0,23
Repetibilidade (r^2)	0,27	0,01	0,25	0,30

Agrupamento k-médias e aleatório

A seleção genômica em programas de melhoramento genético permite prever os valores genéticos genômicos dos indivíduos e para isso, duas populações devem ser definidas, população de treinamento e de validação. Para a formação desses subconjuntos alguns procedimentos para os agrupamentos são utilizados, como o agrupamento k-médias e aleatório. A estratégia de agrupamento k-médias utiliza o nível de parentesco dos indivíduos para formação dos grupos e que pode usar a matriz de relacionamento com base no pedigree (Saatchi et al., 2011, 2013) e matriz de relacionamento genômica (Bodhireddy et al., 2014).

O número de indivíduos, coeficiente de endogamia e o relacionamento entre e dentro de grupos pelo agrupamento k-médias para as raças Braford, Hereford e Braford_Hereford estão na Tabela 4, 5 e 6, respectivamente. Com base no relacionamento genômico o agrupamento k-médias particionou em 4, 3 e 5 grupos, com tamanhos desiguais para animais BO, HH e BO_HH, respectivamente.

Tabela 4: Número de indivíduos (N), coeficiente de endogamia¹, desvio padrão (\pm SD) e relacionamento genético genômico (G_{ij}) entre e dentro de grupos pelo agrupamento k-média para conjunto de dados contendo apenas animais da raça Braford.

Grupos	1	2	3	4
N	913	453	1293	319
Coeficiente de endogamia¹	1,01 \pm 0,04	1,01 \pm 0,04	0,99 \pm 0,04	0,99 \pm 0,04
G_{ij} dentro do grupo	0,05 \pm 0,04	0,01 \pm 0,04	0,01 \pm 0,03	0,11 \pm 0,07
G_{ij} entre grupos	-0,01 \pm 0,05	0,00 \pm 0,04	0,00 \pm 0,03	-0,01 \pm 0,04

¹ Elementos da diagonal da matriz de relacionamento menos 1 ($G_{ii} - 1$)

Independente do banco de dados avaliado, os coeficientes de endogamia foram próximos de zero, ou seja, diagonal próxima de 1,00 (Tabela 4, 5 e 6). A média dos valores de G_{ij} dentro de grupos foram maiores que os valores de G_{ij} entre grupos, o que era esperado, já que o agrupamento k-médias particiona indivíduos relacionados em grupos e conseqüentemente diminui o relacionamento entre grupos (Saatchi et al. 2011). É possível observar na Tabela 4 que o grupo 4 foi o que obteve o maior valor de G_{ij} dentro de grupo (0,11), e este é o grupo com menor número de animais.

Observando a proporção de Hereford (Tabela 6), o método k-médias dividiu os animais genotipados em dois subgrupos: o primeiro formado pelo grupo 1 e 3, que apresenta 86% e 94,05% de Hereford e o segundo formado pelos demais grupos, apresentando um total de 4,29% de proporção de animais da raça Hereford. No segundo subgrupo, o grupo 2 foi o que apresentou o maior (0,05) e menor (-0,02) valor de G_{ij} dentro e entre grupos, respectivamente.

Tabela 5: Número de indivíduos (N), coeficiente de endogamia¹, desvio padrão (\pm SD) e relacionamento genético genômico (G_{ij}) entre e dentro de grupos pelo agrupamento k-média para o conjunto de dados contendo apenas animais da raça Hereford.

Grupos	1	2	3
N	333	85	644
Coeficiente de endogamia¹	1,03 \pm 0,05	1,04 \pm 0,03	1,04 \pm 0,05
G_{ij} dentro do grupo	0,08 \pm 0,05	0,34 \pm 0,03	0,11 \pm 0,04
G_{ij} entre grupos	0,01 \pm 0,06	0,00 \pm 0,07	-0,02 \pm 0,06

¹ Elementos da diagonal da matriz de relacionamento menos 1 ($G_{ii} - 1$)

Na Tabela 6, o grupo 3 se destaca por apresentar a maior proporção de bovinos Hereford (94,72%) e maior coeficiente de endogamia (0,05) em relação aos demais. No grupo 3, os valores de relacionamento genético dentro e entre grupos foram 0,11 e -0,02, respectivamente.

Tabela 6: Número de indivíduos (N), coeficiente de endogamia¹, proporção de Hereford, desvio padrão (\pm SD) e relacionamento genético genômico (G_{ij}) entre e dentro de grupos pelo agrupamento k-médias para conjunto de dados contendo as duas raças Braford_Hereford.

Grupos	1	2	3	4	5
N	497	1356	656	962	569
Proporção de Hereford	86,12	0,59	94,05	2,29	1,41
Coeficiente de endogamia¹	1,03 \pm 0,05	1,01 \pm 0,04	1,05 \pm 0,05	0,98 \pm 0,04	1,00 \pm 0,05
G_{ij} dentro de grupo	0,08 \pm 0,04	0,05 \pm 0,04	0,11 \pm 0,05	0,01 \pm 0,03	0,02 \pm 0,04
G_{ij} entre grupos	-0,01 \pm 0,06	-0,02 \pm 0,06	-0,02 \pm 0,06	0,00 \pm 0,03	0,00 \pm 0,03

¹ Elementos da diagonal da matriz de relacionamento menos 1 ($G_{ii} - 1$)

Utilizando o agrupamento aleatório, os grupos foram divididos de forma equilibrada para Braford com 745 (grupo 1 e 2) e 744 (grupos 3 e 4) indivíduos, para Hereford todos os três grupos continham 354 animais e para Braford_Hereford com 808 animais para todos os cinco grupos. Não houve diferenças nos valores médios de G_{ij}

entre e dentro de grupos para os conjuntos de dados (resultados não apresentados no presente estudo).

Capacidade preditiva via pedigree e genômica

Os valores de acurácia mínimo, máximo, médio e desvio padrão dos animais genotipados sem fenótipos para os métodos PBLUP e ssGBLUP para o conjunto de dados Braford, Hereford e Braford_Hereford estão apresentados na Tabela 7. A acurácia média para o método ssGBLUP teve um aumento de 53,44%, 55% e 58% em relação ao PBLUP para o conjunto de dados BO, HH e BO_HH, respectivamente, ou seja, o uso da matriz de relacionamento genômica foi superior ao método baseado no pedigree. Com a inclusão da raça Hereford ao banco de dados de Braford, teve uma diminuição de 0,02 na média da acurácia tanto para os métodos PBLUP e ssGBLUP. Hayes et al. (2009) demonstraram que ao substituir a matriz de relacionamento média derivada do pedigree pela matriz de relacionamento genômica, a acurácia de predição dos valores genéticos foi superior. Villanueva et al. (2005) também encontraram resultados semelhantes.

Tabela 7: Valor de acurácia mínimo (min), máximo (max), médio e desvio padrão dos animais genotipados sem fenótipo para os métodos PBLUP e ssGBLUP para o conjunto de dados BO, HH e BO_HH

	Acurácia PBLUP				Acurácia ssGBLUP			
	Min	Max	Média	SD	Min	Max	Média	SD
BO	0,01	0,95	0,31	± 0,25	0,05	0,95	0,52	± 0,13
HH	0,01	0,73	0,21	± 0,20	0,09	0,74	0,38	± 0,12
BO_HH	0,01	0,94	0,29	± 0,25	0,15	0,94	0,50	± 0,13

Métodos genômico single-step versus o PBLUP pela validação cruzada

Utilizando o painel de 50K, os métodos single-step não apresentaram diferença em relação ao PBLUP independente do agrupamento utilizado para o conjunto de dados Braford_Hereford (Tabela 8). Para o conjunto de dados contendo somente animais Braford, apenas o método WssGBLUP_3 e o WssGBLUP_4 foram diferentes do PBLUP para k-médias, já para o agrupamento aleatório todos os métodos genômicos foram diferentes do PBLUP (Tabela 8). Para o conjunto de dados contendo animais Hereford, somente o método WssGBLUP foram superiores ao PBLUP para os dois tipos de agrupamento (Tabela 8). Cardoso et al. (2015) trabalhando com a mesma população em uma análise considerando as duas raças juntas (BO_HH), verificaram que os métodos genômicos foram similares e superiores ao PBLUP. Lourenco et al. (2015) testaram a viabilidade do método ssGBLUP em bovinos da raça Angus, verificaram que a depender da característica, o uso da informação genômica na avaliação genética levou um aumento de 0,01 a 0,16 na capacidade preditiva.

O viés da regressão (β) para o painel 50K variou de 0,44 a 1,84 para k-médias e 0,26 a 1,43 para aleatório para os três conjuntos de dados. Entre métodos, o viés do PBLUP para o painel 50K independente dos agrupamentos, foi mais próximo de 1,00 em comparação aos outros métodos, exceto para HH em que o valor do ssGBLUP para k-médias foi mais próximo de 1,00.

Os valores de β abaixo e acima de 1,00 observados nos dois tipos de agrupamento e nos três conjuntos de dados (Tabela 8) tendem a prover previsões genômicas superdispersas. Cardoso et al. (2015) trabalhando com a mesma população (Braford e Hereford), também observaram comportamento semelhante ao presente estudo. Saatchi et al. (2013) também observou a superdispersão dos valores genéticos genômicos em estudo com bovinos da raça Angus na América do Norte para a característica de peso.

Comparação entre métodos single-step usando o painel de 50K

Com o conjunto de dados contendo as duas raças (BO_HH), os métodos single-step não apresentaram diferença entre eles, independente do agrupamento utilizado. O método ssGBLUP ponderado apresentou um aumento na capacidade preditiva em relação ao ssGBLUP e SSBR para o k-médias e o aleatório. A utilização da abordagem bayesiana (SSBR) não mostrou ganho em capacidade preditiva em relação aos métodos frequentista. Quanto ao viés da regressão, os valores foram muito abaixo de 1,00 e próximos entre os métodos para os dois tipos de agrupamento (Tabela 8).

Para o conjunto de dados contendo somente animais Braford, não teve diferença entre os métodos ssGBLUP, WssGBLUP_2 e SSBR para o grupo aleatório. Já o WssGBLUP_3 e o WssGBLUP_4 apresentaram diferenças para os métodos ssGBLUP e SSBR. Quanto ao agrupamento k-médias o ssGBLUP foi diferente somente para o WssGBLUP_3 e WssGBLUP_4, já os WssGBLUP ponderados foram diferentes ao SSBR. Não houve diferença entre os WssGBLUP ponderados para os dois tipos de agrupamento. Houve ganho na capacidade preditiva quando se utilizou WssGBLUP_i (i= 2, 3 e 4) em relação ao ssGBLUP e o método bayesiano SSBR. Os valores de viés da regressão foram abaixo de 1,00, mas próximos entre os métodos para K-médias e aleatório (Tabela 8).

Para os animais Hereford, teve diferenças entre os métodos que utilizam a informação genômica para os agrupamentos k-médias e aleatório, exceto entre o WssGBLUP_2 e WssGBLUP_3 para o aleatório que não teve diferença. O WssGBLUP_3 apresentou ganho de 0,18 e 0,13 em capacidade preditiva em relação ao ssGBLUP e 0,06 e 0,05 para o WssGBLUP_2 para k-médias e para o aleatório, respectivamente. Os valores dos coeficientes de regressão para o ssGBLUP ponderado foram maiores que 1,00 para os dois tipos de agrupamento. Para o método ssGBLUP no agrupamento k-médias o valor do viés da regressão foi abaixo de 1,00 e para o aleatório o valor foi muito próximo a 1,00. Para o método SSBR os valores da capacidade preditiva foram negativos (-0,01 e -0,02) para os agrupamentos k-médias e aleatório. O viés da regressão foram 0,00 para k-médias e -0,23 para aleatório (Tabela 8).

Independente do tipo de agrupamento e método que utiliza a informação genômica na avaliação, a capacidade preditiva diminuiu para os conjuntos de dados Hereford e Braford_Hereford. Essa diminuição está relacionada ao fato da raça Hereford

apresentar baixa variabilidade genética para a característica resistência aos carrapatos e também por ter poucos indivíduos com informações fenotípicas e genotípicas.

Tabela 8: Capacidade preditiva média (r_{aa}) e o coeficiente de regressão (β) da validação cruzada dos valores genéticos para diferentes métodos, utilizando o painel de 50K

Conjunto de dados	Métodos	Agrupamento k-médias		Agrupamento aleatório	
		r_{aa}	$\hat{\beta}$	r_{aa}	$\hat{\beta}$
Braford_Hereford	PBLUP	0,10 ± 0,04	0,91	0,17 ± 0,04	1,00
	ssGBLUP	0,12 ± 0,05	0,44	0,18 ± 0,04	0,45
	WssGBLUP_2	0,15 ± 0,05	0,47	0,20 ± 0,04	0,45
	WssGBLUP_3	0,16 ± 0,06	0,45	0,21 ± 0,03	0,45
	SSBR	0,12 ± 0,09	0,47	0,17 ± 0,02	0,46
Braford	PBLUP	0,13 ± 0,05	0,80	0,19 ± 0,03	0,88
	ssGBLUP	0,16 ± 0,03	0,48	0,25 ± 0,01	0,67
	WssGBLUP_2	0,20 ± 0,03	0,52	0,27 ± 0,02	0,66
	WssGBLUP_3	0,23 ± 0,01	0,54	0,29 ± 0,01	0,66
	WssGBLUP_4	0,24 ± 0,01	0,57	0,29 ± 0,01	0,67
	SSBR	0,15 ± 0,03	0,55	0,23 ± 0,01	0,73
Hereford	PBLUP	0,02 ± 0,07	1,05	0,02 ± 0,04	0,26
	ssGBLUP	0,05 ± 0,03	0,62	0,09 ± 0,04	0,94
	WssGBLUP_2	0,17 ± 0,01	1,84	0,17 ± 0,04	1,43
	WssGBLUP_3	0,23 ± 0,03	1,60	0,22 ± 0,03	1,30
	SSBR	-0,01 ± 0,07	0,00	-0,02 ± 0,06	-0,23

PBLUP = BLUP com base no pedigree; ssGBLUP = single step GBLUP; WssGBLUP_i = single-step GBLUP ponderado para iteração i = 2 e 3; SSBR = single step regressão bayesiana BayesB ($\pi=0.99$).

