RODRIGO REIS MOTA

GENOMIC REACTION NORMS FOR TICK RESISTANCE IN HEREFORD AND BRAFORD BEEF CATTLE

Thesis presented to the Universidade Federal de Viçosa as part of the requirements of Genetics and Breeding Graduate Program for the achievement of the title of *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRAZIL
2015

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

T
M917g
2015

Mota, Rodrigo Reis, 1985-
Genomic reaction norms for tick resistance in Hereford and Braford beef cattle / Rodrigo Reis Mota. – Viçosa, MG, 2015.
xii, 86f. : il. ; 29 cm.

Inclui apêndices.
Orientador: Paulo Sávio Lopes.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Bovino de corte. 2. Interação genótipo-ambiente. 3. Variação (Genética). 4. Carrapato . I. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-graduação em Genética e Melhoramento. II. Título.

CDD 22. ed. 636.213

RODRIGO REIS MOTA

GENOMIC REACTION NORMS FOR TICK RESISTANCE IN HEREFORD AND BRAFORD BEEF CATTLE

Thesis presented to the Universidade Federal de Viçosa as part of the requirements of Genetics and Breeding Graduate Program for the achievement of the title of *Doctor Scientiae*.

APPROVED: April 15th, 2015.

Fernando Flores Cardoso
(Co-advisor)

Robert John Tempelman

Fabyano Fonseca e Silva

Cláudio Napolis Costa

Paulo Sávio Lopes
(Advisor)

I dedicate this work to my parents
José Carlos and Célia,
my sister Débora
and my cousin Bárbara (in memoriam).

ACKNOWLEDGMENTS

The Universidade Federal de Viçosa (UFV), especially the Genetics and Breeding Graduate Program and the Department of Animal Science, for the opportunity of carrying out the course;

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support;

Conexão Delta G for providing the data and to have become this work possible;

Embrapa Pecuária Sul, Instituto Nacional de Investigacion Agropecuaria (INIA – Uruguay) and Michigan State University for theoretical and technical support;

Agriculture and Food Research Initiative Competitive Grant no. 2011-67015-30338 from the USDA National Institute of Food and Agriculture for support this work;

Professor Dr. Paulo Sávio Lopes, my advisor, for his valuable supervision, teachings, patience, friendship, advice and opportunities given to me;

My co-advisor Dr. Fernando Flores Cardoso for his supervision, teachings and friendship;

Professor Dr. Robert John Tempelman, my adviser at Michigan State University for his excellent suggestions, support, teachings and friendship;

Professor Dr. Ignacio Garcia Aguilar, for hosting me in Montevideo - Uruguay, his supervision, support, helps, teachings and friendship;

Professor Dr. Fabyano Fonseca e Silva for his teachings and friendship;

Professor Dr. Marcos Deon Vilela de Resende for his helps, teachings and friendship;

Professor Dr. Simone Eliza Facioni Guimarães for her teachings, support and friendship;

Researcher Dr. Cláudio Napolis Costa for his helps and support;

Researcher Dr. Marco Antônio Machado for his helps and support;

Professor Dr. Robledo de Almeida Torres for friendship and teachings;

Professor Dr. Juan Pedro Steibel for his helps, teachings and friendship;

Researcher Dr. Marcos Jun it Yokoo for his helps and support;

Professors, technicians and employees from the Genetics and Breeding Graduate Program and Department of Animal Science;

Professors from Universidade Federal do Espírito Santo – UFES, for their support and teachings;

My family, José Carlos Souza Mota, Célia Maria Reis Mota and Débora Reis Mota for caring, support, comfort and also for believe in me all my life;

My friends from UFV, Ali, Aline, Ana Lúcia, André Hidalgo, André Mauric, Bruno, Carla, Carol, Cris, Débora, Daniele, Edson, Felipe, Giovani, Henrique, Hinayah, Jeferson, Joashlleny, Lais, Leonardo, Lucas, Luciano, Luiz, Marcos Lagrotta, Marcos Lopes, Mariele, Matilde, Nadson, Renata, Rodrigo Pacheco, Rogério and Vinícius, for good times, jokes, pleasant workplace and for their friendship;

My friends from Bagé, Bruno, Ândrea, Rodrigo Azambuja, Rodrigo Costa, Patrícia, Robert, Rafael, Tassiani, Luiza and Gerusa;

My friends from United States Angélica, Ana Paula, Ayra, Carol, Cameron, Cássia, Guilherme, Hunter, João Paulo, Lei, Luiz, Marília, Marina, Miguel, Sandeep, Vinícius, Wesley and especially to Lilian and Yeni for their friendship help and support;

My friends from Alegre-ES, especially from República Vira-latas;

My friends from Viçosa and Ponte Nova for their support and friendship, I am really thankful;

José Geraldo Costa for his friendship and good times together;

My relatives for helping me every time in my whole life;

God to become it possible;

To everyone who directly or indirectly contributed to this work.

BIOGRAPHY

RODRIGO REIS MOTA, son of José Carlos Souza Mota and Célia Maria Reis Mota, was born May 23th, 1985 in Ponte Nova, Minas Gerais, Brazil.

In October 13th, 2004 he started his undergraduate in Animal Science at Centro de Ciências Agrárias Universidade Federal do Espírito Santo (CCAUFES), in Alegre, Espírito Santo. He graduated in August 21th, 2009.

During his undergraduate course period, he was a statistics tutor for three years. Besides he was also tutor in Zoology applied to Animal Science in 2005 and taught some Animal Breeding classes under Professor Dr. Luiz Fernando Aarão Marques supervision in 2008.

In additon, he was volunteer in a undergraduate project titled as “Gain weight test in Espírito Santo and Bahia states” under Professor Dr. Luiz Fernando Aarão Marques supervision from July, 2008 to June, 2009.

He has completed two academic traineeship. In 2008, from February to March, at Universidade Federal de Viçosa (UFV), and from March to May, 2009 at Centro Nacional de Pesquisa de Gado de Leite (EMBRAPA GADO DE LEITE), in Coronel Pacheco-MG, Brazil.

In March, 2010, he started his graduate program at Animal Science Graduate Program by UFV to obtain his degree of *Magister Scientiae* in Animal Science. He presented his dissertation in February 15th, 2012.

In March, 2012, started his PhD at Genetics and Breeding Graduate Program by UFV, to obtain his degree of *Doctor Scientiae* in Genetics and Breeding presenting his thesis in April 15th, 2015.

TABLE OF CONTENTS

RESUMO

MOTA, Rodrigo Reis, P.h.D., Universidade Federal de Viçosa, April, 2015. **Normas de reação genômicas para resistência ao carrapato em bovinos Hereford e Braford via modelos de normas de reação**. Orientador: Paulo Sávio Lopes. Coorientador: Fernando Flores Cardoso.

O “carrapato do boi” é um parasito que causa danos substanciais na produção de bovinos em áreas tropicais. Embora países como o Brasil tenham progredido em avaliações genéticas para a resistência ao carrapato, essas avaliações normalmente não tem considerado a interação genótipo x ambiente (G*A), o que pode afetar diretamente no ganho genético uma vez que a comparação entre os valores genéticos dos animais é dependente do ambiente. O objetivo desse estudo foi investigar a presença de G*A, utilizando modelos com diferentes pressuposições de variância genética e residual. Foram utilizados 10.673 contagens de carrapatos de 4.363 animais Hereford e Braford e um pedigree que continha 11.967 indivíduos. Nove modelos, sendo dois modelos animais tradicionais (MA) e sete modelos hierárquicos de normas de reação (MHNR) foram investigados. Modelos de um passo e dois passos foram usados para inferir sobre a sensibilidade dos valores genéticos ao ambiente via MHNR. O critério de informação da deviance (DIC) foi utilizado como critério estatístico na escolha do melhor modelo. O modelo de melhor ajuste foi o modelo de normas de reação de um passo com 10 classes de variâncias residuais baseados em percentis das estimativas de grupo de contemporâneos utilizadas como gradiente ambiental. Os modelos de normas de reação de um passo apresentaram as maiores estimativas de variância genética. As estimativas de variância do efeito de ambiente permanente foram, em geral, similares entre os modelos testados e variaram de 0,007 a 0,010. As estimativas de correlações genéticas entre o intercepto e a inclinação para ambos os efeitos variaram de baixa a média magnitude e apresentaram altos desvios padrão o que pode ser um indicativo de independência paramétrica. Estimativas de herdabilidades foram maiores para MHNR em comparação com MA. As estimativas de repetibilidade variaram ao longo do gradiente ambiental (de 0,18 a 0,45), o que implica na importância do efeito de ambiente permanente para a característica de resistência ao carrapato. As médias a *posteriori* das