Lee et al. (2017) comparando o ssGBLUP e o SSBR em populações de bovinos de corte da raça Hanwoo, verificaram que os dois métodos apresentaram acurácia de predição semelhantes para características de espessura de gordura, área de olho de lombo e escore de marmoreio, mas para peso da carcaça, o SSBR foi superior ao ssGBLUP.

Capacidade de predição entre grupos da validação cruzada

A capacidade de predição para cada grupo de validação cruzada pelo agrupamento k-médias, com painel de 50K para a característica contagem de carrapatos estão apresentados na Figura 2, 3 e 4.

A capacidade preditiva para o grupo 1 do conjunto de dados BO_HH foram baixas em relação aos outros grupos para todos os métodos, exceto para o método PBLUP que o menor valor foi para o grupo 3 (Figura 4). É possível observar também que para os métodos que utiliza a informação genômica na predição dos valores genéticos, o grupo 3, que possui mais de 90% de animais Hereford, apresentou valores maiores em relação alguns outros grupos. Isso se deve ao fato desses animais terem um maior grau de relacionamentos com alguns animais Braford e Hereford de outros grupos (2, 4 e 5), em comparação aos animais que compõem o grupo 1. Fernando et al. (2015) trabalhando com o mesmo conjunto de dados e utilizando a mesma metodologia para formação dos grupos, também encontrou menores valores de acurácia para grupos onde a maior quantidade de animais da raça Hereford.

Comparando a abordagem frequentista com a bayesiana, para os métodos que utiliza a informação genômica, a abordagem bayesiana foi o que apresentou menores valores para todos os grupos. Entre os métodos frequentista, o WssGBLUP_3 foi o que apresentou maiores valores dos grupos, entre os métodos.

Para o conjunto de dados contem animais somente da raça BO, o PBLUP foi o que apresentou menores valores entre grupos em relação aos outros métodos, exceto ao grupo 4, que foi o que apresentou maior valor (Figura 2). Entre os métodos que utiliza a informação genômica, o bayesiano foi o que apresentou menores valores em comparação aos outros, chegando a ter valores similares apenas com o ssGBLUP. Entre o método frequentista, WssGBLUP_4 foi o que apresentou os maiores valores de capacidade de predição dos grupos em relação aos outros métodos. Essa diferença observada entre grupos para os métodos, é esperado em situação em que os candidatos a seleção diminuem à medida que a média de parentesco dos indivíduos do conjunto de treinamento diminui (Habier et al., 2007).

Para o conjunto de dados contendo somente animais Hereford, os valores de capacidade preditiva tiveram consideráveis variações entre os grupos pelo agrupamento

k-médias para os métodos PBLUP, ssGBLUP e SSBR. O grupo 1 do método PBLUP apresentou valor negativo (-0,05) e os demais grupos foram valores positivos (0,08 e 0,03). No SSBR o grupo 1 e 3 também apresentou valores negativos (-0,06 e -0,05) e o grupo 2 apresentou valor positivo (0,07). Em relação aos métodos single-step, os WssGBLUP foram os que apresentaram maiores valores para todos os grupos (Tabela 3). SAATCHI et al. (2013), utilizando uma população de treinamento com dados de bovinos Hereford da América do Norte, também não produziram previsões acuradas para os rebanhos da mesma raça no Uruguai e na Argentina.

A superioridade observada para o método WssGBLUP para diferente iteração em relação ao SSBR, é porque os métodos bayesianos resultam em previsões reduzidas devido a ignorar SNP de pequeno efeito (Su et al., 2010) e o WssGBLUP considerar todos os SNPs, mesmo que ele tenha pequeno efeito. Isso foi observado no presente estudo porque a característica é influenciada por muito QTL de pequeno efeito.

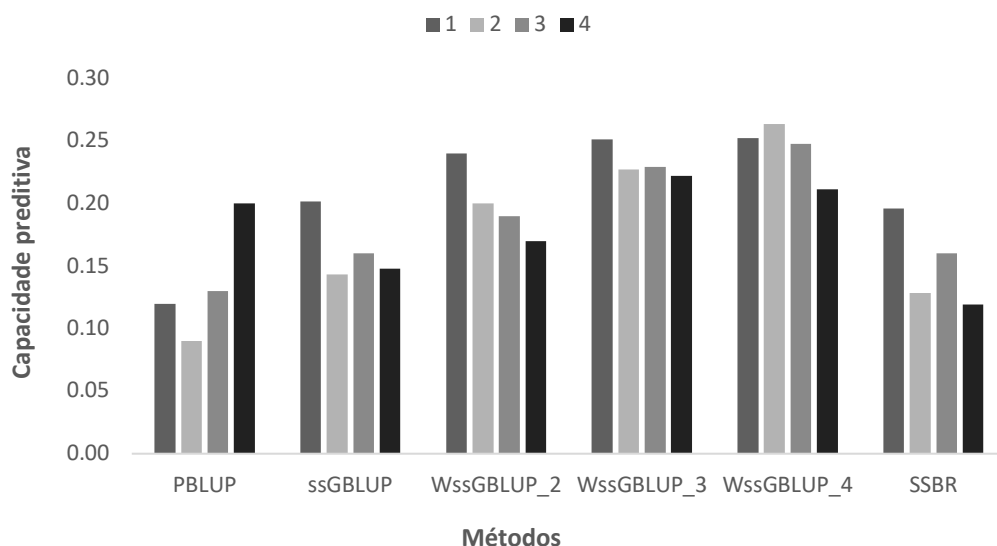


Figura 2: Capacidade preditiva para cada grupo de validação cruzada pelo agrupamento k-médias de acordo com o método de predição: PBLUP = BLUP pedigree; ssGBLUP = single-step GBLUP; WssGBLUP = single-step GBLUP ponderado para a iteração 2, 3 e 4; SSBR = single-step de regressão bayesiana para conjunto de dados contendo apenas animais Braford usando o painel de 50K.

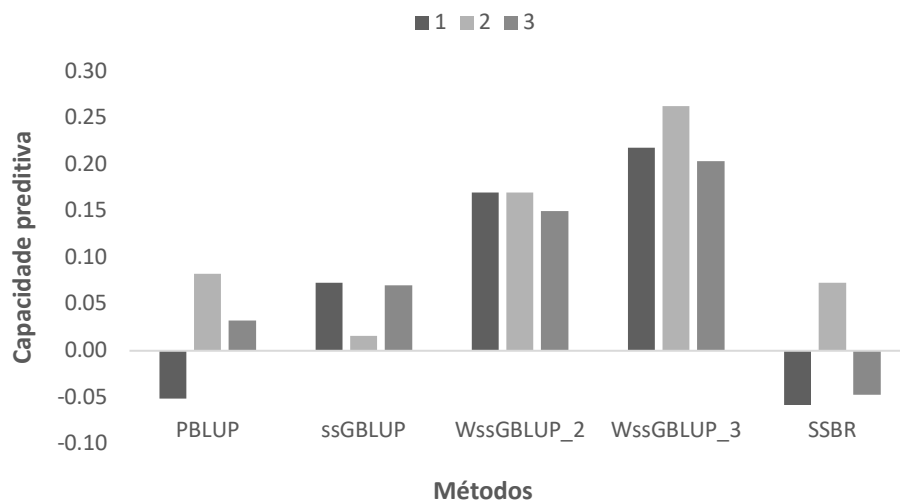


Figura 3: Capacidade preditiva para cada grupo de validação cruzada pelo agrupamento k-médias de acordo com o método de predição: PBLUP = BLUP pedigree; ssGBLUP = single-step GBLUP; WssGBLUP = single-step GBLUP ponderado para a iteração 2 e 3; SSBR = single-step de regressão bayesiana para conjunto de dados contendo apenas animais Hereford usando o painel de 50K.

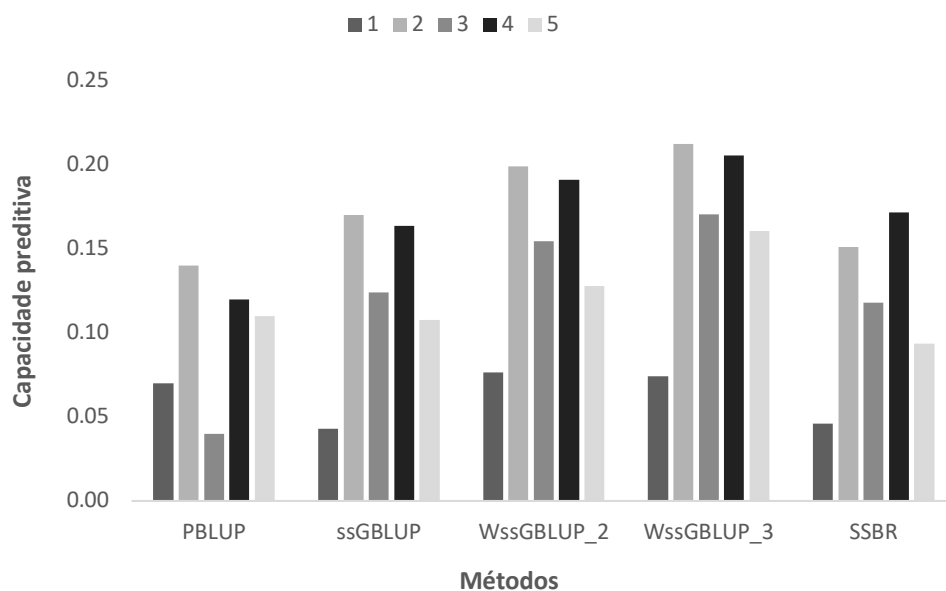


Figura 4: Capacidade preditiva para cada grupo de validação cruzada pelo agrupamento k-médias de acordo com o método de predição: PBLUP = BLUP pedigree; ssGBLUP = single-step GBLUP; WssGBLUP = single-step GBLUP ponderado para iteração 1 e 2; SSBR = single-step de regressão bayesiana para o conjunto de dados contendo as duas raças Braford_Hereford usando o painel de 50K.

CONCLUSÃO

As metodologias que utilizam as informações de pedigree, fenótipos e genômica simultaneamente, resultaram em valores de capacidade de predição superiores ao método baseado somente na informação de pedigree e fenótipo.

A depender da população e raça utilizada, o método WssGBLUP foi superior em relação aos outros métodos single-step.

AGRADECIMENTO

À CAPES e ao CNPq pelo fornecimento de bolsa durante o período de doutorado, a Embrapa Pecuária Sul-Bagé e a Conexão Delta G – Bagé, RS.

REFERÊNCIAS

Aguilar, I.; Misztal, I.; Johnson, D. L.; Legarra, A.; Tsuruta, S.; Lawlor, T. J. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. **Journal of Dairy Science**, v. 93, p. 743-752, 2010.

Ayres, D. R.; Pereira, R. J.; Boligon, A. A.; Silva, F. F.; Schenkel, F. S.; Roso, V. M.; Albuquerque, L. G. Linear and Poisson models for genetic evaluation of tick resistance in cross-bred Hereford x Nelore cattle. **J. Anim. Bree. Genet.** 2013; 130:417-424.

Bodhireddy, P.; Kelly, M. J.; Northcutt, S.; Prayaga, K. C.; Rumph, J.; DeNise, S. Genomic predictions in Angus cattle: Comparisons of sample size, response variables, and clustering methods for cross-validation. **J. Anim. Sci.** 2014; 92:485-497.

Budeli M.A., Nephawe K.A., Norris D., Selapa N.W., Bergh L., Maiwashe A. Genetic parameter estimates for tick resistance in Bonsmara cattle. S. Afr. **J Anim. Sci.** 2009; 39:321–327.

Cardoso, F. F.; Gomes, C. C. G.; Sollero, B. P.; Oliveira, M. M.; Roso, V. M.; Piccoli, M. L.; Higa, R. H.; Yokoo, M. J.; Caetano, A. R.; Aguilar, I. Genomic prediction for tick resistance in Braford and Hereford cattle. **Journal of Animal Science**, v. 93, n. 6, p. 2693-2705, 2015.

Cavani, L.; Cardoso, F. F.; Gomes, C. G.; Caetano, A. R.; Giglioti, R. Estimates of genetic parameter for tick count and infection level of Babesia Bovis traits in Braford and Hereford cattle. **J. Anim Sci**. 2017; 95:101-102.

Cheng H.; Garrick, D. J.; Fernando, R. L. JWAS: Julia implementation of whole-genome analyses software using univariate and multivariate Bayesian mixed effects model. <http://QTLrocks> (2016).

Cornacchia, G. et al. Estimativas do coeficiente de repetibilidade para características fenotípicas de procedências de Pinus tecunumanii (Schw.) Eguluz, Perry e Pinus caribaea var. hondurensis Barret, Golfari. **Revista Árvore**, Viçosa, v.19, n.3, p.333-345, 1995.

Habier D.; Fernando, R. L.; Dekkers, J. C. M. The impact of genetic relationship information on genome-assisted breeding values. **Genetics**, v. 177, p. 2389-2397, 2007.

Hayes, B. J.; Visscher, P. M.; Goddard, M. E. Increased accuracy of artificial selection by using the realized relationship matrix. **Genet. Res**. 91:47-60.

Fernando, R. L.; Dekkers, J. C. M.; Garrick, D. J. A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. **Genetics Selection Evolution**. 2014; 46:50.

Henderson, C. R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, v.31, p.423, 1975.

Lee, J.; Cheng, H.; Garrick, D.; Golden, B.; Dekkers, J.; Park, H.; Lee, D.; Fernando, R. L. Comparison of alternative approaches to singles-trait genomic prediction using genotyped and non-genotyped Hanwoo beef cattle. **Genet Sel Evol.** 2017; 49:2.

Legarra, A., Robert-Granié, C., Manfredi, E., Elsen, J. M., Performance of genomic selection in mice. **Genetics**, v. 180, p. 611-618, 2008.

Legarra A.; Aguilar, I.; Misztal, I. A relationship matrix including full pedigree and genomic information. **J. Dairy Sci** 2009; 92:4656-4663.

Lourenco, D. A. L.; Tsuruta, S.; Fragomeni, B. O.; Masuda, Y.; Aguilar, I. et al. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. **J. Anim Sci.** 2015; 93:2653-2662.