correlações genéticas ao longo do gradiente ambiental apresentaram um grande platô com valores acima de 0,80 para ambientes de baixa infestação de carrapatos. Os MHNR são uma poderosa ferramenta na identificação e quantificação da G*A além de ser uma alternativa promissora para as avaliações genéticas para resistência ao carrapato em bovinos Hereford e Braford, podendo elevar a eficiência de seleção e progresso genético. Melhores respostas à seleção são também esperadas em MHNR que consideram heterogeneidade de variância residual. Em um segundo estudo, foi incorporada a informação de marcador para comparar a eficiência de modelos animais convencionais e modelos de normas de reação utilizando o procedimento de um passo que combina a informação de marcador a de pedigree e também comparar o desempenho das predições de valores genéticos genômicos (GEBV) obtidos utilizando apenas o fenótipo e a informação de pedigree, como também incorporando a informação de marcador. Quatro diferentes modelos foram testados: dois modelos convencionais (BLUP) e dois de normas de reação de um passo (MNR), sendo um BLUP e um MNR com e sem informação de marcador SNP. Os modelos convencionais apresentaram um pior ajuste em comparação com os modelos de normas de reação. O modelo de normas de reação que incluiu a informação de marcador apresentou estimativa de variância genética inferior ao modelo de normas de reação que não a incluía. Estimativas de herdabilidade e repetibilidade foram, em geral, similares em ambos os modelos e variaram ao longo do gradiente ambiental de 0,07 a 0,46 e 0,20 a 0,60, respectivamente. As correlações genéticas foram notoriamente baixas entre ambientes extremos, o que indica a presença de interação genótipo x ambiente (G*A) para a característica de resistência ao carrapato. As predições de acurácias em um estudo de validação cruzada para os modelos testados foram altas e superiores a 0,55 e 0,59 para os procedimentos de partições “K-means” e partições aleatórias, respectivamente. Esses resultados sugerem que a informação de marcador não contribui para o aumento da acurácia de predição em que estas decrescem à medida que a infestação de carrapatos aumenta e, ou a relação de parentesco entre animais na população de referência e da população de validação diminui. Em um terceiro estudo, os objetivos foram:

obter predições de valores genéticos em bovinos Hereford e Braford usando o procedimento de passo único que combina informação de pedigree a de marcador (ssBLUP), estimar os efeitos de marcador das normas de reação associadas com a resistência ao carrapato, bem como identificar genes candidatos derivados dos marcadores SNP mais relevantes. Um modelo de normas de reação de um passo foi ajustado para a estimação dos componentes de (co)variância e parâmetros genéticos. Para estudar os efeitos de marcadores SNP ao longo de diferentes níveis de infestação de carrapato, foram identificados os 1% SNPs mais relevantes em cada nível de infestação de carrapato e apontadas a similaridade entre estes marcadores ao longo dos níveis. Os efeitos genéticos e de ambiente permanente apresentaram significantes inclinações confirmando a presença de G*A. As correlações entre o intercepto e a inclinação foram positivas e de alta (0,52±0,18) e média (0,26±0,15) magnitudes, respectivamente, para os efeitos genéticos e de ambiente permanente. Dos 410 (1%) de SNPs identificados, 75 foram constantemente relevantes em todos os níveis ambientais e indicaram presença de interação SNP x ambiente. Os SNPs mais relevantes estão localizados nos cromossomos 1, 2, 6, 7, 9, 11, 14, 21 e 23 e genes encontrados próximos a esses marcadores apresentaram variadas funções como metabolismo energético, pigmentação do epitélio da retina, integridade e manutenção de células fotorreceptoras e diferenciação celular.

ABSTRACT

MOTA, Rodrigo Reis, P.h.D., Universidade Federal de Viçosa, April, 2015. **Genomic Reaction Norms for tick resistance in Hereford and Braford beef cattle**. Advisor: Paulo Sávio Lopes. Co-advisor: Fernando Flores Cardoso.

The cattle tick is a parasite that adversely affects livestock performance in tropical areas. Although countries such as Australia and Brazil have provided genetic evaluations for tick resistance, these evaluations have not typically considered genotype by environment interaction (G*E); hence genetic gains could be adversely affected as breedstock comparisons are environmentally-dependent in the presence of G*E, particularly if residual variability is also heterogeneous across environments. The objective of this study was to investigate the existence of G*E based on various models with different assumptions on genetic and residual variability. Data were collected by the Delta G Connection improvement program including 10,673 tick count phenotypes on 4,363 animals. Nine models including two traditional animal models (AM) and seven different hierarchical Bayesian reaction norm models (HBRNM) were investigated. One-step and two-step modeling approaches were used to infer upon G*E. Model choice was based on the deviance criterion information (DIC). The best-fitting model specified heterogeneous residual variances across 10 subclasses as delimited by every decile of the contemporary group estimates of tick count effects. One-step models generally had the highest estimated genetic variances. Estimates of heritabilities were generally higher for HBRNM than AM. Furthermore, one-step models based on heterogeneous residual variances also generally lead to higher heritability estimates, especially in harsh environments. Estimates of repeatability varied along the environmental gradient (range 0.18-0.45) implying that the relative importance of additive and permanent environment effects for tick resistance is environmentally influenced. The posterior means of the genetic correlations across environmental tick infestation surface plot demonstrated a large plateau above 0.80. HBRNM represent powerful tools to infer G*E and account for their effects for genetic evaluations of tick resistance. Additional increases in accuracies on estimated breeding values

are also expected based on HBRNM analyses that additionally consider heterogeneity of residual variances across environments. In a second study, we incorporated marker information to compare a conventional genomic-based single step BLUP model with its one-step genomic reaction norm model extension on tick infestation phenotypes and to compare the performance of genomic estimates breeding values (GEBV) predictions obtained from using only phenotypes and phenotypes plus marker information. Four different models were tested: two conventional animal models, and two one-step reaction norm model with and without genomics. The non reaction norm models seem to be poorer fitting in comparison with its one-step extensions. The reaction norm model including marker information presented lower intercept and slope genetic variance estimates in comparison with the models that included the pedigree-based relationship matrix. Heritability and repeatability estimates were, in general, similar for both models and ranged over the environmental gradient (EG) from 0.07 to 0.46 and from 0.20 to 0.60, respectively. Genetic correlations were remarkably low between extreme EG, indicating the presence of G*E for tick resistance. Cross validation estimates were in average 0.66±0.02, 0.67±0.02, 0.67±0.02 and 0.66±0.02 for BLUP, GBLUP, GLRNM and LRNM, respectively, based on K-means partitioning, whereas GLRNM was 0.71±0.01 and tend to better than BLUP (0.67±0.01), GBLUP (0.70±0.01) and LRNM (0.70±0.01) based on random partitioning. However, no statistical significance was reported between GLRNM and LRNM. Our results also suggest that marker information do not lead for higher prediction accuracies which decreased as the tick infestation level increased and as the relationship between animals in training and validation datasets decreased. In third and last study, was aimed to perform genome-enabled predictions for tick resistance in Hereford and Braford cattle by using single step genomic BLUP methodology (ssGBLUP), to estimate marker effects from reaction norms associated with tick resistance as well as to identify candidate genes derived from the most relevant SNP markers. A one-step reaction norm model was fitted to estimate the (co)variance components and genetic parameters. To study SNP effects across different tick infestation (TI) levels, we identified the top 1% of SNPs in each TI and pointed out to the similarity

between these markers across the levels. The additive genetic and permanent environment effects showed significant slope confirming the presence of G*E. Correlations between intercept and slope were positive with high (0.52±0.18) and moderate (0.26±0.15) magnitude for genetic and permanent environment effects, respectively. From the top 1% SNPs (410), 75 were consistently relevant across TI and indicated SNP by environment interaction. The most relevant SNPs were located on chromosomes 1, 2, 6, 7, 9, 11, 14, 21 and 23 and the annotated genes closest these markers showed functions related to energy metabolism, retinal pigment epithelium, maintenance and integrity of the photoreceptor cells, and cell differentiation.

GENERAL INTRODUCTION

Brazil is one of the world's main suppliers of agricultural products. The Brazilian beef cattle production and exports have been rising over time. Brazil has the second largest beef cattle production and is the largest exporter of meat (http://www.brazilianbeef.org.br/).

The country has approximately, 210 million cattle herd in continuous growth and has made progress in productivity rates (http://www.brazilianbeef.org.br/). However, to remain competitive in foreign markets and seek for better prices, Brazil has to improve beef quality and safety standards.

Losses in livestock production due to external parasites have long been a major concern to beef cattle producers in tropical and subtropical regions (Seifert, 1984a;b). Tick-borne disease from *Rhipicephalus* (*Boophilus*) *microplus* is one of the major causes of losses in cattle and has caused decrease in animal performance by hematophagism, devaluation of hide price and high costs for supplies (Grisi et al., 2002).

Tick chemical treatments have a crucial role in the control of this parasite, but can result in meat not free from chemical residues and also jeopardizing food security. Furthermore, losses are aggravated by the transmission of other infectious agents such as *Babesia bovis*, *B. bigemina* and *Anaplasma spp* causing tick fever (Ribeiro, 1989; Regitano et al., 2006), and acaricide resistance (Jonsson and Hope, 2007). Estimates suggest that Brazil has two billion dollars in annual losses because of this tick (Grisi et al., 2002).

In addition, the cattle tick has limited the expansion of production from purebred and crossbreed cattle such as those from British origin. The use of these breeds would be a strategy to improve meat quality through zebu populations in crossbreeding systems and generation of synthetic breeds. The feasibility of this alternative is largely limited by the higher susceptibility of *Bos taurus* infestation by ticks compared to *Bos indicus*, whereas *Bos indicus x Bos taurus* crossbred has intermediate susceptibility (Frisch, 1999).

Repeated use of treatments with acaricides for tick control in susceptible populations has historically generated parasite resistance with large costs in ineffective treatments and investments for new drugs

discovery. Moreover, although bovines are able to develop an immune response after repeated exposure to ticks (Dossa et al., 1996), vaccines development has not been totally successful due to great ability of parasites to modulate the host immune system (Brossard and Wikel, 2004). Thus, the use of vaccines would succeed only in cases in which there was already high resistance and low infestation (Frisch, 1999).

Nevertheless, the existence of genetic variability for tick resistance has been verified implying the feasibility of improvement for this trait (Oliveira et al., 1989; Gomes, 1992; Veríssimo et al., 1997; Andrade et al., 2001; Silva et al., 2006, Oliveira et al., 2012; Sollero et al., 2014). Therefore, knowledge of the level of resistance for each breed, between different breeds and crossbreed individuals becomes an important alternative for improvement of beef cattle production.

Furthermore, the discovery of genes that control tick resistance should increase the selection efficiency because the breedstock selection for tick resistance will reduce the use of acaricides and chemical residues, avoiding commercial, environmental and human health problems. Thus, increasing host resistance has been identified as the only permanent solution (Frisch, 1999).

Traditionally, genetic evaluations for tick resistance has been done using the Best Linear Unbiased Prediction (BLUP) methodology (traditional or Bayesian) without genomics (Cardoso et al., 2006; Budeli et al., 2009; Biegelmeyer, 2012). However, the availability of dense single nucleotide polymorphisms (SNP) and dramatic reduction in the cost of acquiring this information has allowed genome wide association study (GWAS) in economically important quantitative traits and the use of new alternatives of genetic evaluation and selection, such as genomic wide-selection (GWS) that includes the genome wide marker information in the prediction of genomic estimated breeding values (GEBV) of animals (Verbyla et al., 2009).

Genomic wide-selection was introduced by Meuwissen et al. (2001) as a selection strategy based on prediction of GEBV by simultaneously evaluating and summing large numbers of marker effects across the entire genome. Such selection is expected to improve the accuracy of GEBV because the method makes the assumption that some markers from a dense

SNP panel will be in linkage disequilibrium (LD) with large number of multiple quantitative trait loci (QTLs) that control economically important quantitative traits (Hayes et al., 2009). In addition, higher marker density implies higher LD level between markers and QTLs. Thus, higher proportion of genetic variance can be explained by markers effects (Verbyla et al., 2009). Hence, GWS would lead to faster genetic gain in comparison with traditional genetic evaluations with higher prediction accuracy and potential interval generation decrease.

On the other hand, markers effects are not necessarily the same in different studies and also environments. In GWS, markers effects are estimated and used for selection for each population in a specific environment. However, models including genotype by environment interaction (G*E) can also be used to investigate the possibility of higher prediction accuracy for a set of environments (Resende et al., 2013).

Genotype by environment interaction occurs when different genotypes will respond differentially to environmental changes, possibly leading to re-rankings in the performance of animals in different sets of environments (Falconer and Mackay, 1996; Lynch and Walsh, 1998; Kolmodin et al., 2002). It implies that failing to consider G*E in genetic evaluations can affect genetic gains since breedstock comparisons may be environment-dependent. That is, animals identified as top breeders in a certain environment may not be best for other environmental conditions; this issue is further exacerbated if progeny are managed in an environment different from their selected parents (Cardoso and Tempelman, 2012).

Several studies have been performed that part of the variability shown is attributed to different environmental conditions inherent to each study. Fraga et al. (2003), Silva et al. (2010) and Biegelmeyer (2012) have reported that tick resistance can be influenced by several environmental factors such as statistical modeling, (co)variance components and genetic parameters estimation in different populations and/or regions, or even different genetic compositions. In these studies, G*E has been specified by defining the phenotype in different environments as different traits (Falconer, 1952).

Reaction norm models (RNM) may describe G*E by assuming traits change along a set of environmental gradient (Falconer and Mackay, 1996).

It is possible partly because it can accommodate a very large number of environments with few parameters through the use of covariance functions (Kirkpatrick et al., 1990). Several studies have examined the importance of G*E in different traits in beef cattle (Pégolo et al., 2009; Corrêa et al., 2010; Ambrosini et al., 2012). Most of these studies have revealed the presence of G*E leading to re-rankings in the performance of animals in different environments. In addition, combining marker information with reaction norm models is a promising strategy to investigate the presence of G*E and also increase prediction accuracy, leading to faster and larger genetic gains.

Objectives:

The general objectives in this study were to investigate the presence of genotype by environment interaction via reaction norm models applied to tick resistance in Hereford and Braford beef cattle, to verify whether the incorporation of marker information to data would increase prediction accuracy as well as to identify chromosome regions that affect tick resistance in those cattle populations.

The specific objectives were to investigate the nature of G*E based on models with different genetic and residual variance assumptions and different ways for accounting for uncertainty on environmental sensitivities, to develop a routine in Fortran 90 to include the single step BLUP methodology (ssBLUP; Legarra et al., 2009; Misztal et al., 2009; Aguilar et al., 2010a,b) in INTERGEN program developed by (Cardoso, 2013) to obtain solutions to mixed model equations in which the numerator relationship matrix is modified by a known matrix accounting for the genomic information, to compare a conventional genomic-based BLUP model with its one-step genomic reaction norm model extension on tick infestation data, to compare estimates obtained from using non genomic and genomic approaches, to investigate the efficiency of GEBV prediction through cross validation studies, to perform genome-enabled predictions for tick resistance by using ssBLUP, to estimate marker effects from reaction norms associated with tick resistance as well as to identify candidate genes derived from the most relevant SNP markers.

REFERENCES

AGUILAR, I.; MISZTAL, I.; JOHNSON, D.L. et al. 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, 2010a.

AGUILAR, I.; MISZTAL I.; LEGARRA, A. et al. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. **Journal Animal Breeding and Genetics**, v.128, p.422-428, 2010b.

AMBROSINI, D.P.; CARNEIRO, P.L.S.; BRACCINI NETO, J. et al. Genotype × environment interaction for yearling weight in Polled Nellore cattle in Northeast Brazil. **Pesquisa Agropecuária Brasileira**, v.47, p.1489-1495, 2012.

ANDRADE, A.B.F.; ALENCAR, M.M.; RAZOOK, A.G. et al. Genetic and phenotypic parameters for measures of infestation of females of the Caracu cattle breed by cattle tick (B*oophilus microplus*) ***Proceedings*** of 38th Reunião Anual da Sociedade Brasileira de Zootecnia, Piracicaba, Brazil: SBZ, 2001. p.575-577.

BIEGELMEYER, P. **Genetic resistance to *Rhipicephalus* (Boophilus) *microplus* natural and artificial infestation in Hereford and Braford cattle**. 2012, 96p. Dissertation (Master in Science) - Universidade Federal de Pelotas, Pelotas, RS, Brazil.

BROSSARD, M.; WIKEL, S.K. Tick immunobiology. **Parasitology**, v.129, p.S161-S176, 2004.

BUDELI, M.A.; NEPHAWE, K.A.; NORRIS, D. et al. Genetic parameter estimates for tick resistance in Bonsmara cattle. **South African Journal of Animal Science**, v.39, p.321-327, 2009.