Meuwissen, T. H. E.; Hayes, B. J.; Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, 2001; 157:1819-182.

Misztal, I., Computational techniques in animal breeding. University of Georgia, Athens, GA, v., p., 2000.

Misztal, I., Reliable computing in estimation of variance components. **Journal of animal breeding and genetics**, v. 125, p. 363-370, 2008.

R Core Team. The R Project for Statistical Computing. 2013.

Saatchi, M.; McClure, S. D.; McKay, S. D.; Rolf, M. M.; et al. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. **Genet Sel Evol.** 2011; 43:40.

Saatchi, M., Ward, J., Garrick, D. J., Accuracies of direct genomic breeding values in Hereford beef cattle using national or international training populations. **Journal of animal science**, v. 91, p. 1538-1551, 2013.

Sollero, B. P.; Junqueira, V. S.; Gomes, C. C. G.; Caetano, A. R. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using Bayesian methods. **Genet Sel Evol.** 2017; 49:49.

Su, G., Guldbbrandtsen, B., Gregersen, V. R., and Lund, M. S. (2010). Preliminary investigation on reliability of genomic estimated breeding values in the Danish Holstein population. **J. Dairy Sci.** 93, 1175–1183. doi: 10.3168/jds.2009-2192.

VanRaden, P. M., Efficient methods to compute genomic predictions. **Journal of dairy science**, v. 91, p. 4414-4423, 2008.

Villanueva, B., Pong-Wong, R., Fernández, J. & Toro, M. A. (2005). Benefits from marker-assisted selection under an additive polygenic genetic model. **Journal of Animal Science** 83, 1747–1752.

Wang, H.; Misztal, I.; Aguilar, I.; Legarra, A.; Muir, W. M. Genome-wide association mapping including phenotypes from relatives without genotypes. **Genetics Research**, v. 94, p. 73-83, 2012.

Weigel, D.A.; Gianola, D. Estimation of heterogeneous within-herd variance components using empirical bayes methods: a simulation study. **Journal of Dairy Science**, v.75, p.2824-2833, 1992.

Zhang, Z.; Liu, J.; Ding, X.; Bijma, P.; De Koning, D. J.; Zhang, Q. Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. **PLoS One**, v. 5, n. 9, 2010.

CAPÍTULO 2

SELEÇÃO DE MARCADORES MOLECULARES UTILIZANDO ESTUDO DE ASSOCIAÇÃO GENÔMICA AMPLA BAYESIANA PARA CARACTERÍSTICA CONTAGEM DE CARRAPATOS EM BOVINOS DA RAÇA BRAFORD E HEREFORD

Resumo: Com a informação molecular, estudos de associação genômica ampla (GWAS) permitem identificar loci ligados a determinada característica. O GWAS explora a existência de desequilíbrio de ligação entre SNPs e variantes causais, e assim permitindo selecionar SNPs mais informativos para customização de painéis de marcadores SNP de baixa densidade. Os métodos Bayesianos permitem estimar os efeitos dos SNPs de alta densidade e quando aplicado em GWAS, pode melhorar o mapeamento de regiões ao longo do genoma que contem variantes causais, especialmente no caso de características complexas para as quais a maioria dos SNPs explicam uma pequena proporção da variância total observada. O objetivo desse estudo foi identificar regiões genômicas e selecionar os SNPs mais representativos associados com a características resistência aos carrapatos. Adicionalmente, verificou a capacidade preditiva para os painéis customizados com base em animais da raça Braford e Hereford. Para isso, foram utilizados o método BayesB ($\pi = 0,99$) para a seleção dos marcadores altamente informativos para a característica e o single-step de regressão bayesiana para prever os valores genéticos genômicos. A capacidade preditiva foi obtida pela correlação entre o fenótipo ajustado e o valor genético predito. Os softwares Gensel e o JWAS foram utilizados para a realização desse estudo. Os resultados observados nesse estudo sugerem a possibilidade de selecionar janelas superiores que explicam mais que 0,2% da variância genética para a característica. Os painéis customizados (tag SNP) para animais da raça Braford, Hereford e Braford_Hereford continham 35, 44 e 61 marcadores do tipo SNP que apresentaram maior efeito sobre a característica. O uso das tags SNP para animais representaria ao menos 67%, 72% e 28% da capacidade de predição obtidas pelo painel de 50K para animais Braford_Hereford, Braford e Hereford, respectivamente.

Palavras-chaves: bayesB, bayesiana, bovino de corte, SSBR, tag SNP

INTRODUÇÃO

O controle dos carrapatos é importante porque parasitam uma ampla gama de hospedeiros, são vetores de doenças virais, protozoárias, bacterianas e diminuem a produtividade dos bovinos (Mapholi et al., 2016). No Brasil, a espécie de carrapato *Rhipicephalus (Boophilus) microplus* provoca diversos prejuízos e o controle continua sendo um desafio para a pecuária.

Com o objetivo de controlar a incidência de carrapatos em bovinos, diversos estudos voltados a imunologia foram realizados, para o desenvolvimento de carrapaticidas. Com o uso indiscriminado, a resistência dos carrapatos a esses produtos aumentou, além disso os resíduos químicos na carne, no meio ambiente e o custo de produção também aumentaram. Diante desse cenário, o interesse na genética de adaptação a resistência aos carrapatos vem crescendo em resposta à necessidade de identificar e multiplicar bovinos que são adaptados e eficientes em diferentes ambientes (Porto Neto et al., 2011).

Com o propósito de detectar associação entre as variantes genéticas e o fenótipo, painéis de polimorfismo de nucleotídeos único (SNP) de diferentes densidades vem sendo usados. Os SNPs distribuídos por todo genoma permitem identificar e mapear os principais genes que estão envolvidos no controle à resistência aos carrapatos por um processo chamado estudo de associação genômica ampla (genome wide association study - GWAS) (Mapholi et al., 2016, Sollero et al., 2017).

A predição do valor genético genômico dos indivíduos é baseada em um grande número de SNPs, com uma cobertura de alta densidade. Devido a alta densidade de marcadores, espera-se que cada locus de característica quantitativa (quantitative trait loci – QTL) esteja em desequilíbrio de ligação (DL) com pelo menos um SNP, ou seja, espera-se que os efeitos dos QTL sejam capturados por SNP (Meuwissen et al., 2001). Alguns estudos de GWAS identificaram associação de QTL à resistência aos carrapatos em regiões genômicas para bovinos (Barendse, 2007; Porto Neto et al., 2010 e 2011; Turner et al., 2010; Sollero et al., 2014 e 2017)

Os métodos Bayesianos permitem estimar os efeitos dos SNPs de alta densidade e quando aplicado em GWAS, pode melhorar o mapeamento de regiões ao longo do

genoma que contem variantes causais, especialmente no caso de características complexas para as quais a maioria dos SNPs explicam uma pequena proporção da variância total observada. Identificar esses marcadores mais informativos associados as características complexas pode contribuir para desenvolvimento de um painel de SNP de baixa densidade com alto desempenho preditivo (Sollero et al., 2017). O que seria altamente desejável, uma vez que é necessário soluções de baixo custo para a seleção genômica a ser implementada na maioria dos setores de produção animal (Van Eenennaam et al., 2014).

Portanto os objetivos desse estudo foram: 1) Identificar regiões genômicas e SNPs associados à resistência aos carrapatos, utilizando método Bayesiano BayesB em um painel de 50k SNP de bovinos corte da raça Braford e Hereford; 2) Testar o desempenho preditivo do painel de baixa densidade baseado em um subconjunto de SNPs selecionados e comparar com os resultados obtidos com o painel de SNP de 50K.

MATERIAL E MÉTODOS

A aprovação do Comitê de Cuidados e Uso de Animais não foi obtida para este estudo porque o conjunto de dados foi gerado a partir de vários estudos anteriores.

Arquivo de dados fenotípicos, genotípicos e pedigree

Todos os 13.874 registros de contagem de carrapatos, coletados no lado esquerdo do corpo, foram registrados para cada animal nascido entre 2008 e 2014, uma, duas e três vezes consecutivas, durante o período pós-desmame. Estes registros utilizados nas análises foram obtidos de animais da raça Braford e Hereford de diferentes fazendas do programa de melhoramento da conexão Delta G (Rio Grande do Sul, Brasil).

A característica contagem de carrapato foi transformada por $\log_{10}(x_i + 1.0)$, onde x_i é a quantidade de carrapatos.

Como a população é composta por duas raças (Braford e Hereford), foram gerados três bancos de dados para realização dessa pesquisa. O primeiro conjunto de

dados (BO) continha somente animais da raça Braford com 9.439 fenótipos (4.011 animais), o segundo conjunto de dados (HH) continha somente animais da raça Hereford com 4.460 registros (1.856 animais) e o terceiro conjunto de dados (BO_HH) continha animais das duas raças com 13.874 fenótipos (5.867 animais). O agrupamento da raça Braford e Hereford no terceiro conjunto de dados, se deve ao fato do objetivo comum de melhoramento pelo Programa de Melhoramento da Conexão Delta G e devido a similaridade genética das duas raças, já que os Brafords são compostos com uma contribuição de 62,5% da raça Hereford (Piccoli et al., 2014; Biegelmeyer et al., 2016).

Um total de 4.529 indivíduos foram genotipados em painéis de diferentes densidades. Sendo que 230 reprodutores foram genotipados no painel de alta densidade (BovineHD – Illumina® bread chip – 777.962 SNPs) e 162, 3.728, 287, 122 animais genotipados com painéis de SNPs de 150K, 50K, 30K e 27K, respectivamente. Todos os painéis foram filtrados para selecionar apenas SNP que estavam presentes no painel de 50K. O controle de qualidade foi realizado usando o pacote R/snpStats (Clayton, 2015). Diferentes critérios de controle de qualidade foram utilizados para SNPs e amostras. Os critérios adotados para exclusão de SNPs foram: alelo menos frequente (MAF) ($< 3\%$), taxa de genotipagem (Call rate) ($< 98\%$) e Equilíbrio de Hardy-Weinberg ($P > 10^{-7}$). Para animais, os critérios utilizados para avaliação da exclusão foram: Call rate ($< 90\%$), desvio de heterozigose com três desvios padrões acima ou abaixo da média observada dos animais genotipados, genótipos idênticos ($> 99,5\%$), erro de identificação sexual e dados duplicados. Após a edição, permaneceram um total de 39.922 marcadores SNP e 4.493 animais e que foram utilizadas nas análises. Dos 4.493 animais, 3.232 animais da raça Braford e 1.261 da raça Hereford. Os genótipos perdidos foram imputados utilizando o algoritmo sliding window com o software FImpute versão 2.2 (Sargolzaei et al., 2011). Com o objetivo de obter uma maior acurácia no processo de imputação, o pedigree da população foi utilizado.

Estudo de associação genômica ampla (GWAS) bayesiano

Primeiramente, os componentes de variância e posteriormente os valores genéticos preditos (Estimated breeding values - EBV) foram obtidos ajustando-se um

modelo animal de repetibilidade baseado em pedigree aos dados de contagem de carrapatos, com os programas REMLF90 e BLUPF90 (Misztal et al., 2002). O seguinte modelo de repetibilidade, em notação matricial, foi utilizado nas análises:

$$y = X\beta + Za + Wpe + e \quad [1]$$

em que, y é o vetor de observações; X , Z e W são matrizes de incidência para os efeitos fixos, genético e de ambiente permanente, respectivamente; β são os efeitos de grupo contemporâneo (fixo), proporção racial (covariável), heterozigosidade (covariável), perda por recombinação (covariável), idade no dia da contagem linear e quadrático (covariável); a é o efeito genético aditivo, assumindo que $a \sim N(0, A\sigma_a^2)$, em que A é a matriz de numeradores dos coeficientes de parentesco e σ_a^2 é a variância genética aditiva; pe é o efeito de ambiente permanente, assumindo que $pe \sim N(0, I\sigma_p^2)$, em que I é a matriz identidade e σ_p^2 é a variância de ambiente permanente; e e é o termo residual, assumindo que $e \sim N(0, I\sigma_e^2)$, em que σ_e^2 é a variância genética residual.

Posteriormente, os valores genéticos preditos para contagem de carrapatos foram desregredidos (deregression breeding values – DEBV) e correspondentes ponderadores (w) = {wj} para todos os animais, livre do efeito médio dos pais, de acordo com Garrick et al. (2009). Estes “pseudo fenótipos” ou DEBV foram então analisados em um modelo aleatório com efeito de substituição alélica de SNP (Sollero et al., 2017) utilizando o programa GenSel versão 4.0 (Fernando & Garrick, 2009) para seleção de janelas top e tag SNP. Os métodos Bayesianos BayesA e BayesB foram aplicados para analisar os dados de DEBV usando genótipos como variáveis explicativas (Meuwissen et al., 2001). Em BayesA e BayesB, cada SNP tem uma variância específica para cada locus, que considera uma distribuição qui-quadrado invertida escalonada $X^{-2}(v, S)$ com graus de liberdade v e escala S determinados no GenSel por padrão. A especificação a priori para efeito de SNP no BayesB assume que uma proporção de SNP tenha efeito zero (nulo), com probabilidade fixa π , enquanto que os SNPs remanescentes têm efeito normalmente distribuídos com variância de locus específica e uma probabilidade $1 - \pi$. Por outro lado, o método BayesA todos os SNPs foram ajustados com $\pi = 0$, para cada ciclo de Monte

Carlos via Cadeia de Markov (MCMC). O modelo estatístico utilizado para as duas análises Bayesianas foi:

$$y = \sum_{i=1}^{k=39.922} \delta_i z_i a_i + e,$$

em que, y é o vetor de fenótipo (DEBV); k é o total de marcadores SNPs no painel; δ_i é o indicador de que o SNP i foi incluído ($\delta_i = 1$) ou excluído ($\delta_i = 0$) do modelo para uma dada iteração do MCMC; z_i é o vetor de genótipos ajustado para cada SNP i , codificado como -10/0/10; a_i é o efeito de substituição aleatória para o SNP i ajustado com sua própria variância $\sigma_{a_i}^2$ e um efeito a priori nulo com probabilidade π ou um efeito não-nulo com probabilidade $1 - \pi$; e e é o vetor aleatório residual com distribuição normal. Enquanto para BayesA, δ_i foi sempre igual a 1. Para BayesB o parâmetro π foi assumido ser igual a 0,99 (Sollero et al. 2017). Uma cadeia de 52.000 iterações foi utilizada, em que as primeiras 2.000 foram descartadas como burn-in. A convergência da cadeia MCMC foi verificada pelo teste Geweke (1992) usando o pacote “boa” (Bayesian output analysis) do program R (Smith, 1992).