CARDOSO, V.; FRIES, L.A.; ROSO, V. M. et al. Estimates of heritability for resistance to *Boophilus microplus* tick evaluated by an alternative method in a commercial Polled Hereford x Nelore population in Brazil. ***Proceedings*** of 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil, 2006. CD-ROM.

CARDOSO, F.F.; TEMPELMAN, R.J. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. **Journal of Animal Science**, v.90, p.2130-2141, 2012.

CARDOSO, F.F. **Application of Bayesian inference in animal breeding using the Intergen program: Manual of version 1.22**. Documents 112, Embrapa South Livestock, Bagé, RS, Brazil, 2013.

CORRÊA, M.B.B.; DIONELLO, N.J.L.; CARDOSO, F.F. Genetic evaluation of Devon cattle using a reaction norms model. **Revista Brasileira de Zootecnia**, v.39, p.128-133, 2010.

DOSSA, S.C.; KAAYA, G.P.; ESSUMAN, S. et al. Acquisition of resistance to the tick Amblyommavariegatum in Boran cattle, *Bos indicus* and the effects of *Trypanosomacongolense* and *Babesia bigemina* on host resistance. **Veterinary Parasitology**, v.62, p.317-330, 1996.

FALCONER, D.S. The problem of environmental and selection. **American Natutalist**, v.86, p.293-298, 1952.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to Quantitative Genetics**. 4 ed. Longman Group Ltd, Harlow, UK, 1996.

FRAGA, A.B.; ALENCAR, M.M.; FIGUEIREDO, L.A. et al. Genetic Analysis of the Infestation of Caracu Female Cattle Breed by Cattle Tick (*Boophilus microplus*). **Revista Brasileira de Zootecnia**, v.32, p.1578-1586, 2003.

FRISCH, J.E. Towards a permanent solution for controlling cattle ticks. **International Journal for Parasitology**, v.29, p.57-71, 1999.

GOMES, A. **Resistência à infestação natural por larvas, ninfas e adultos de *Boophilus microplus* em vacas zebuínas da raça Gir, em função de sua idade, da gestação, da lactação e da seleção para produção leiteira, com e sem tratamento carrapaticida, ao longo de 12 estações consecutivas de um triênio**. 1992. 90p. Thesis (PhD in Science) – Universidade de São Paulo, São Paulo, Brazil.

GRISI, L.; MASSARD, L.C.; BORJA, G.E.M. et al. Impacto econômico das principais ectoparasitoses em bovinos no Brasil. **A Hora Veterinária**, Porto Alegre, v.125, p.8-10, 2002.

HAYES, B.J.; BOWMAN, P.J.; CHAMBERLAIN, A.J. et al. Invited review: Genomic selection in dairy cattle: progress and challenges. **Journal of Dairy Science**, v.92, p.433-443, 2009.

JONSSON, N.N.; HOPE, M. Progress in the epidemiology and diagnosis of amitraz resistance in the cattle tick *Boophilus microplus*. **Veterinary Parasitology**, v.146, p.193-198, 2007.

KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, p.979-993, 1990.

KOLMODIN, R.; STRANDBERG, E.; MADSEN, P. et al. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. **Acta Agriculturae Scandinava Section A Animal Science**, v.52, p.11-24, 2002.

LEGARRA, A.; AGUILAR, I.; MISZTAL, I. A relationship matrix including full pedigree and genomic information. **Journal of Dairy Science**, v.92, p.4656-4663, 2009.

LYNCH, M.; WALSH, B. **Genetics and Analysis of Quantitative Traits**. 1st ed. Sinauer Associates, Inc., Sunderland, MA, 1998.

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.; 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.

OLIVEIRA, G.P.; ALENCAR, M.M.; FREITAS, A.R. Resistência de bovinos ao carrapato Boophilus microplus. II. Infestação Natural. **Pesquisa Agropecuária Brasileira**, v.24, p.1267-1271, 1989.

OLIVEIRA, M.M.; GOMES, C.C.G.; ROSO, V.M. et al. Genomic selection for tick resistance in Braford and Hereford cattle ***Proceedings*** of 9th Simpósio Brasileiro de Melhoramento Animal, João Pessoa, Brazil: SBMA, 2012.

PÉGOLO, N.T.; OLIVEIRA, H.N.; ALBUQUERQUE, L.G. et al. Genotype by environment interaction for 450-day weight of Nelore cattle analyzed by reaction norm models. **Genectics and Molecular Biology**, v.32, p.281-287, 2009.

REGITANO, L.C.A.; OLIVEIRA, M.C.S.; ALENCAR, M.M. et al. Avaliação da resistência de bovinos de diferentes grupos genéticos ao carrapato e à babesiose. **Boletim de Pesquisa e Desenvolvimento,** Embrapa Pecuária Sudeste, n.9, 2006.

RESENDE, M.D.V.; SILVA, F.F.; LOPES, P.S. et al. **Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência Bayesiana (MCMC), Regressão Aleatória Multivariada e Estatística Espacial**. Universidade Federal de Viçosa, Viçosa, 2013, 291p.

RIBEIRO, J.M.C. Role of saliva in tick/host interactions. **Experimental Applied Acarology**, v.7, p.15-20, 1989.

SEIFERT, G.W. Research and practical experience in selection for resistance to the cattle tick (*Boophilus microplus*) and gastrointestinal helminths in northern Australian beef cattle. ***Proceedings*** of 2nd World Congress of Sheep and Beef Cattle Breeding. Pretoria, South Africa. p.149-159, 1984a.

SEIFERT, G.W. Selection of beef cattle in Northern Australia for resistance to the cattle tick (*Boophilus microplus*): Research and application. **Preventative Veterinary Medicine**, v.2, p.553-558, 1984b.

SILVA, A.M.; ALENCAR, M.M.; REGITANO, L.C.A. et al. Heritability and repeatability estimates for degree of infestation by External parasites in females of four beef cattle genetic groups. ***Proceedings*** of 43th Reunião Anual da Sociedade Brasileira De Zootecnia, João Pessoa, 2006.

SILVA, A.M.; ALENCAR, M.M.; REGITANO, L.C.A. et al. Natural infestation by external parasites in beef cattle females in southern Brazil. **Revista Brasileira de Zootecnia**, v.39, p.1477-1482, 2010.

SOLLERO, B.P.; GOMES, C.C.G.; ROSO, V.M. et al. Goodness of Fit Comparisons among Five Bayesian Models in Genome-Wide Association of Tick Resistance in Brazilian Hereford and Braford Beef Cattle. ***Proceedings*** of 10th World Congress on Genetics Applied to Livestock Production, Vancouver, Canada, 2014.

VERBYLA, K.L.; HAYES, B.J.; BOWMAN, P.J. et al. Accuracy of genomic selection using stochastic search variable selection in Australian Holstein Friesian dairy cattle. **Genetics Research**, v.91, p.307-311, 2009.

VERÍSSIMO, C.J.; SILVA, R.G.; OLIVEIRA, A.A.D. et al. Resistance and susceptibility of a crossbred dairy herd to the cattle tick *Boophilus microplus*. **Boletim de Indústria Animal**, v.54. p.1-10, 1997.

Chapter 1

GENOTYPE BY ENVIRONMENT INTERACTION IN TICK RESISTANCE OF HEREFORD AND BRAFORD BEEF CATTLE VIA REACTION NORM MODELS

Abstract: The cattle tick is a parasite that adversely affects livestock performance in tropical areas. Although countries such as Australia and Brazil have provided genetic evaluations for tick resistance, these evaluations have not typically considered genotype by environment interaction (G*E); hence genetic gains could be adversely affected as breedstock comparisons are environmentally-dependent in the presence of G*E, particularly if residual variability is also heterogeneous across environments. The objective of this study was to investigate the existence of G*E based on various models with different assumptions on genetic and residual variability. Data were collected by the Delta G Connection improvement program including 10,673 tick count phenotypes on 4,363 animals. Nine models including two traditional animal models (AM) and seven different hierarchical Bayesian reaction norm models (HBRNM) were investigated. One-step and two-step modeling approaches were used to infer upon G*E. Model choice was based on the deviance information criterion (DIC). The best-fitting model specified heterogeneous residual variances across 10 subclasses as delimited by every decile of the contemporary group estimates of tick count effects. One-step models generally had the highest estimated genetic variances. Estimates of heritabilities were generally higher for HBRNM than AM. Furthermore, one-step models based on heterogeneous residual variances also generally lead to higher heritability estimates, especially in harsh environments. Estimates of repeatability varied along the environmental gradient (range 0.18-0.45) implying that the relative importance of additive and permanent environment effects for tick resistance is environmentally influenced. The posterior means of the genetic correlations across environmental tick infestation surface plot demonstrated a large plateau above 0.80. HBRNM represent powerful tools to infer G*E and account for their effects for genetic evaluations of tick resistance. Additional

increases in accuracies on estimated breeding values are also expected based on HBRNM analyses that further consider heterogeneity of residual variances across environments.
Keywords: Bayesian inference, beef cattle, breeding value, heritability, residual variance.

INTRODUCTION

The cattle tick is a parasite adversely affecting beef cattle production in tropical areas, such as Brazil. Retail beef markets are imposing greater restrictions such as, for example, on meat that is free of chemical residues that are perceived to negatively impact the environment, health and human welfare. Hence, to remain competitive in foreign beef markets, Brazil must aspire to some of these higher standards.

To assure market competitiveness, one strategy might be to increase the breed composition of *Bos taurus* in Brazilian herds since such breeds seem to be more productive (Ayres et al., 2013); furthermore, they improve meat quality traits, sexual precocity and culling. However, *Bos taurus* breeds tend to have greater susceptibility to tick infestation compared to *Bos indicus* breeds (Prayaga et al., 2005; Silva et al., 2010). Hence, selection of animals for tick resistance may be useful to reduce chemical control while increasing productivity.

Some evidence for additive genetic variability of tick resistance in cattle (Burrow et al., 2001; Prayaga et al., 2005; Budeli et al., 2009; Ibelli et al., 2012) have been based on reported heritabilities ranging between 0.05 and 0.42. Genetic evaluations for tick resistance are routinely provided in countries such as Australia that has a similiar climate to Brazil and where the cattle tick is also present. Examples of such evaluations include breeds such as Adaptator, Belmont Red, AFS (*Australian Friesian Sahiwal*) and AIS (*Australian Illawarra Shorthorn*). In Brazil, the Conexão Delta G (Delta G Connection) company has had a genetic improvement program based on selection for tick resistance in Hereford and Braford cattle since 2003.

These and other research studies and genetic evaluations have typically not considered genotype by environment interaction (G*E). Failing to consider G*E in genetic evaluations can adversely affect breeding programs if relative genetic merit is environmentally affected (Falconer and Mackay, 1996; Lynch and Wlash, 1998; Cardoso and Tempelman, 2012). That is, animals identified as top breeders in a certain environment may not be ideal in other environments; this issue is further exacerbated if progeny are managed in environments different from that of their selected parents (Côrrea

et al., 2009). Additionally, most current genetic evaluation systems assume homogeneous residual variances across environments, even though investigators have reported evidence of residual heteroskedascity, defined as heterogeneity of residual variances across contemporary groups, for different traits such as milk yield (Torres et al., 2000) and post-weaning gain (Carvalheiro et al., 2002; Cardoso and Tempelman, 2012). Linear reaction norm models capture a simple form of G*E. This modeling approach is based on the use of covariance functions (Kirkpatrick et al., 1990), allowing for the prediction of relative genetic merit of animals as a function of gradual linear changes in an environmental covariate. Sometimes this environmental covariate is not known with certainty and needs to be estimated; Su et al. (2006) demonstrated how this inference uncertainty could be formally accounted for using Bayesian methods. If G*E is important for tick resistance, reaction norm models could be used to fine-tune genetic improvement for tick resistance in Brazilian Hereford and Braford beef cattle. However, if residual variability is also heterogeneous across environments, then reaction norm inferences on G*E could still be adversely affected if this phenomena is not accounted for since G*E itself contributes to heterogeneous genetic variability across environments.

The objective of this study was to investigate the nature of G*E based on models with different assumptions regarding the nature of genetic and residual variation and different ways for accounting for uncertainty on environmental covariates or sensitivities.

MATERIAL AND METHODS

Tick counts data

Data used in this current study were obtained from a breeding program conducted by Conexão Delta G (Delta G Connection). Data included records of tick counts (TC) on Hereford and Braford beef cattle from herds raised in the Rio Grande do Sul state, Brazil. TC were obtained on each animal from 326 to 729 days of age using the method described by Wharton and Utech (1970), whereby all female ticks ≥ 4.5 mm were counted on the entire left side of the animal. Up to 3 such counts were acquired from each animal with each such count separated by a minimum of 30 days similar as in

other studies (Frisch and O'Neill, 1998; Budeli et al., 2009; Biegelmeyer, 2012). The distribution of the number of measurements taken per animal was 241, 1,934 and 2,188 animals having one, two and three TC measurements, respectively, for a total of 10,673 records. The average age during the evaluation period was 524 ± 65 days and the overall mean TC was 34.99 with a standard deviation of 42.15 (range 0- 532).

The 4,363 animals having records were born between 2008 and 2011 and originated from 197 sires and 3,966 dams with 10 generations of pedigree depth. A total of 11,967 animals remained after pedigree pruning (i.e. removing any ancestors occurring only once in the pedigree file). Furthermore, pedigree information was sparse due to use of multiple-sire matings, which resulted in 65% of the animals with TC observations having unknown paternity. A detailed breakdown of the pedigree structure is provided in Table 1.

Table 1 Pedigree structure as defined by parentage certainty

	With records	**Without records**	**Total**
Both parents known	1,515	1,917	3,432
Both parents unknown	4	3,807	3,811
Only sire known	7	4	11
Only dam known	2,837	1,876	4,713
Total	4,363	7,604	**11,967**

Because TC were not typically normally distributed (Figure 1), the log transformation was used such that LTTC = $\log_{10}$ (TC+1.001) was the response variable (Ayres et al., 2013; Biegelmeyer, 2012). The constant 1.001 was included in this transformation as some of the TC were equal to zero (Ayres et al., 2013; Biegelmeyer, 2012). Note that data transformation does not guarantee that the log transformed data are normal, but might improve properties of estimates, predictions and inferences (Budeli et al., 2009).

A contemporary group (CG) was defined as a group of animals being within the same herd, year of birth, season of birth (April to July; August to

November and December to March), of the same sex and from the same management group. Each CG was required to have at least five animals and with each LTTC record being within 3.5 Standard Deviation (SD) from their respective CG means. Moreover, connectedness among CG was assessed by the AMC software (Roso and Schenkel, 2006) such that CG with less than 10 genetic links, were removed. Finally, solutions for CG effects on LTTC were assumed to define the environmental covariates for a linear reaction norm model because they seemed to be the most appropriate entities to describe environmental conditions most important for beef cattle production (Cardoso et al., 2011; Mattar et al., 2011; Cardoso and Tempelman, 2012).

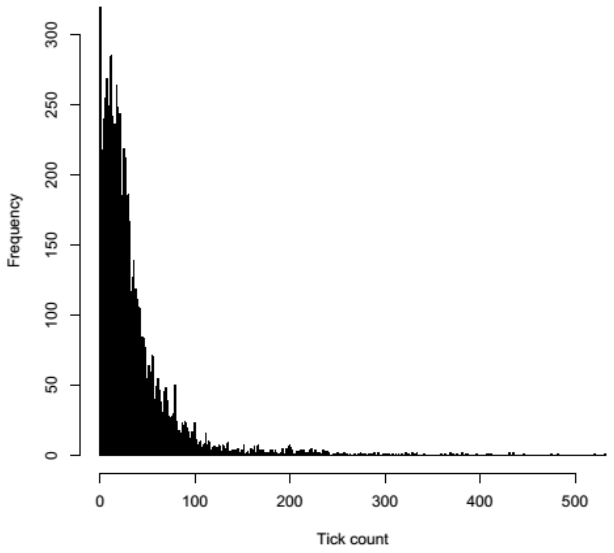


Figure 1 Distribution of tick counts

Statistical models

Nine different analyses based on different models and/or inferential methodologies and specifications on residual variability were conducted on the data: These analyses are described further below as M_1-M_9.

Traditional animal model (AM)

Consider the following simple linear traditional animal model (M_1):

$$y_{ijk} = \mathbf{x}'_j \boldsymbol{\beta} + w_i + a_j + c_j + e_{ijk}. \quad [1]$$

Here y_{ijk} is the k^{th} phenotypic record (LTTC) of animal *j* recorded within CG *i*, $\boldsymbol{\beta}$ is the vector of fixed effects that includes an overall intercept, linear regression coefficients for Nellore breed proportion, heterozygosity and

recombination loss as well as linear and quadratic regression coefficients on age of calf, $\mathbf{x}'_j$ is the known incidence row vector of covariates connecting $\boldsymbol{\beta}$ to y_{ijk}, w_i is the random effect of CG i (i =1,2,…,146 levels), a_j is the random additive genetic effect of animal j, c_j is the random permanent environment effect of animal j and e_{ijk} is the residual error.