Com base nas análises estatística e nos resultados dos componentes de variância obtidos com a equação [1], estes foram utilizados como priori para rodar as análises de GWAS bayesiana no GenSel.

Seleção das Janelas Top e Tag SNPs

Os SNPs foram alocados em 2.530 janelas do genoma de 1-Mb não sobrepostas, que continham diferentes quantidade de SNP com base na ordem do mapa físico derivada do conjunto de genoma bovino (*Bos taurus*) UMD (Zimin et al., 2009). A variância genética explicada conjuntamente por cada subconjunto de SNP atribuídos a cada janela (“variação da janela”), foi estimada na iteração da cadeia MCMC e

posteriormente convertida na proporção da variância genética total explicada pela janela (Wolc et al., 2012).

As regiões do genoma que potencialmente inclui loci de características quantitativa (QTL) associadas à resistência aos carrapatos, referida como janela top, foram identificadas a partir de um limiar definido com base na contribuição da variância genética descrito por Schurink et al., (2012). As janelas tops foram identificadas no GWAS considerando 4.493 animais e 39.922 marcadores SNP, utilizando o método BayesB ($\pi = 0,99$) para os três conjuntos de dados. Assumindo contribuição igual a todas as regiões genômicas, a proporção esperada da variância genética explicada por cada uma das 2.530 janelas é de 0,04%. Assim, janelas com tamanho de 1-Mb que explica pelo menos 0,2% da variância genética, o que corresponde a cinco vezes a variância esperada ($0,04\% * 5 = 0,2\%$), foram consideradas como região potencialmente associada a QTL (Onteru et al., 2013; Zare et al., 2014, Sollero et al. 2017) e selecionadas para posteriores análises.

Para identificar potenciais SNP para construção de painéis de baixa densidade, a estratégia de seleção de tag SNP foi testada dentro das janelas tops, considerando os parâmetros frequência do modelo (Model Frequency – MF), t-like (TL), desequilíbrio de ligação (linkage disequilibrium - LD) e alelos de menor frequência (Minor Allele Frequency - MAF). No software GenSel, o parâmetro MF reflete a proporção de cadeias que incluíram determinado marcador no modelo, enquanto o TL reflete a razão do efeito médio a posteriori sobre o desvio padrão das distribuições destes efeitos, considerando somente as cadeias que incluíram o efeito no modelo. O pacote R/snpStats (Clayton, 2012) foi utilizado para obter os valores de LD.

A estratégia utilizada para selecionar SNP foi a seguinte: 1) Identificação de SNP com maior valor de MF dentro de cada janela top; 2) Identificação do menor valor dentre os maiores valores de MF dos SNP top selecionados nas janelas tops; 3) Seleção de SNP com MF acima do menor valor dentre os maiores valores de MF para SNP top – este passo teve o objetivo de selecionar SNP que não era tops de suas janelas, mas que tinham valor de MF que excedia o valor de MF dos SNPs tops localizados em outras janelas selecionadas. Uma semelhante abordagem foi utilizada para avaliar a consistência do efeito do SNP considerando TL; 4) Dos SNPs pré-selecionados baseado

no MF, o valor mínimo de TL foi determinado e definido como o limiar para selecionar os SNP remanescentes dentro das janelas tops que excediam esse valor mínimo de TL; 5) Com o propósito de remover SNP redundantes devido ao alto LD observado entre os subconjuntos de SNP pré-selecionados por MF e TL, os SNPs com valores r^2 maior que 0,4 (Badke et al., 2012), apenas o SNP com maior MAF foi selecionado.

Capacidade preditiva dos painéis tag SNP selecionados

Os 4.040 animais genotipados e fenotipados (2.978 Braford e 1.062 Hereford) foram divididos em quatro, três e cinco subgrupos para os conjuntos de dados Braford, Hereford e Braford_Hereford, respectivamente. Duas estratégias foram utilizadas: agrupamento k-médias de acordo com o relacionamento genômico e agrupamento aleatoriamente (preservando o mesmo número de grupos pelo k-médias) de acordo com Saatchi et al. (2011), usando o pacote R/base. Para a validação cruzada dentro da estratégia de agrupamento, primeiramente foram selecionados subconjuntos de SNP utilizando o método BayesB ($\pi = 0,99$) como descrito acima, usando três dos quatros, dois dos três e quatro dos cincos subgrupos do conjunto de dados de BO, HH e BO_HH, respectivamente.

Para checar a eficiência da escolha dos SNPs mais informativos para predição genética para resistência aos carrapatos, o método BayesA foi usado na metodologia de regressão bayesiana de único passo (single step bayesian regression - SSBR), denominando no presente estudo de SSBR-A, a cada um dos painéis tag SNP proposto, na estratégia de validação cruzada para cada grupo de forma independente, assumindo que todos os SNPs tops possuem efeito não nulo ($\pi = 0,00$).

A capacidade de predição dentro de cada grupo da validação cruzada, foi estimada como a correlação entre os fenótipos ajustados para os efeitos fixo (α^*) e o valores genéticos genômico preditos (\hat{a}). O viés da predição foi avaliado pela regressão simples do fenótipo ajustado e o valor genético predito. Para a metodologia SSBR foi utilizado JWAS: pacote Julia para análise de seleção genômica (Cheng et al., 2016) com base no método de Monte Carlos via Cadeias de Markov, utilizando 100.000 ciclos, período de descarte inicial de 10.000 amostras e intervalo de amostragem de 100. Para

os métodos PBLUP e ssGBLUP foi utilizado o software da família BLUPF90 (Misztal, 2000, 2008)

RESULTADOS E DISCUSSÕES

Grupos de animais

O número de indivíduos e o relacionamento genômico entre e dentro dos grupos, baseado no agrupamento k-médias, para animais Braford, Hereford e Braford_Hereford estão na Tabela 1, 2 e 3, respectivamente. Para os conjuntos de dados contendo apenas animais Braford e Hereford, o maior relacionamento dentro de grupo foram observados nos grupos com menor número de indivíduos. Já para o conjunto de dados contendo as duas raças Braford_Hereford, o maior relacionamento observado dentro de grupo foi no grupo três, onde apresenta a maior porcentagem de animais com proporção da raça Hereford.

Tabela 1. Número de animais (N), relacionamento genômico (Gij) dentro e entre grupos (\pm SD) para agrupamento k-médias para o conjunto de dados Braford.

Grupo	N	Gij dentro de grupo	Gij entre grupos
1	913	0,05 \pm 0,04	-0,01 \pm 0,05
2	453	0,01 \pm 0,04	0,00 \pm 0,04
3	1293	0,01 \pm 0,03	0,00 \pm 0,03
4	319	0,11 \pm 0,07	-0,01 \pm 0,04

Para conjunto de dados Braford o número de animais a partir do agrupamento aleatório para cada um dos quatro grupos continha 808 animais. Para o conjunto de dados Hereford a quantidade de animais para cada um dos três grupos foi de 354 animais. Já para o conjunto de dados Braford_Hereford, a quantidade de animais foram de 745 (grupos um e dois) e 744 (grupos três e quatro). O relacionamento genômico foram similares dentro e entre grupos para os três conjuntos de dados.

Tabela 2. Número de animais (N), relacionamento genômico (Gij) dentro e entre grupos (\pm SD) para agrupamento k-médias para o conjunto de dados Hereford.

Grupo	N	Gij dentro de grupo	Gij entre grupos
1	333	0,08 ± 0,05	0,01 ± 0,06
2	85	0,34 ± 0,03	0,00 ± 0,07
3	644	0,11 ± 0,04	-0,02 ± 0,06

Tabela 3. Número de animais (N), proporção da raça Hereford, relacionamento genômico (Gij) dentro e entre grupos (\pm SD) para agrupamento k-médias para o conjunto de dados Braford_Hereford.

Grupo	N	Proporção de Hereford	Gij dentro de grupo	Gij entre grupos
1	497	86,12	0,08 ± 0,04	-0,01 ± 0,06
2	1356	0,59	0,05 ± 0,04	-0,02 ± 0,06
3	656	94,05	0,11 ± 0,05	-0,02 ± 0,06
4	962	2,29	0,01 ± 0,03	0,00 ± 0,03
5	569	1,41	0,02 ± 0,04	0,00 ± 0,03

Estudo de associação genômica ampla

Para a análise BayesB, que inclui todos os animais e marcadores SNP simultaneamente, o parâmetro $\pi = 0,99$ foi utilizado, o que corresponde a 1% dos SNPs são ajustados no modelo a cada iteração da cadeia MCMC. As herdabilidade (h^2) estimadas foram de 0,27, 0,05 e 0,13 para os conjuntos de dados Braford, Hereford e Braford_Hereford, respectivamente. A h^2 para BO foi superior em comparação a herdabilidade baseada no pedigree ($h^2 = 0,24$) e para HH e BO_HH os valores foram menores ($h^2 = 0,07$ e 0,18, respectivamente). Utilizando o mesmo método BayesB ($\pi = 0,99$), Sollero et al (2017) reportaram herdabilidade menor (0,11) utilizando a mesma população (BO_HH). Essa diferença provavelmente se deve ao fato na diferença de quantidade de registro fenotípicos entre os dois estudos (10.673 registro de contagem de carrapatos contra 13.874). Quando os mesmos autores no mesmo estudo mudaram o valor do parâmetro π para 0,9999, o valor de h^2 foi de 0,02. A diferença observada nos valores de h^2 , deve-se a fato que maiores valores de π pode ser mais discriminante na identificação de QTL e, assim, influenciando na proporção da variância genética explicada pelos marcadores (Fernando & Garrick, 2013). O baixo valor de h^2 para

Hereford encontrado no presente estudo, assemelha-se ao encontrado por Mapholi et al., (2016), que analisaram contagem de carrapatos *A. hebraeum* em bovinos Nguni da África do Sul. Simultaneamente, o método BayesA foi utilizado nos três conjuntos de dados, e os valores de herdabilidade foram maiores para BO (0,37) e semelhante para HH e BO_HH (0,05 e 0,13, respectivamente) comparando com o BayesB.

A proporção da variância genética explicada por cada uma das 2.530 janelas de 1-Mb, incluindo todos os 39.922 SNPs em todo o genoma para os conjuntos de dados Braford, Hereford e Braford_Hereford estão apresentados na Figura 1, 2 e 3, respectivamente.

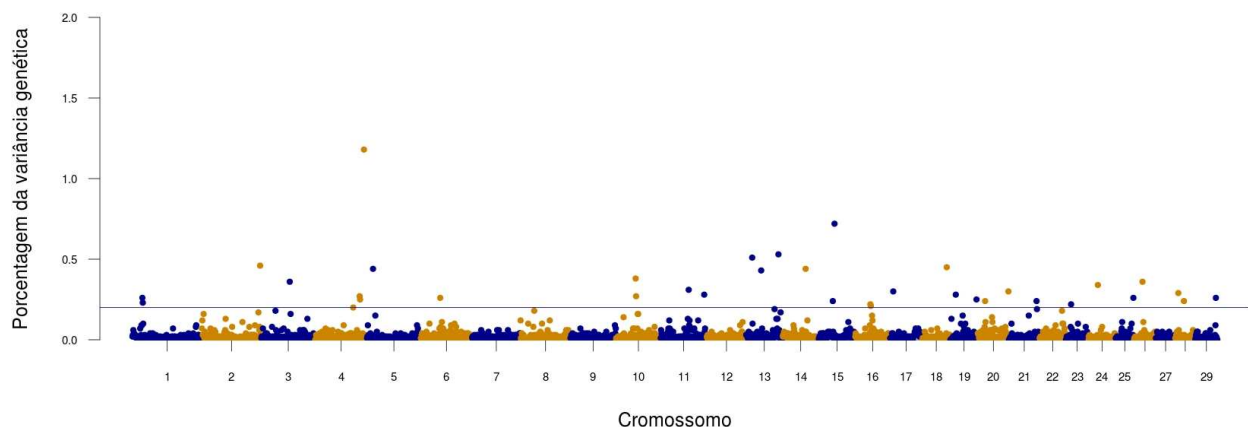


Figura 1. Gráfico de Manhattan plot mostrando os resultados para o estudo de associação genômica ampla realizado pelo método BayesB ($\pi = 0,99$) para resistência aos carrapatos dos animais do conjunto de dados Braford. O eixo Y representa a proporção da variância genética total explicada pela janela de 1-Mb através do genoma bovino e o eixo X representa a localização cromossômica das janelas (2530 janelas não sobrepostas). Janelas explicando mais de 0,2 % da variância genética estão acima da linha preta horizontal.

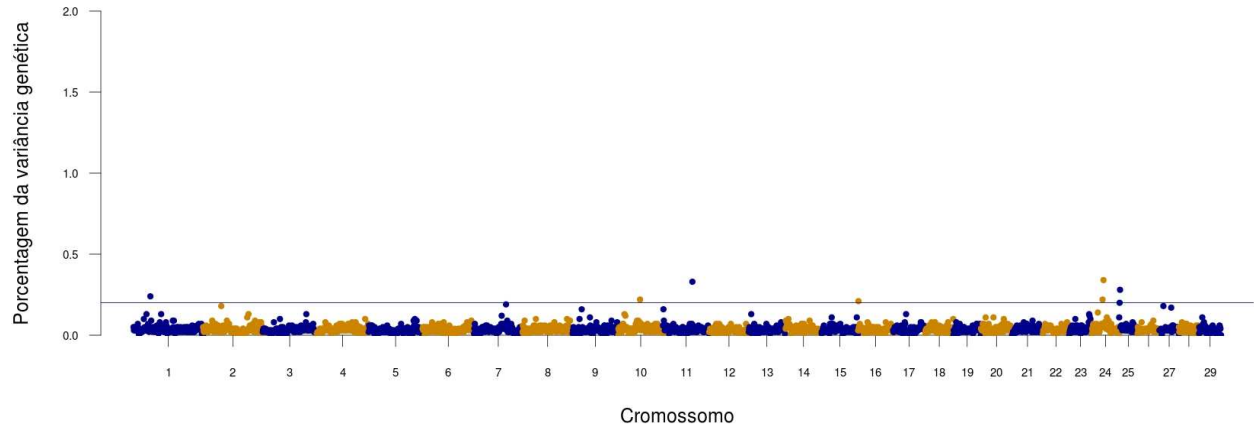


Figura 2. Gráfico de Manhattan plot mostrando os resultados para o estudo de associação genômica ampla realizado pelo método BayesB ($\pi = 0,99$) para resistência aos carrapatos dos animais do conjunto de dados Hereford. O eixo Y representa a proporção da variância genética total explicada pela janela de 1-Mb através do genoma bovino e o eixo X representa a localização cromossômica das janelas (2530 janelas não sobrepostas). Janelas explicando mais de 0,2 % da variância genética estão acima da linha preta horizontal.