The following distributional assumptions were assumed: $\mathbf{w} = \{w_i\} \sim N(\mathbf{0}, \mathbf{I}\sigma^2_w)$, $\mathbf{a} = \{a_j\} \sim N(\mathbf{0}, \mathbf{A}\sigma^2_a)$ and $\mathbf{c} = \{c_j\} \sim N(\mathbf{0}, \mathbf{I}\sigma^2_c)$ and $\mathbf{e} = \{e_{ijk}\} \sim N(\mathbf{0}, \mathbf{I}\sigma^2_e)$ by which $\sigma^2_w, \sigma^2_a, \sigma^2_c$ and σ^2_e represent, variances components due to CG, additive genetics, permanent environment and residual terms, respectively. Here, **A** represents the numerator relationship matrix between the animals in the pedigree and **I** is the identity matrix.

Hierarchical Bayesian reaction norm models (HBRNM)

Two somewhat different approaches were used to estimate environmental sensitivities of genotypes. One approach was based on a commonly used two-step modeling process (Calus et al., 2002; Kolmodin et al., 2002). The first step is based on using the regular animal model (M_1) from Equation [1] to estimate CG effects $\hat{w}_i$. The second step consists of using these $\hat{w}_i$ estimates as if they were "known" environmental covariates in a linear reaction norm model. More specifically, posterior means of $\hat{w}_i$ obtained by M_1, were used as covariate values in the following reaction norm model (M_2).

$$y_{ijk} = \mathbf{x}'_j \boldsymbol{\beta} + \phi \hat{w}_i + a_j + b_j \hat{w}_i + c_j + d_j \hat{w}_i + e_{ijk}. \quad [2]$$

Here ϕ is an overall linear regression coefficient of y_{ijk} on $\hat{w}_i$, a_j is the additive genetic intercept of animal j pertaining to genetic merit for an average environment ($\hat{w}_i$=0), b_j is random additive genetic effect of the reaction norm slope of animal j on $\hat{w}_i$, c_j is the non-genetic (e.g., permanent environment effect) intercept of animal j as defined for an average environment ($\hat{w}_i$=0), d_j is random permanent environment effect of the reaction norm slope of animal j on $\hat{w}_i$. Note that y_{ijk}, $\mathbf{x}'_j \boldsymbol{\beta}$ and e_{ijk} are defined as before.

Another two step modeling strategy (M_3) that is very similar to M_2 is given in Equation [3]:

$$y_{ijk} = \mathbf{x}'_j \boldsymbol{\beta} + w_i + a_j + b_j \hat{w}_i + c_j + d_j \hat{w}_i + e_{ijk}. \quad [3]$$

Note that in M_3, contemporary group effects are refitted as random effects rather than being treated as known covariates in an overall partial regression using Model M_2 based on their respective posterior means ($\hat{w}_i$) in a first stage model M_1. Hence M_3 may be more flexible than M_2 for modeling CG effects. Nevertheless, $\hat{w}_i$ were again used as "known" covariates in the random regression portion of the model.

Now including $\hat{w}_i$ as if they are "known" covariates in the second stage model is clearly a limitation that may understate statistical uncertainty and lead to biased predictions on animal genetic merit. These biases may be due to genetic trend or differences between environmental values across CG or both (Su et al., 2006; Cardoso and Tempelman, 2012).

An appealing one-step approach alternative to the two-step approach limitation was proposed by Su et al. (2006). This approach is purely Bayesian in that the covariate associated with the reaction norm is treated as unknown, thereby allowing inference for all unknowns together within a one-step linear reaction norm model (M_4):

$$y_{ijk} = \mathbf{x}'_j\boldsymbol{\beta} + w_i + a_j + b_j w_i + c_j + d_j w_i + e_{ijk}. \quad [4]$$

Model M_4 can be rewritten in matrix notation as below [14]:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{P}\mathbf{w} + \mathbf{Z}_\mathbf{a}\mathbf{a} + \mathbf{Z}_\mathbf{b}\mathbf{b} + \mathbf{Z}_\mathbf{c}\mathbf{c} + \mathbf{Z}_\mathbf{d}\mathbf{d} + \mathbf{e}, \quad [5]$$

in which $\mathbf{y} = \{y_{ijk}\}$ is the *nx1* vector of observations, $\boldsymbol{\beta}$ is the fixed effects vector of order *p*, $\mathbf{w} = \{w_i\}_{i=1}^{n_w}$ is the vector of environmental effects, $\mathbf{a} = \{a_j\}_{j=1}^{q}$ is the vector of random genetic intercepts, $\mathbf{b} = \{b_j\}_{j=1}^{q}$ is the vector of random genetic slopes, $\mathbf{c} = \{c_j\}_{j=1}^{q}$ is the vector of random permanent environment intercepts, $\mathbf{d} = \{d_j\}_{j=1}^{q}$ is the vector of random permanent environment slopes and $\mathbf{e}$ is the *nx1* vector of residuals. Furthermore, $\mathbf{X}$, $\mathbf{P}$, $\mathbf{Z_a}$ and $\mathbf{Z_c}$ are known incidence matrices, whereas the column address of matrices $\mathbf{Z_b}$ and $\mathbf{Z_d}$ has exactly one element equal to the environmental covariate (w_i or an estimate of w_i) for that CG in the row address of the observation, with all other elements in that row equal to zero.

Prior distributional specifications

To infer environmental sensitivities using a hierarchical Bayesian model, three stages are required: the first stage defines the distribution of the observed data conditional on all other parameters (Su et al., 2006):

$\mathbf{y}|\boldsymbol{\beta},\mathbf{w},\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{R} \sim N(\mathbf{X\beta} + \mathbf{Pw} + \mathbf{Z_a a} + \mathbf{Z_b b} + \mathbf{Z_c c} + \mathbf{Z_d d}, \mathbf{R})$. [6]

For homoscedastic residual specification such as for M_1, M_2 M_3 and M_4, $\mathbf{R} = \mathbf{I}\sigma^2_e$, where σ^2_e is the residual variance and **I** is the identity matrix. However, as we previously noted, it might be important to model residual heteroscedasticity. We propose two alternative strategies along this front. The first heteroskedastic residual specification (S_1) is such that $\mathbf{R} = diag(\mathbf{I}_{n_i}\sigma^2_{e_i})$, a diagonal matrix $\mathbf{I}_{n_i}$ denoting an identity matrix of order n_i being the number of records in the CG *i*, $\sigma^2_{e_i} = \sigma^2_e * \eta^{\hat{w}_i}$ for η being a hyperparameter that characterizes the heterogeneity of residual variance across environments (Cardoso et al. 2005).

Based on S_1, two alternative two-step reaction norm models (M_5 and M_6) were tested based on using $\sigma^2_{e_i} = \sigma^2_e * \eta^{\hat{w}_i}$ with both models based on using inferred values $\hat{w}_i$ from Model M_1 as if they were known. Model M_5 was then a residual heteroskedastic extension of Model M_2 whereas M_6 was a residual heteroskedastic extension of Model M_3.

Another heteroskedastic residual alternative (S_2) was based on residual variances subclasses determined by a decile-based classification of $\hat{w}_i$. That is, CG were ordered into one of 10 categories based on decile delimiters of $\hat{w}_i$ obtained from M_1, such that $\mathbf{R} = diag\left(\mathbf{I}_{n_k}\sigma^2_e\gamma_k\right)$ where order n_k denotes the number of records delimited by deciles *k*-1 and *k*, being respectively 1,157, 1,174, 1,047, 765, 1,188, 1,192, 1,208, 918, 1,150 and 874, respectively, for *k* = 1,2,…10. This strategy was also considered by Cardoso and Tempelman (2012). This specification was used to extend the two-step model M_2 in M_7 and the one-step model M_4 in M_8 but with this particular heteroskedastic residual specification.

Finally, the last model [M_9] implemented was a traditional animal model but with also a decile-based classification specification on residual heteroscedasticity (S_2) as a control model in order to determine between the

consequences of failing to model G*E versus failing to model residual heteroscedasticity (Cardoso and Tempelman, 2012). A summary of all models implemented in this study is presented in Table 2.

The second stage is represented by the prior distributions of the location parameters, as follows:

$$\beta \sim \mathrm{p}(\boldsymbol{\beta}) \quad [7]$$

$$\mathbf{w} \mid \sigma_{\mathrm{w}}^2 \sim \mathrm{N}(0, \mathbf{I}\sigma_{\mathrm{w}}^2) \quad [8]$$

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_{\mathbf{a}}^2 & \sigma_{\mathbf{ab}} \\ \sigma_{\mathbf{ab}} & \sigma_{\mathbf{b}}^2 \end{bmatrix} \otimes \mathbf{A} \right) \quad [9]$$

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_{\mathbf{c}}^2 & \sigma_{\mathbf{cd}} \\ \sigma_{\mathbf{cd}} & \sigma_{\mathbf{d}}^2 \end{bmatrix} \otimes \mathbf{I} \right), \quad [10]$$

where $\mathrm{p}(\boldsymbol{\beta}) \propto 1$, σ_{w}^2 is the environmental effect variance, σ_{a}^2 and σ_{b}^2 are, respectively, the additive genetic variances due to the reaction norm intercept and slope, σ_{c}^2 and σ_{d}^2 are, respectively, permanent environment variances due to reaction norm intercept and slope; σ_{ab} is the genetic covariance between reaction norm, intercept and slope, and σ_{cd} is the permanent environment covariance between reaction norm intercept and slope.

Then, $\mathbf{r_{ab}} = \sigma_{\mathrm{ab}} / \sqrt{\sigma_{\mathrm{a}}^2 * \sigma_{\mathrm{b}}^2}$ and $\mathbf{r_{cd}} = \sigma_{\mathrm{cd}} / \sqrt{\sigma_{\mathrm{c}}^2 * \sigma_{\mathrm{d}}^2}$ are the corresponding genetic and permanent environment correlations between intercept and slope, respectively.

Finally, the third stage of this model was based on specifying a scaled inverse (IG) distribution for the variance of the contemporary group effects; i.e. $\sigma_{\mathrm{w}}^2 \mid \alpha_{\mathrm{w}}, \beta_{\mathrm{w}} \sim \mathrm{IG}(\alpha_{\mathrm{w}} = 1, \beta_{\mathrm{w}} = 2800)$ where the mean of this distribution is:

$$\mathrm{E}\left(\mathrm{w_i} \mid \alpha_{\mathrm{w}}, \beta_{\mathrm{w}}\right) = \frac{\alpha_{\mathrm{w}}}{\beta_{\mathrm{w}}}. \quad [11]$$

Similarly, we specify $\sigma_{\mathrm{e}}^2 \mid \alpha_{\mathrm{e}}, \beta_{\mathrm{e}} \sim \mathrm{IG}(\alpha_{\mathrm{e}} = 1, \beta_{\mathrm{e}} = 600)$.

Likewise, an inverted Wishart distribution (IW) prior distribution was presumed, respectively, for the permanent environment and additive genetic covariance matrices:

$$\mathbf{U_0} = \begin{bmatrix} \sigma_{\mathbf{c}}^2 & \sigma_{\mathbf{cd}} \\ \sigma_{\mathbf{cd}} & \sigma_{\mathbf{d}}^2 \end{bmatrix} \sim \mathrm{IW}(\mathbf{T_0}, \mathrm{v}) \quad [12]$$

$$\mathbf{G}_0 = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix} \sim \mathrm{IW}(\mathbf{T}_0, \mathrm{v}). \quad [13]$$

Following Cardoso and Tempelman (2012), v=4 represents a presumed known degrees of freedom, $\mathbf{T}_0 = \begin{bmatrix} 200 & 4 \\ 4 & 0.1 \end{bmatrix}^{-1}$ a presumed scale matrix such that $\mathrm{E}(\mathbf{G}_0) = \frac{\mathbf{T}_0^{-1}}{\mathrm{v} - \mathrm{p} - 1}$ is the prior mean for $v>p+1$, where p is the number of parameters.

In the models where we specified heterogeneity of residual variances, additional hierarchical specifications were required, depending on the nature of the function (S_1 or S_2) chosen; i.e., $\eta|\alpha_\eta, \beta_\eta \sim \mathrm{p}(\eta|\alpha_\eta, \beta_\eta) = \mathrm{IG}(\alpha_\eta, \beta_\eta)$, for S_1 or $\gamma_k|\alpha_\gamma \sim \mathrm{p}(\gamma_k|\alpha_\gamma) = \mathrm{IG}(\alpha_\gamma, \alpha_\gamma - 1)$, k = 1,2,…,10 for S_2 (Cardoso et al., 2005; Kizilkaya and Tempelman, 2005; Cardoso and Tempelman, 2012).

We specified $\alpha_\eta = -1$ and $\beta_\eta = 0$ whereas the prior p(α_γ) on α_γ was a gamma with shape and scale hyperparameters values of 0.03 and 0.01, respectively (Cardoso and Tempelman, 2012). This assumption leads to a prior mean of α_γ equal to 3 ($\mathrm{E}(\alpha_\gamma) = 3$) and a large prior variance ($\mathrm{var}(\alpha_\gamma) = 300$) (Cardoso et al., 2005).

Due to absence of relevant previous knowledge, flat or highly dispersed prior densities were assumed for all parameters of all models implemented and hyperparameters for variance components priors were specified on the basis of REML estimates obtained by M_1 and M_2 (not shown).

Bayesian inference

Analyses were conducted using Bayesian analyses, specifically Monte Carlo Markov Chain (MCMC), to sample all parameters from their fully conditional posterior distributions. Gibbs sampling was generally used except for the w_i's and η in M_5 and M_6 and also for α_γ (S_2) in M_7, M_8 and M_9. MCMC sampling of these parameters required a random walk Metropolis-Hastings steps due to all of their fully conditional posterior distributions being unrecognizable (see Cardoso and Tempelman, 2012 for further details).

MCMC was implemented by INTERGEN software (Cardoso, 2013) by saving every 10th cycle from a total of 1,000,000 cycles, after 100,000 cycles of burn-in. Global convergence was checked using the Geweke's Z criterion (Geweke, 1991) applied to the conditional distribution of the data, as proposed by Brooks and Roberts (1998). In addition, visual inspection of trace plots was conducted and a minimum effective sample size of 100 for all unknown parameters was always obtained as well.

Table 2 Implemented model according approach, contemporary group effect and heteroscedasticity specification

M_x	two-step	CG	HET
M_1	N/A	R	N/A
M_2	Y	C	N/A
M_3	Y	R	N/A
M_4	N	R	N/A
M_5	Y	C	S_1
M_6	Y	R	S_1
M_7	Y	C	S_2
M_8	N	R	S_2
M_9	N/A	R	S_2

M_x: model number x; two-step (Y= Yes, N=No, N/A=non-applicable since model is not reaction norm model); CG: specification on contemporary group effects (C= covariate, R= random classification effect); HET: heterogeneous residual variance; S_1: exponential function on heteroscedastic residual variance; S_2: discrete subclasses based on classification function on heteroscedastic residual variance.

Model comparison

In order to identify models with the best fit to the data Deviance Information Criterion (DIC) was used (Spiegelhalter et al., 2002) :

$$DIC = \overline{D}(\theta) + p_D = 2\overline{D}(\theta) - D(\bar{\theta}), \quad [14]$$

where $\overline{D}(\theta) = E_{\theta|y}[D(\theta)]$ is the posterior expectation of Bayesian Deviance; $p_D = \overline{D}(\theta) - D(\bar{\theta})$ corresponds to the penalty for increasing model complexity where θ is the model parameters vector and $D(\bar{\theta})$ is the Bayesian Deviance as a function of the posterior mean of the parameters. Smaller values of DIC indicate better-fitting model.

Variance components and genetic parameters

The additive genetic variance for a specific environment *i* with effect w_i is obtained as follows:

$$\sigma_a^2 \mid w_i = \mathrm{var}\left(a_j + b_j w_i\right) = \sigma_a^2 + w_i^2\sigma_b^2 + 2w_i\sigma_{ab}. \quad [15]$$

Thus, the heritability (h_a^2) and repeatability (*r*) for a specific environment was determined as:

$$h_a^2 \mid w_i = \frac{\sigma_a^2 \mid w_i}{\sigma_a^2 \mid w_i + \sigma_c^2 \mid w_i + \sigma_e^2 \mid w_i}, \quad [16]$$

and $r \mid w_i = \frac{\sigma_a^2 \mid w_i + \sigma_c^2 \mid w_i}{\sigma_a^2 \mid w_i + \sigma_c^2 \mid w_i + \sigma_e^2 \mid w_i}$, [17]

respectively, where $\sigma_c^2 \mid w_i$ and $\sigma_e^2 \mid w_i$ are permanent environment and residual variances in environment *i*, respectively. For a homoscedastic error models (from M_1 to M_4) $\sigma_e^2 \mid w_i$ is constant, i.e., $\sigma_e^2 \mid w_i = \sigma_e^2 \forall i$. For heteroscedastic error models $\sigma_{e_i}^2 \mid w_i = \sigma_e^2 * \eta^{\hat{w}_i}$ for M_5 and M_6, and $\sigma_{e_i}^2 \mid w_i = \sigma_e^2 * \gamma_{k:i}$, where *k*:*i* denotes the decile-based classification *k* for CG *i* in models M_7, M_8 and M_9.