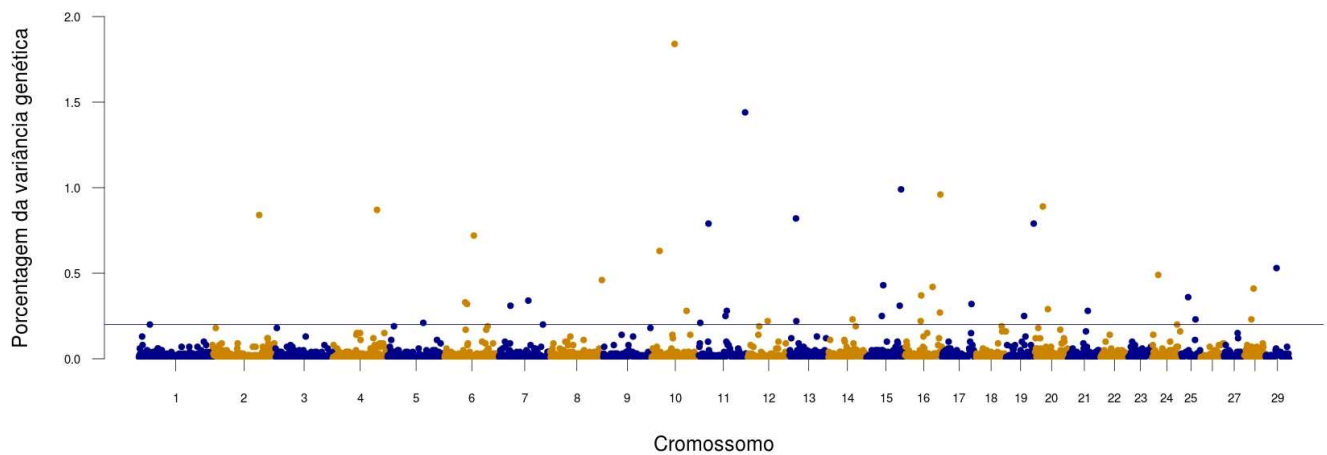


Figura 3. Gráfico de Manhattan plot mostrando os resultados para o estudo de associação genômica ampla realizado pelo método BayesB ($\pi = 0,99$) para resistência aos carrapatos dos animais do conjunto de dados Braford_Hereford. O eixo Y representa a proporção da variância genética total explicada pela janela de 1-Mb através do genoma bovino e o eixo X representa a localização cromossômica das janelas (2530 janelas não sobrepostas). Janelas explicando mais de 0,2 % da variância genética estão acima da linha preta horizontal.

Usando o método BayesB ($\pi = 0,99$) no GWAS, as janelas mais representativas para contagem de carrapatos foram identificadas, nas quais estas, foram consideradas contendo QTL.

Para conjunto de dados contendo somente animais Braford, a quantidade de SNP incluídos nas janelas variou de 10 a 28. Um total de 36 janelas representadas por 658 SNP explicaram 12,50% da variância genética e foram consideradas como janelas superiores contendo QTL (Tabela 4). As oito primeiras janelas juntas, explicam aproximadamente 5% da variância genética e que podem ser observadas nos BTA4, 15, 13, 2, 18, 14 e 5. A primeira janela superior explica individualmente mais que 1% da variância aditiva para resistência aos carrapatos, e está localizada no BTA4 (1,18%). Apesar de não encontrar na literatura uma relação da região cromossômica BTA4 com resistência aos carrapatos, Hanotte et al. (2003) relataram QTLs relacionados à resistência à infecção pelos tripanossomas, em bovinos na mesma região. O tripanossoma é um protozoário transmitido através do sangue, que também acomete as células sanguíneas do seu hospedeiro e desencadeia sintomatologia bem parecida com a observada na babesiose, que é transmitida pelos carrapatos.

Para o conjunto de dados contendo apenas animais Hereford, o número de SNPs nas janelas variou de 14 a 27. Um total de 8 janelas representadas por 142 SNPs explicam juntos mais que 2% da variância genética (Tabela 5). Nenhuma janela top individualmente explicou mais que 0,4% da variância aditiva, indicando que para animais Hereford não há cromossomos que tenha grande importância e onde possivelmente teria presença de genes que exercesse grande efeito sobre a característica. Diversos trabalhos têm reportado a região cromossômica BTA11 associada à resistência aos carrapatos em gado leiteiro australiano e Brahman usando marcadores SNPs, e em bovinos brasileiros utilizando marcadores microssatélites (Turner et al., 2010; Machado et al., 2010; Porto Neto et al., 2011 e Sollero et al., 2014). Na presente pesquisa a região cromossômica BTA11 explica somente 0,33% da variância genética aditiva.

Já para o conjunto de dados BO_HH a quantidade de SNPs nas janelas variou de 9 a 30. Um total de 45 janelas representadas por 770 SNPs foram responsáveis por explicar conjuntamente mais que 21,49% da variância genética e também foram considerados como janelas tops contendo QTL (Tabela 6). As nove primeiras janelas

juntas, explicam aproximadamente 10% da variância genética e que podem ser observadas nas regiões cromossômicas BTA10, 11, 15, 16, 20, 4, 2 e 13. As duas primeiras janelas superiores explicaram individualmente mais que 1% da variância aditiva para resistência aos carrapatos, e estão localizadas na região BTA10 (1,84%) e BTA11 (1,44%). Mapholi et al., (2016) também encontraram SNPs no BTA10 que tiveram efeito significativo para resistência aos carrapatos *A. hebraeum*, coletado no corpo inteiro dos animais.

Tabela 4. Janelas tops que explicam as maiores porcentagens da variância genética para resistência aos carrapatos para o conjunto de dados da raça bovina Braford

Obs	Window	Nome SNP Inicial	Nome SNP Final	N SNP	%Var	Cum Var	chr_Mb	Nome TOP SNP	ModelFreq	t.like	efeito padrão
1	536	Hapmap57291-ss46526771	Hapmap22875-BTA-155031	14	1.18	1.18	4_113	Hapmap22875-BTA-155031	0.8265	1.567	1.620
2	1631	ARS-BFGL-BAC-27751	ARS-BFGL-NGS-115263	16	0.72	1.91	15_37	ARS-BFGL-NGS-5811	0.625	1.185	1.014
3	1500	Hapmap31079-BTA-128662	ARS-BFGL-NGS-114018	22	0.53	2.44	13_75	Hapmap31079-BTA-128662	0.5647	1.125	0.944
4	1439	Hapmap48542-BTA-97857	ARS-BFGL-NGS-27497	11	0.51	2.95	13_14	Hapmap44228-BTA-34185	0.4566	1.049	0.698
5	296	BTA-52274-no-rs	ARS-BFGL-NGS-63882	25	0.46	3.41	2_135	ARS-BFGL-NGS-2645	0.3612	1.001	0.572
6	1893	BTA-43856-no-rs	ARS-BFGL-NGS-53433	16	0.45	3.86	18_57	BTA-43856-no-rs	0.4228	1.032	0.634
7	1563	ARS-BFGL-NGS-115527	ARS-BFGL-NGS-1112	14	0.44	4.3	14_54	BTB-00915241	0.3572	1.003	0.525
8	558	Hapmap36482- SCAFFOLD163485_1458	BTA-87049-no-rs	13	0.44	4.74	5_14	Hapmap52967-rs29017027	0.353	1.001	0.503
9	1460	ARS-BFGL-BAC-14435	ARS-BFGL-NGS-19638	20	0.43	5.17	13_35	ARS-BFGL-NGS-117085	0.378	1.008	0.573
10	1167	ARS-BFGL-NGS-113665	ARS-BFGL-NGS-10383	21	0.38	5.55	10_46	ARS-BFGL-NGS-60054	0.2553	0.964	0.352
11	365	Hapmap40185-BTA-68018	ARS-BFGL-NGS-2777	23	0.36	5.91	3_66	BTB-01946657	0.3497	0.987	0.630
12	2353	ARS-BFGL-NGS-10103	Hapmap46411-BTA-15820	19	0.36	6.27	26_20	BTB-00930925	0.3688	1.005	0.545
13	2248	Hapmap38610-BTA-57584	ARS-BFGL-NGS-114549	16	0.34	6.61	24_22	ARS-BFGL-NGS-26399	0.3305	0.99	0.469
14	1291	ARS-BFGL-NGS-44192	Hapmap60779-rs29022104	13	0.31	6.93	11_65	Hapmap60779-rs29022104	0.2102	0.951	0.299
15	2038	Hapmap44700-BTA-34998	ARS-BFGL-NGS-19159	16	0.3	7.23	20_71	ARS-BFGL-NGS-13702	0.1883	0.938	0.298
16	1768	ARS-BFGL-NGS-5880	BTA-122662-no-rs	22	0.3	7.52	17_7	ARS-BFGL-BAC-27352	0.3128	0.984	0.444
17	2438	Hapmap38171-BTA-64624 Hapmap35067-	BTB-00975363	28	0.29	7.82	28_7	BTB-00975169	0.2948	0.977	0.420
18	1915	BES1_Contig631_717	UA-IFASA-8072	16	0.28	8.1	19_13	ARS-BFGL-NGS-101272	0.222	0.948	0.340
19	1327	ARS-BFGL-NGS-112243	ARS-BFGL-NGS-12954	17	0.28	8.39	11_101	ARS-BFGL-NGS-111179	0.3164	0.987	0.427
20	1168	UA-IFASA-5422	ARS-BFGL-NGS-64260	22	0.27	8.66	10_47	ARS-BFGL-NGS-30594	0.2042	0.943	0.330
21	526	BTA-25514-no-rs	ARS-BFGL-NGS-117579	17	0.27	8.93	4_103	ARS-BFGL-NGS-113196	0.2938	0.977	0.406
22	2331	Hapmap30960-BTC-030209	ARS-BFGL-NGS-43920	21	0.26	9.19	25_41	ARS-BFGL-NGS-33887	0.1693	0.94	0.212
23	23	BTB-01153166	ARS-BFGL-NGS-103546	19	0.26	9.44	1_23	BTB-01153166	0.1249	0.911	0.210
24	2526	Hapmap24672-BTA-140771	ARS-BFGL-NGS-29493	28	0.26	9.7	29_48	BTA-66199-no-rs	0.2996	0.978	0.399

Tabela 4. Continuação

Obs	Window	Nome SNP Inicial	Nome SNP Final	N SNP	%Var	Cum Var	chr_Mb	Nome TOP SNP	ModelFreq	t.like	efeito padrão
25	712	Hapmap30420-BTC-039335	Hapmap48206-BTA-119876	19	0.26	9.96	6_45	Hapmap47403-BTA-76048	0.2081	0.951	0.266
26	1963	ARS-BFGL-NGS-119228	ARS-BFGL-NGS-27079	22	0.25	10.21	19_61	Hapmap43271-BTA-46356	0.2083	0.953	0.265
27	527	ARS-BFGL-NGS-105153	ARS-BFGL-NGS-21106	15	0.25	10.46	4_104	ARS-BFGL-NGS-32907	0.1931	0.945	0.263
28	1627	Hapmap50943-BTA-36514	BTA-46694-no-rs	20	0.24	10.7	15_33	Hapmap24803-BTA-129848	0.2795	0.974	0.377
29	2103	ARS-BFGL-NGS-43652	ARS-BFGL-NGS-63564	12	0.24	10.94	21_64	ARS-BFGL-NGS-75411	0.1571	0.937	0.205
30	1984	ARS-BFGL-NGS-13160	BTB-00774644	19	0.24	11.18	20_17	BTB-00770554	0.1468	0.934	0.196
31	2451	BTB-00980670	Hapmap34915- BES7_Contig278_1082	17	0.24	11.42	28_20	Hapmap34915- BES7_Contig278_1082	0.2053	0.952	0.251
32	24	UA-IFASA-5116	BTA-49319-no-rs	22	0.23	11.65	1_24	BTA-39462-no-rs	0.1155	0.913	0.185
33	1714	Hapmap47935-BTA-38630	BTA-38698-no-rs	16	0.22	11.87	16_35	Hapmap45766-BTA-38696	0.1115	0.917	0.159
34	2185	ARS-BFGL-NGS-116348	ARS-BFGL-NGS-111735	16	0.22	12.09	23_12	ARS-BFGL-NGS-70126	0.1456	0.926	0.207
35	1715	BTA-38712-no-rs	BTB-00637815	21	0.21	12.3	16_36	Hapmap38223-BTA-38319	0.1457	0.93	0.209
36	511	Hapmap40678-BTA-122051	UA-IFASA-2769	10	0.2	12.5	4_88	Hapmap58408-rs29020693	0.2178	0.951	0.293

Tabela 5. Janelas tops que explicam as maiores porcentagens da variância genética para resistência aos carrapatos para o conjunto de dados da raça bovina Hereford

Obs	Window	Nome SNP Inicial	Nome SNP Final	N SNP	%Var	Cum Var	chr_Mb	Nome TOP SNP	ModelFreq	t.like	efeito padrão
1	2253	ARS-BFGL-NGS-57713	BTB-00866491	14	0.34	0.34	24_27	ARS-BFGL-NGS-96359	0.0175	0.502	0.0130
2	1294	Hapmap48260-BTA- 24589	Hapmap42399-BTA- 102134	18	0.33	0.67	11_68	ARS-BFGL-NGS-10829	0.0331	0.764	0.0477
3	2292	ARS-BFGL-NGS-23131	ARS-BFGL-NGS-10250	18	0.28	0.95	25_2	ARS-BFGL-NGS-25277	0.0391	0.772	0.0578
4	39	ARS-BFGL-NGS-84347	BTB-00018345	15	0.24	1.2	1_39	Hapmap47738-BTA-79246	0.0392	0.789	0.0655
5	2251	ARS-BFGL-NGS-85255	ARS-BFGL-NGS-63047	14	0.22	1.42	24_25	Hapmap22882-BTA-161137	0.0240	0.628	0.0309
6	1172	ARS-BFGL-NGS-1854	ARS-BFGL-NGS-24556	27	0.22	1.64	10_51	ARS-BFGL-NGS-101657	0.0309	0.759	0.0422
7	1681	ARS-BFGL-NGS-108427	ARS-BFGL-NGS-32184	20	0.21	1.85	16_1	ARS-BFGL-NGS-10856	0.0458	0.796	0.0739
8	2291	ARS-BFGL-NGS-110520	Hapmap23849-BTC- 016077	16	0.2	2.04	25_1	ARS-BFGL-BAC-44214	0.0240	0.749	0.0289

Tabela 6. Janelas tops que explicam as maiores porcentagens da variância genética para resistência aos carrapatos para o conjunto de dados das duas raças bovina Braford_Hereford

Obs	Window	Nome SNP Inicial	Nome SNP Final	N SNP	%Var	Cum Var	chr_Mb	Nome TOP SNP	ModelFreq	t.like	efeito padrão
1	1172	ARS-BFGL-NGS-1854	ARS-BFGL-NGS-24556	27	1.84	1.84	10_51	Hapmap58695-rs29019899	0.814	1.524	1.789
2	1327	ARS-BFGL-NGS-112243	ARS-BFGL-NGS-12954	17	1.44	3.28	11_101	ARS-BFGL-NGS-111179	0.816	1.524	1.702
3	1670	ARS-BFGL-NGS-112176	BTB-00616941	22	0.99	4.27	15_76	ARS-BFGL-NGS-114766	0.361	0.996	0.701
4	1756	ARS-BFGL-NGS-23665	BTB-01698874	22	0.96	5.23	16_77	Hapmap48320-BTA-40139 Hapmap34041-	0.646	1.195	1.378
5	1984	ARS-BFGL-NGS-13160	BTB-00774644	19	0.89	6.12	20_17	BES1_Contig298_838	0.578	1.129	1.151
6	520	Hapmap25270-BTA-142450	BTB-00203584	11	0.87	6.99	4_97	ARS-BFGL-NGS-103036	0.570	1.122	1.105
7	263	Hapmap48190-BTA-114376	BTA-48503-no-rs	12	0.84	7.83	2_102	BTA-92857-no-rs	0.324	0.974	0.642
8	1439	Hapmap48542-BTA-97857	ARS-BFGL-NGS-27497	11	0.82	8.65	13_14	Hapmap44228-BTA-34185	0.566	1.119	1.108
9	1247	Hapmap43962-BTA-86597	Hapmap42711-BTA-87541	24	0.79	9.45	11_21	Hapmap43962-BTA-86597	0.504	1.071	0.956
10	1963	ARS-BFGL-NGS-119228	ARS-BFGL-NGS-27079	22	0.79	10.23	19_61	BTA-123114-no-rs	0.336	0.971	0.712
11	731	Hapmap49442-BTA-111073	ARS-BFGL-NGS-28491	19	0.72	10.96	6_64	Hapmap40687-BTA-21304	0.310	0.974	0.594
12	1140	ARS-BFGL-NGS-94247	ARS-BFGL-NGS-32828 Hapmap36784-	24	0.63	11.59	10_18	ARS-BFGL-NGS-107048	0.453	1.036	0.903
13	2502	BTB-02076946	SCAFFOLD25027_24456	19	0.53	12.12	29_24	ARS-BFGL-NGS-109817	0.274	0.96	0.513
14	2239	ARS-BFGL-NGS-80841	BTA-34314-no-rs	10	0.49	12.61	24_13	BTA-34314-no-rs	0.366	0.994	0.654
15	1013	ARS-BFGL-NGS-107896	BTB-00377449	11	0.46	13.07	8_113	ARS-BFGL-NGS-106379	0.394	1.008	0.699
16	1631	ARS-BFGL-BAC-27751	ARS-BFGL-NGS-115263 Hapmap30034-BTA-	16	0.43	13.5	15_37	ARS-BFGL-NGS-5811	0.367	0.994	0.671
17	1739	ARS-BFGL-NGS-111532	154964 Hapmap34915-	15	0.42	13.92	16_60	Hapmap50348-BTA-39557 Hapmap34915-	0.152	0.927	0.235
18	2451	BTB-00980670	BES7_Contig278_1082	17	0.41	14.34	28_20	BES7_Contig278_1082	0.198	0.938	0.341
19	1714	Hapmap47935-BTA-38630	BTA-38698-no-rs	16	0.37	14.71	16_35	Hapmap54735-ss46526095	0.188	0.935	0.317
20	2305	ARS-BFGL-NGS-26313	ARS-BFGL-NGS-34801	18	0.36	15.07	25_15	ARS-BFGL-NGS-84660	0.206	0.943	0.327
21	851	ARS-BFGL-NGS-109962	BTB-00316348 Hapmap48206-BTA-	21	0.34	15.41	7_64	ARS-BFGL-NGS-109962	0.182	0.929	0.316
22	712	Hapmap30420-BTC-039335	119876	19	0.33	15.73	6_45	Hapmap30420-BTC-039335	0.120	0.916	0.184

Tabela 6. Continuação

Obs	Window	Nome SNP Inicial	Nome SNP Final	N SNP	%Var	Cum Var	chr_Mb	Nome TOP SNP	ModelFreq	t.like	efeito padrão
23	1825	Hapmap24163-BTA-160726	ARS-BFGL-BAC-36625	18	0.32	16.05	17_64	ARS-BFGL-NGS-50172	0.202	0.937	0.347
24	716	BTB-01688071	ARS-BFGL-NGS-58275	23	0.32	16.37	6_49	BTB-02002785	0.198	0.939	0.331
25	1667	Hapmap42191-BTA-37608	ARS-BFGL-NGS-119332	11	0.31	16.68	15_73	Hapmap40672-BTA-121035	0.279	0.961	0.501
26	812	BTB-00299459	Hapmap59434-rs29012267	12	0.31	16.98	7_25	Hapmap42675-BTA-78561	0.208	0.939	0.350
27	1995	BTA-50190-no-rs	BTB-01328740	13	0.29	17.27	20_28	Hapmap53513-rs29024419	0.208	0.938	0.331
28	2082	ARS-BFGL-NGS-40594	BTB-01924219	19	0.28	17.55	21_43	ARS-BFGL-NGS-105547	0.259	0.948	0.516
29	1287	Hapmap39736-BTA-26188	BTA-11701-rs29017459	12	0.28	17.84	11_61	ARS-BFGL-NGS-52541	0.154	0.922	0.257
30	1198	ARS-BFGL-NGS-38839	UA-IFASA-1616	15	0.28	18.12	10_77	ARS-BFGL-NGS-111871	0.285	0.964	0.493
31	1755	ARS-BFGL-NGS-119489	ARS-BFGL-NGS-49089	24	0.27	18.39	16_76	ARS-BFGL-NGS-37846	0.212	0.94	0.349
32	1628	ARS-BFGL-NGS-22671	Hapmap49972-BTA-36466	19	0.25	18.64	15_34	ARS-BFGL-NGS-111085	0.180	0.927	0.314
33	1284	ARS-BFGL-NGS-15285	Hapmap32152-BTA-29315	15	0.25	18.88	11_58	BTA-33625-no-rs	0.158	0.926	0.268
34	1942	BTB-00752956	BTA-45471-no-rs	11	0.25	19.13	19_40	BTB-00752956	0.244	0.946	0.450
35	2446	ARS-BFGL-NGS-112617	BTB-01878683	22	0.23	19.36	28_15	BTB-00979245	0.165	0.924	0.291
36	2321	Hapmap30681-BTC-028325	ARS-BFGL-NGS-8024	30	0.23	19.59	25_31	Hapmap26309-BTC-057772	0.180	0.935	0.284
37	1563	ARS-BFGL-NGS-115527	ARS-BFGL-NGS-1112	14	0.23	19.82	14_54	BTB-00915241	0.166	0.926	0.254
38	1440	BTA-34170-no-rs	ARS-BFGL-NGS-16065	15	0.22	20.04	13_15	ARS-BFGL-NGS-113489	0.203	0.94	0.324
39	1713	Hapmap41467-BTA-18750	BTB-00635702	21	0.22	20.27	16_34	BTB-00636278	0.094	0.896	0.148
40	1377	Hapmap55276-rs29021589	Hapmap33524-BTA-21440	9	0.22	20.49	12_43	Hapmap52274-rs29017133	0.185	0.925	0.334
41	622	BTA-74159-no-rs	BTA-74084-no-rs	10	0.21	20.7	5_78	ARS-BFGL-NGS-11787	0.209	0.922	0.428
42	1229	ARS-BFGL-NGS-31804	ARS-BFGL-NGS-18450	20	0.21	20.9	11_3	ARS-BFGL-NGS-18450	0.211	0.939	0.344
43	883	Hapmap33036-BTA-149699	ARS-BFGL-NGS-15731	13	0.2	21.1	7_96	Hapmap55378-rs29025271	0.225	0.943	0.390
44	24	UA-IFASA-5116	BTA-49319-no-rs	22	0.2	21.3	1_24	BTB-00008277	0.099	0.91	0.131
45	2280	BTA-117811-no-rs	ARS-BFGL-NGS-85334	10	0.2	21.49	24_54	BTA-117811-no-rs	0.171	0.932	0.273

Seleção de tag SNPs

Com o propósito de selecionar SNP mais informativos, evitando redundância através do LD, que estão ligados aos QTL relacionados à resistência aos carrapatos das janelas superiores e que tivesse uma capacidade preditiva adequada para formar um painel de baixa densidade, os parâmetros frequência do modelo (model frequency - MF) e a estatística t-like (TL) fornecida pelo Gensel foram utilizadas. Esta estratégia favorece os marcadores que foram mais frequentemente incluídos no modelo Bayesiano (maior MF) e com um efeito mais consistente (maior TL).

No primeiro passo, utilizando um valor mínimo de MF de 0,0939 foram selecionados 63 SNPs e, no segundo passo, através do valor mínimo de TL de 0,89, foram selecionados 69 SNPs. Após pré-selecionar com base no MF e TL, os 69 SNPs foram subsequentemente analisados em termos de LD com outro SNP selecionado e apresentarem menor MAF, o que resultou em um painel de 61 SNP selecionados como os mais representativos para a característica resistência aos carrapatos para o conjunto de dados contendo animais das duas raças Braford_Hereford. Esses SNPs selecionados foram distribuídos na maioria dos cromossomos bovinos, exceto BTA3, 9, 18, 22, 23, 26 e 27. É importante mencionar que nove dos 61 SNPs estavam localizados no BTA16. Mota et al (2017) em estudo de associação genômica com análise de normas de reação com a mesma população (BO_HH), reportaram a região cromossômica BTA16 associada a resistência aos carrapatos.

Para animais do conjunto de dados Braford, a seleção de SNP utilizando o limiar de MF mínimo de 0,1115, resultou em 39 SNPs e após a segunda seleção pelo valor mínimo de TL de 0.911, 50 SNPs foram selecionados. Após pré-selecionar com base nos parâmetros MF e TL, os 50 SNPs foram subsequentemente analisados em termos de LD e menor MAF, o que resultou em um painel de 44 SNPs selecionados como os mais representativos para a característica resistência aos carrapatos para o conjunto de dados Braford. Esses SNPs selecionados foram distribuídos na maioria dos cromossomos bovinos, exceto BTA7, 8, 9, 12, 22 e 27. Dos 44 SNP, quatro estavam localizados no BTA4 e BTA20. Apesar de haver relacionamento da região BTA20 com resistência aos

carrapatos, Gonda et al. (2007) encontrou a presença de um QTL nessa região ligada a infecção por *Mycobacterium paratuberculosis* em bovinos.

Com o menor valor mínimo de MF de 0,0175 foram selecionados 20 SNPs e, posteriormente ao valor mínimo de TL de 0,502, 52 SNPs. Do total de 52 SNPs selecionados pelos dois parâmetros (MF e LD), 17 SNPs foram excluídos por apresentarem LD com outro SNP e menor MAF, o que resultou em um painel de 35 SNPs selecionados mais representativos para a característica em estudo, para animais do conjunto de dados Hereford. Esses SNPs selecionados foram distribuídos nos cromossomos BTA1, 10, 11, 16, 24 e 25. É importante mencionar que dos 35 SNPs, onze estavam localizados no BTA25. Apesar de não encontrar associação da região BTA25 com resistência aos carrapatos, Höglund et al. (2015) encontrou SNP significativo associado a característica de fertilidade em fêmeas em bovinos da raça Jersey nessa mesma localização.

Geralmente, SNP com alto valores de MF nas janelas superiores estão potencialmente associados com o fenótipo em estudo (Zare et al., 2014). Considerando todos os SNPs com maior MF dentro de cada uma das principais janelas, de acordo com a análise BayesB (Tabela 4, 5 e 6), os valores médios de MF para a característica contagem de carrapatos foram $0,26 \pm 0,15$, $0,02 \pm 0,01$ e $0,25 \pm 0,17$ para os três conjunto de dados Braford, Hereford e Braford_Hereford, respectivamente. Este resultado indica que a maioria dos SNPs explicou uma pequena proporção da variância fenotípica da característica. Resultado semelhante foram relatados por Porto Neto et al. (2011), que concluíram que programas de seleção que usa SNP, teria que ser como um painel de marcadores, em vez de único marcador com alto valor preditivo.

A correlação entre MF e TL foram de moderada a alta (0,50, 0,71 e 0,53 para Braford, Hereford e Braford_Hereford, respectivamente). Apesar da correlação do TL ser incompleta com MF e o TL ser uma medida alternativa do efeito do SNP, combinamos ambos os parâmetros, MF e TL, para selecionar SNPs informativos

A mesma estratégia de seleção de SNPs foi aplicada aos grupos formados para validação cruzada. A partir dos métodos de agrupamento, foram realizados 6, 8 e 10 GWAS dependendo do número de grupos para cada conjunto de dados, resultado em pequenas listas de SNPs (sub-listas) (Tabela 7, 8 e 9).

Capacidade preditiva dos painéis de SNP selecionados

A herdabilidade a posteriori estimada (h^2), para o conjunto de dados Braford_Hereford, com o método BayesA usando o painel tag SNP selecionados para cada grupo (k-médias e aleatório) variou de 0,11 a 0,16 e foi similar as h^2 estimadas com o BayesB ($\pi= 0.99$) como todos os 39.992 marcadores (Tabela 9). O número de SNPs selecionados variou de 45 até 89 em cada grupo. Com o mesmo conjunto de dados (Braford_Hereford). Sollero et al. (2017) utilizando a mesma população do presente estudo, reportaram valores de herdabilidade para o painel tag SNP com o método BayesA, para os grupos (k-médias e aleatório), semelhante ao BayesB ($\pi= 0.99$) com o painel de 41.045 SNPs.

Tabela 7. Proporção média da variância *a posteriori* explicada pelos marcadores (h^2) usando diferentes métodos bayesianos e número de segmentos cromossômicos e SNPs envolvidos na seleção de painel de densidade muito baixa por grupo de validação cruzada k-médias e aleatória para conjunto de dados com animais da raça Braford.

Grupo	h2			Seleção do painel SNP		
	BayesB ($\pi= 0.99$) ¹	BayesA ¹	BayesA tag	Janelas Top ^a	SNPs Top ^b	Tag SNPs ^c
K-médias 1	0.21	0.31	0.20	49	925	67
K-médias 2	0.21	0.33	0.18	46	860	52
K-médias 3	0.23	0.33	0.25	43	755	74
K-médias 4	0.20	0.31	0.18	49	902	58
Aleatório 1	0.21	0.32	0.18	39	697	54
Aleatório 2	0.22	0.33	0.18	42	750	50
Aleatório 3	0.20	0.31	0.16	40	753	43
Aleatório 4	0.23	0.33	0.19	41	717	53

^a Janelas top representa o número de janelas que explicam acima de 0.2% da variância genética na análise de GWAS usando o método BayesB ($\pi= 0.99$).

^b SNPs top representa o número de SNPs incluídos nesta janela top

^c Tag SNPs representa o número de SNPs selecionados como mais informativos de acordo com os critérios baseado na frequência do modelo e na estatística t.like, desequilíbrio de ligação e menor frequência do alelo

¹ Análise utilizando o painel com 39.922 SNPs

Para o conjunto de dados composto somente de animais da raça Braford, a herdabilidade a posteriori com o painel tag SNP selecionados para cada grupo variou de 0,16 a 0,25 e foram levemente inferiores em comparação ao BayesB ($\pi= 0.99$) com o

painel de 39.922 marcadores (Tabela 7). A quantidade de SNPs selecionados variou de 43 até 74 marcadores em cada grupo. Já para animais somente da raça Hereford a herdabilidade a posteriori do painel tag SNP selecionados para cada grupos variou de 0,03 a 0,07 e foram semelhantes as h^2 utilizando o BayesB ($\pi= 0.99$) com painel de marcadores completo. O número de SNPs selecionados para cada grupo, variou de 4 até 30 (Tabela 8).

Estes resultados demonstram que tanto o painel de muito baixa densidade (tag SNPs) com BayesA quanto o painel de marcadores completo com BayesB, apresenta a mesma ou similar proporção da variância fenotípica explicada pelos SNPs.

Tabela 8. Proporção média da variância *a posteriori* explicada pelos marcadores (h^2) usando diferentes métodos bayesianos e número de segmentos cromossômicos e SNPs envolvidos na seleção de painel de densidade muito baixa por grupo de validação cruzada k-médias e aleatória para o conjunto de dados contendo animais da raça Hereford.

Grupo	h^2			Seleção do painel SNP		
	BayesB ($\pi= 0.99$) ¹	BayesA ¹	BayesA tag	Janelas Top ^a	SNPs Top ^b	Tag SNPs ^c
k-médias 1	0.05	0.05	0.03	3	56	5
k-médias 2	0.06	0.07	0.06	10	206	18
k-médias 3	0.05	0.06	0.06	3	55	4
Aleatório 1	0.06	0.07	0.06	7	126	10
Aleatório 2	0.06	0.06	0.07	8	171	17
Aleatório 3	0.05	0.06	0.04	6	88	30

^a Janelas top representa o número de janelas que explicam acima de 0.2% da variância genética na análise de GWAS usando o método BayesB ($\pi= 0.99$).

^b SNPs top representa o número de SNPs incluídos nesta janela top

^c Tag SNPs representa o número de SNPs selecionados como mais informativos de acordo com os critérios baseado na frequência do modelo e na estatística t.like, desequilíbrio de ligação e menor frequência do alelo.

¹ Análise utilizando o painel com 39.922 SNPs

Tabela 9. Proporção média da variância *a posteriori* explicada pelos marcadores (h^2) usando diferentes métodos bayesianos e número de segmentos cromossômicos e SNPs envolvidos na seleção de painel de densidade muito baixa por grupo de validação cruzada k-médias e aleatória para o conjunto de dados contendo as duas raças Braford_Hereford.

Grupo	h^2			Seleção do painel SNP		
	BayesB ($\pi= 0.99$) ¹	BayesA ¹	BayesA tag	Janelas Top ^a	SNPs Top ^b	Tag SNPs ^c
k-médias 1	0.16	0.14	0.14	51	893	63
k-médias 2	0.12	0.12	0.11	36	649	45
k-médias 3	0.16	0.14	0.15	50	885	89
k-médias 4	0.14	0.14	0.13	45	788	48
k-médias 5	0.16	0.14	0.13	48	884	54
Aleatório 1	0.14	0.13	0.14	49	854	68
Aleatório 2	0.16	0.14	0.16	59	1004	78
Aleatório 3	0.14	0.13	0.12	45	772	51
Aleatório 4	0.15	0.14	0.13	38	694	49
Aleatório 5	0.14	0.13	0.14	51	869	70

^a Janelas top representa o número de janelas que explicam acima de 0.2% da variância genética na análise de GWAS usando o método BayesB ($\pi= 0.99$).

^b SNPs top representa o número de SNPs incluídos nesta janela top

^c Tag SNPs representa o número de SNPs selecionados como mais informativos de acordo com os critérios baseado na frequência do modelo e na estatística t.like, desequilíbrio de ligação e menor frequência do alelo

¹ Análise utilizando o painel com 39.922 SNPs

Os valores de capacidade preditiva foram estimados pela correlação genética entre o fenótipo transformado corrigido e o valor genético genômico pelo método GBLUP de único passo a partir da validação cruzada com o BayesA (SSBR-A), utilizando os SNPs selecionados (painel tag SNP). Para o conjunto de dados Braford_Hereford, a capacidade preditiva média para o grupo k-médias foi $0,07 \pm 0,05$ e para grupo aleatório foi de $0,15 \pm 0,05$. Para Braford, a capacidade preditiva média para o grupo k-médias foi de $0,09 \pm 0,04$ e para o grupo aleatório foi de $0,10 \pm 0,01$. Já para animais Hereford a capacidade preditiva para k-médias foi $0,05 \pm 0,03$ e para aleatório foi $0,04 \pm 0,05$.

Os valores de capacidade preditiva dentro de cada grupo variaram de 0,03 a 0,14 para Braford, -0,02 a 0,08 para Hereford e 0,01 a 0,21 para Braford_Hereford (Figura 4, 5, 6). Os valores para os três bancos de dados foram maiores para os grupos formados aleatoriamente em relação ao k-médias, exceto para os grupos 1, 3 e 4 para Braford, Hereford e Braford_Hereford, respectivamente (Tabela 4, 5 e 6). Embora todos os grupos aleatórios tivessem o mesmo número de animais e a mesma distância genética dentro e

entre grupos, a capacidade preditiva para o grupo 4 (Braford_Hereford), 1 (Braford) e 3 (Hereford) foram consideravelmente menores (0,10, 0,09 e -0,02, respectivamente) comparado com outros grupos aleatório. As maiores capacidades preditiva foram observadas para o grupo aleatório 3 com valor de 0,21 (Braford_Hereford), grupo k-médias 1 com valor de 0,14 (Braford) e grupo aleatório 1 e k-médias 3 com valor de 0,08 para ambos (Hereford).

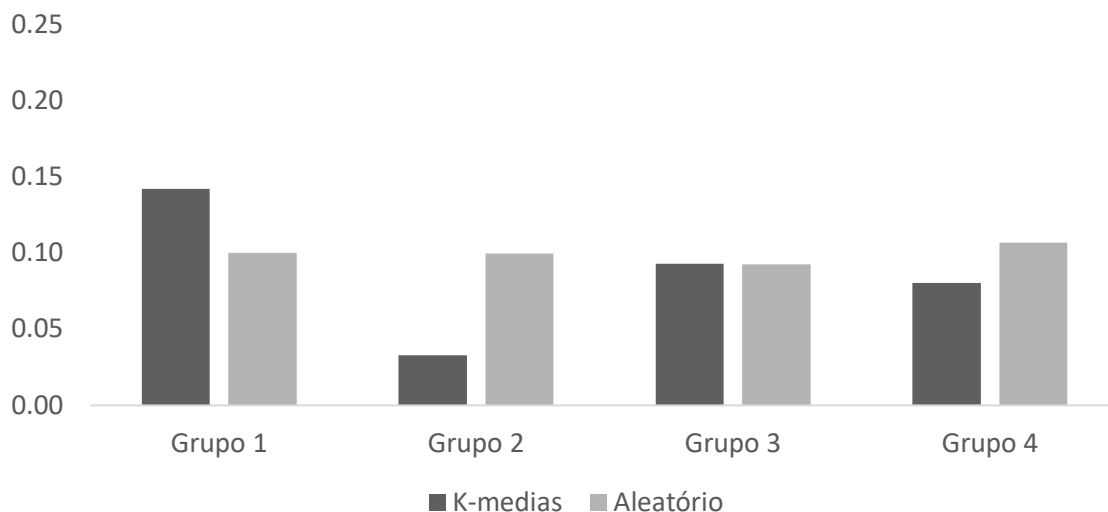


Figura 4. Acurácia dos valores genéticos genômico preditos a partir da validação cruzada para cada agrupamento k-médias e aleatório com o método BayesA com o painel tag SNP para animais do conjunto de dados Braford.

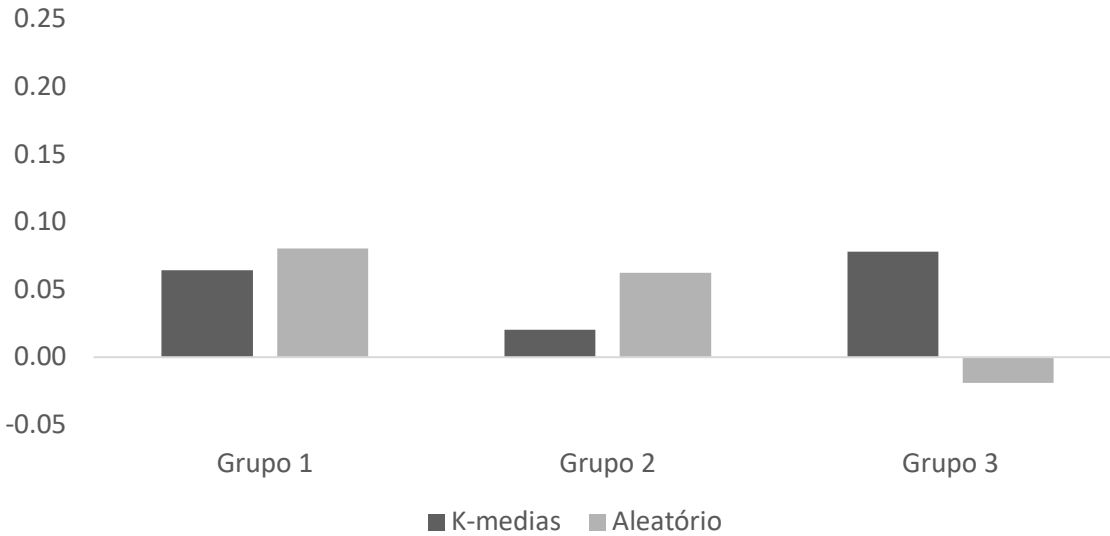


Figura 5. Acurácia dos valores genéticos genômico preditos a partir da validação cruzada para cada agrupamento k-médias e aleatório com o método BayesA com o painel tag SNP para animais do conjunto de dados Hereford.

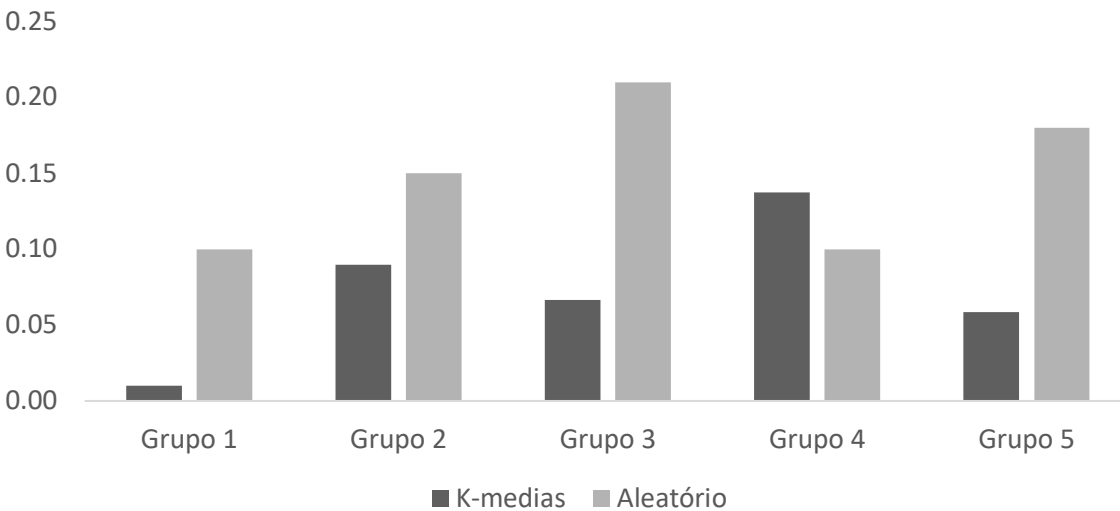


Figura 6. Acurácia dos valores genéticos genômico preditos a partir da validação cruzada para cada agrupamento k-médias e aleatório com o método BayesA com o painel tag SNP para animais do conjunto de dados Braford_Hereford.

Usando o conjunto completo de 39.922 SNPs, a capacidade preditiva média da validação cruzada para agrupamentos k-médias e aleatório, respectivamente, para BayesB com $\pi = 0.99$ foram $0,15 \pm 0,03$ e $0,23 \pm 0,01$ para conjunto de dados BO, $-0,01$

$\pm 0,07$ e $-0,02 \pm 0,06$ para o conjunto de dados HH e $0,12 \pm 0,09$ e $0,17 \pm 0,02$ para conjunto de dados BO_HH.

Comparando os resultados para BO_HH, utilizando o método SSBR-A, a capacidade preditiva que foi derivada usando o painel tag SNP de baixa densidade com 61 SNP representaria pelo menos 67% da capacidade de predição obtida utilizando todos os 39.922 SNPs. Esse resultado demonstra que o tag SNPs pode ser usado para predição genômica para bovinos de corte, como alternativa de baixo custo em relação aos painéis de alta densidade. Sollero et al., (2017) relataram que as acurácias do painel tag SNPs com 58 SNPs (proposto pelos mesmos autores), utilizando o método BayesB, representaria pelo menos 68% em comparação ao painel com 41,045 SNPs. Em comparação ao presente estudo, a diferença dos valores se deve ao fato que Sollero et al. (2017) obtiveram a acurácia de acordo com a proposta do Legarra et al., (2008) o que provoca um aumento nos valores de acurácia, já que a característica apresenta baixa herdabilidade, além do método diferente utilizado.

Para o conjunto de dados BO, utilizando o método SSBR-A, a capacidade preditiva que foi derivada usando o painel tag SNP com 44 SNP representaria pelo menos 72% da capacidade de predição obtidas utilizando todos os 39.922 SNPs. Já para o conjunto de dados HH, utilizando o método SSSBR-A, a capacidade preditiva que foi derivada usando o tag SNP com 35 SNP representaria pelo mens 28% da capacidade de predição obtidas utilizando todos os 39922 SNPs.

Utilizando a proposta do Legarra et al. (2008) a acurácia para o painel tag SNPs (não apresentado no presente estudo) representaria pelo menos 64% em comparação ao método SSBR-B e SSBR-A. Esse resultado mostra que o método de regressão bayesiana de único passo corrobora com os achados por Sollero et al. (2017).

Com o proposito de minimizar os custos de genotipagem e simultaneamente maximizar os ganhos genéticos e os lucros, a seleção de SNPs mais informativos para formação de painel de muito baixa densidade torna-se uma alternativa na seleção genômica. Os dados de genótipos gerados a menores custos, provenientes de pequenos subconjuntos de SNP altamente informativos, poderiam ser usados para genotipar a maioria dos animais em um rebanho e gerar predições de valor genético genômico com base nos efeitos de SNP estimados a partir de dados de treinamento com painéis de alta

densidade (Habier et al., 2009) e ou se completar esse painel através da imputação dos genótipos (Dassonnevilleat al., 2011).

CONCLUSÕES

O método BayesB demonstrou ser adequado para a seleção de SNP com base nos parâmetros frequência do modelo (model frequency – MF) e t-like (TL). Os painéis tag SNP de baixa densidade propostos incluíram SNP que explicaram parte da variação genética e que estão em regiões cromossômicas que possui influência sobre a característica para resistência aos carrapatos para animais da raça Braford e Hereford.

As predições genômicas obtidas com os painéis tag SNP propostos, usando SSBR-A, apresentaram predição de baixa a moderada para a característica resistência aos carrapatos.

AGRADECIMENTO

À CAPES e ao CNPq pelo fornecimento de bolsa durante o período de doutorado, a Embrapa Pecuária Sul - Bagé e a Conexão Delta G – Bagé.

REFERÊNCIAS

Badke, Y. M.; Bates, R. O.; Ernst, C. W.; Schwab, C.; Steibel, J. P. Estimation of linkage disequilibrium in four US pig breeds. **BMC Genomics**. 2012; 13:24.

Barendse, W., 2007. Assessing tick resistance in a bovine animal for selecting cattle for tick resistance by providing a nucleic acid from the bovine animal and assaying for the occurrence of a single nucleotide polymorphism (SNP). **Patent application** WO2007051248-A1, 1–146.

Biegelmeyer, P.; Gullias-Gomes, C. C.; Caetano, A. R.; Steibel, J. P.; Cardoso, F. F. Linkage disequilibrium, persistence of phase and effective population size estimates in Hereford and Braford cattle. **BMC Genet.** 2016; 17:32.

Cheng, H.; Fernando, R. L.; and Garrick, D. J.; JWAS: Julia implementation of whole-genome analysis software. **Proceedings of the World Congress on Genetics Applied to Livestock Production**, 11.859. Auckland, New Zealand, 2018.

Clayton, D. snpStats: SnpMatrix and XSnpmatrix classes and methods. R Package version 1.26.0. 2015.

Dassonneville, R.; Brondum, R. F.; Druet, T.; Fritz, S.; Guillaume, F.; Guldbrandtsen, B.; Lund, M. S.; Ducrocq, V.; Su, G. Effect of imputing markers from a low-density chip on the reliability of genomic breeding values in Holstein populations. **Journal of Dairy Science**, v. 94, p. 3679–3686, 2011.

Fernando, R. L.; Garrick, D. J. Bayesian methods applied to GWAS. In: *Genome-Wide Association Studies and Genomic Prediction*. **Springer**; p. 237– 274, 2013.

Garrick, D. J.; Taylor, J.F; Fernando, R.L. Deregressing estimated breeding values and weighting information for genomic regression analyses. **Genetics Selection Evolution**, v. 41, p. 55, 2009.

Geweke, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4*, (ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith). **Clarendon Press**, Oxford, UK, 1992.

Gonda, M. G.; Kirkpatrick, B. W.; Shook, G. E.; Collins, M. T. Identification of a QTL on BTA20 affecting susceptibility to *Mycobacterium avium* ssp. *Paratuberculosis* infection in US Holsteins. **Animal Genetics**, 38, 389-396, 2007. doi:10.1111/j.1365-2052.2007.01627.x

Habier, D.; Fernando, R. L.; Dekkers, J. C. M. Genomic selection using lowdensity marker panels. **Genetics**, v. 182, p. 343–353, 2009.

Hanotte, O.; Ronin, Y.; Agaba, M.; Nilsson, P.; Gelhaus, A.; Horstmann, R.; Sugimoto, Y.; Kemp, S.; Gibson, J.; Korol, A.; Soller, M.; Teale, A. Mapping of quantitative trait loci controlling trypanotolerance in a cross of tolerant West African N'Dama and susceptible East African Boran cattle Communicated by James E. Womack, Texas A&M University, College Station, TX, April 22, 2003.

Höglund, J. K.; Guldbrandtsen, B. Mogens S. L.; Sahana, G. Identification of genomic regions associated with female fertility in Danish Jersey using whole genome sequence data. **BMC genetics**. 16:60, 2015.

Machado, M. A.; Azevedo, A. L. S.; Teodoro, R. L.; Pires, M. A.; Peixoto, M. G. C. D.; Freitas, C.; Prata, M. C. A.; Furlong, J.; Silva, M. V. G. B.; Guimarães, S. E. F.; Regitano, L. C. A.; Coutinho, L. L.; Gasparin, G.; Verneque, R. S. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (*Bos taurus*, *Bos indicus*). **BMC Genomics**, 11 (2010), p. 280

Mapholi, N. O.; Maiwashe, A.; Matika, O.; Riggio, V.; Bishop, S. C.; MacNeil, M. D.; et al. Genome-wide association study of tick resistance in South African Nguni cattle. *Ticks Tick Borne Dis.* 2016; 7:487–97.

Meuwissen, T. H. E.; Hayes, B. J.; Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, p. 1819–1829, 2001.

Misztal, I., Reliable computing in estimation of variance components. **Journal of animal breeding and genetics**, v. 125, p. 363-370, 2008.

Misztal, I.; Tsuruta, S.; Strabel, T.; Auvray, B.; Druet, T.; Lee, D. H. BLUPF90 and related programs (BGF90) In: **7th World Congress on Genetics Applied to Livestock Production**, Montpellier, France. CD-ROM. Communication number 28-07, 2002.

MisztaL, I., Computational techniques in animal breeding. University of Georgia, Athens, GA, v., p., 2000.

Mota, R. R., Silva, F. F., Lopes, P. S., Tempelman, R. J., Sollero, B. P., Aguilar, I., & Cardoso, F. F. (2017). Analyses of reaction norms reveal new chromosome regions associated with tick resistance in cattle. **Animal**, 12(02), 205–214.[doi:10.1017/s1751731117001562](https://doi.org/10.1017/s1751731117001562)

Onteru, S. K.; Gorbach, D. M.; Young, J. M.; Garrick, D. J.; Dekkers, J. C. M.; Rothschild, M. F. Whole genome association studies of residual feed intake and related traits in the pig. **PLoS One**. 2013; 8:e61756.

Piccoli, M. L.; Braccini, J.; Cardoso, F. F.; Sargolzaei, M.; Larmer, S. G.; Schenkel, F. S. Accuracy of genome-wide imputation in Braford and Hereford beef cattle. **BMC Genetics**. 2014; 15:157.

Porto-Neto, L. R.; Jonsson, N. N.; D’Occhio, M. J.; Barendse, W. Molecular genetic approaches for identifying the basis of variation in resistance to tick infestation in cattle. **Vet. Parasitol.**, 180 (2011), pp. 165-172

Porto-Neto, L. R.; Bunch, R. J.; Harrison, B. E.; Prayaga, K. C.; Barendse, W. Haplotypes that include the integrin alpha 11 gene are associated with tick burden in cattle. **BMC Genetics**, 11 (2010), p. 55.

Saatchi, M.; McClure, S. D.; McKay, S. D.; Rolf, M. M.; et al. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. **Genetics Selection Evolution**. 2011; 43:40.

Sargolzaei, M.; Chesnais, J. P.; Schenkel, F. S. FImpute – An efficient imputation algorithm for dairy cattle populations. **Journal of Animal Science**, v. 89, p. 333, 2011.

Schurink, A.; Wolc, A.; Ducro, B. J.; Frankena, K.; Garrick, D. J.; Dekkers, J. C. M.; et al. Genome-wide association study of insect bite hypersensitivity in two horse populations in the Netherlands **Genetics Selection Evolution**. 2012; 44:31.

Smith, B. J. boa: an R package for MCMC output convergence assessment and posterior inference. **Journal Statistical Software**, v. 21, n. 11, p. 1–37, 2007.

Sollero, B. P.; Junqueira, V. S.; Gomes, C. C. G.; Caetano, A. R.; Cardoso, F. F. Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford clattle breeds using Bayesian methods. **Genetics Selection Evolution**, 49:40, 2017.

Sollero, B. P.; Gomes, C. C. G.; Roso, V. M.; Higa, R. H.; Yokoo, M. J.; Cardoso, L. L.; Caetano, A. R.; Cardoso, F. F. Goodness of fit comparisons among five Bayesian models in genome-wide association of tick resistance in Brazilian Hereford and Braford beef cattle. Proceedings, **10th World Congress of Genetics Applied to Livestock Production**, Vancouver, Canada, 17–22 August (2014)

Sun, X.; Habier, D.; Fernando, R. L.; Garrick, D.; Dekkers, J. C. M. Genomic breeding value prediction and QTL mapping of QTLMAS2010 data using Bayesian Methods. **BMC, Proc** 2011; 5 (Suppl 3):S13.

Turner, L. B.; Harrison, B. E.; Bunch, R. J.; Porto-Neto, L. R.; Barendse, W. A genome-wide association study of tick burden and milk composition in cattle. **Anim. Prod. Sci.**, 50 (2010), pp. 235-245.

Van Eenennaam, A. L.; Weigel, K. A.; Young, A. E.; Cleveland, M. A.; Dekkers, J. C. M. Applied animal genomics: results from the field. **Rev Animal Biosci**. 2014; 2:105–39.

Wolc, A.; Arango, J.; Settar, P.; Fulton, J. E.; O'Sullivan, N. P.; Preisinger, R.; et al. Genome-wide association analysis and genetic architecture of egg weight and egg uniformity in layer chickens. **Animal Genetics**. 2012;43(Suppl 1):87–96.

Zare, Y.; Shook, G. E.; Collins, M. T.; Kirkpatrick, B. W.; Genome-wide association analysis and genomic prediction of *Mycobacterium avium* subspecies paratuberculosis infection in US Jersey cattle. **PLoS One**. 2014; 9:e88380.

Zimin, A. V.; Delcher, A. L.; Florea, L.; Kelley, D. R.; Schatz, M. C.; Puiu, D. et al. A whole-genome assembly of the domestic cow. *Bos taurus*. **Genome Biol**. 2009; 10:R42.

CONCLUSÕES GERAIS

O método single-step ponderado apresentou melhores resultados de capacidade preditiva em relação aos outros métodos single-step. A depender das raças que compõem a estrutura da população, os valores de capacidade preditiva podem variar se as raças são avaliadas conjuntamente ou separadas.

O método BayesB demonstrou ser adequado para a seleção de tag SNPs. Os painéis de baixa densidade (tag SNP) são responsáveis pela maior parte da variação genética na resistência ao carrapato. A capacidade preditiva das predições genômicas derivadas dos painéis SNP de baixa densidade proposto usando SSBR-A pode ser útil para fornecer testes genômicos mais baratos para a indústria.