Furthermore, the genetic covariance between two environmental gradients based on covariate values w_i and $w_{i'}$ was calculated by

$$\mathrm{cov}_a\left(a_j + b_j w_i, a_j + b_j w_{i'}\right) = \sigma_a^2 + \left(w_i + w_{i'}\right)\sigma_{ab} + w_i w_{i'}\sigma_b^2, \quad [18]$$

such that the corresponding correlation between two specific environments was calculated as described below:

$$r_a\left(a_j + b_j w_i, a_j + b_j w_{i'}\right) = \frac{\mathrm{cov}_a\left(a_j + b_j w_i, a_j + b_j w_{i'}\right)}{\sqrt{(\sigma_a^2 + w_i^2\sigma_b^2 + 2w_i\sigma_{ab})(\sigma_a^2 + w_{i'}^2\sigma_b^2 + 2w_{i'}\sigma_{ab})}}. \quad [19]$$

Genetic merit

The genetic merit of sire *j* specific to a given environment *i* was obtained as $g_j \mid w_i = a_j + b_j w_i$ (Cardoso and Tempelman, 2012). Values of b_j close to zero indicate that g_j is relatively constant across various environments (w_i) such that sire *j* has an environmentally robust genetic merit. On the other hand, an environmentally sensitive genetic merit has a

large b_j such that their relative performance should substantially change on the environmental gradient (Falconer, 1990).

The sire breeding value estimates were compared by the ranking of the animals obtained by AM and HBRNM, for low, medium and high environmental levels. These values were respectively defined by the values representative of the 10^{th}, 50^{th} and 90^{th} percentiles for w_i. Potential differences in re-ranking of sires for selection by these models were also determined by the Spearman correlation between breeding values obtained from rankings of all animals as well as the top 10% (60) sires having large numbers (>12) of progeny for tick resistance between low, medium and high environmental levels based on models M_1 and M_9, and the three best fitting linear reaction norm models according DIC.

RESULTS AND DISCUSSION

Model comparison

The models M_1 and M_9 which were the only models not modeling any G*E with a linear reaction norm model showed the poorest fits, along with M_6, with M_1 having the worst fit. The comparison of DIC between model M_1 and M_9 implies that considering heterogeneity of residual variance across environments is important for modeling LTTC. However, these DIC improvements from homoscedastic to heteroskedastic error models were small compared to the improvements found going from regular animal to linear reaction norm models, thereby suggesting that modeling G*E is more important than modeling heterogeneous variances (Table 3).

Furthermore, the two one-step reaction norm models (M_4 and M_8) had lower DIC values compared to the corresponding two-step reaction norm models (M_2,M_3,M_5,M_6 and M_7) Thus, treating all CG effects as uncertain for modeling reaction norms in one-step hierarchical models seems to be important for modeling G*E. The lowest DIC value model (Table 3) was for M_8. Recall that M_8 allows for residual variance groupings into decile-based subclasses corroborating further with Cardoso and Tempelman (2012) which reported this same model as the best-fitting in characterizing post-weaning gain in Angus cattle.

Table 3 Deviance Criterion Information (DIC) value and model ranking of all models

Model	DIC Value	Ranking
M_1	4828.60	9º
M_2	3736.80	3º
M_3	4010.93	6º
M_4	3590.84	2º
M_5	3863.92	5º
M_6	4258.18	8º
M_7	3823.49	4º
M_8	3114.49	1º
M_9	4129.39	7º

M_1 = Linear animal model; M_2 = two-step linear reaction norm model; M_3 = two-step linear reaction norm model with the random contemporary group (CG) effect being re-estimated; M_4 = one-step linear reaction norm model with homoscedastic residual variance; M_5 = two-step linear reaction norm model with exponential function on heteroscedastic residual variance; M_6 = two-step linear reaction norm model with exponential function on heteroscedastic residual variance and with the random CG effect being re-estimated; M_7= two-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance and the random CG effect being re-estimated; M_8= one-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance; M_9 = Linear animal model with classification function grouped in discrete subclasses on heteroscedastic residual variance.

Inferences on contemporary group effects

The M_1 model estimated CG posterior means ($\hat{w}_i$) ranged from -0.849 to 0.805 (Figure 2). In order from the 0-1st to 9th-10th deciles, corresponding values of $\hat{w}_i$ were, respectively, -0.424, -0.224, -0.121, -0.032, 0.032, 0.107, 0.182, 0.240 and 0.316. Following Cardoso and Tempelman (2012), these values were the cutoff points for the decile-based heteroscedastic error subclasses defined in M_7, M_8 and M_9.

The estimates or posterior means $\hat{w}_i$ of w_i were similar for all models, regardless if G*E was considered as in M_3, M_4, M_6, M_7 and M_8 or not as in M_1 and M_9 (Figure 2); Pearson correlations among these estimates always exceeded 0.99 such that they were also not influenced by homoscedastic versus heteroskedastic error modeling. These results did not agree with those reported by Cardoso and Tempelman (2012) for Angus post-weaning gain cattle in which fitting the model with the decile-based heteroscedastic classification function (S_2) has substantially different correlation estimates from this model with heteroscedastic exponential function models (S_1) or even conventional animal models. Furthermore, every single model presented negative skewness on the $\hat{w}_i$, ranging from -0.521 to -0.415.

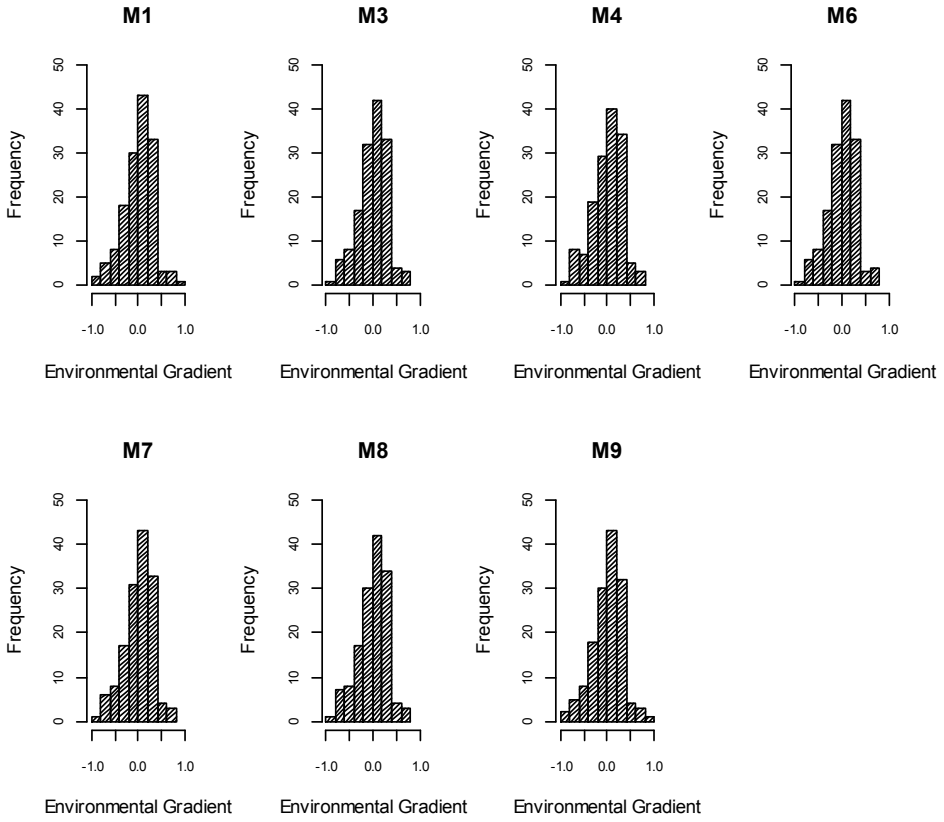


Figure 2 Environmental Gradient for Hereford and Braford tick counts on posterior mean contemporary group effect estimated by different models: M_1 = Linear animal model; M_3 = two-step linear reaction norm model with the random contemporary group (CG) effect being re-estimated; M_4 = one-step linear reaction norm model with homoscedastic residual variance; M_6 = two-step linear reaction norm model with exponential function on heteroscedastic residual variance and with the random CG effect being re-estimated; M_7 = two-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance and the random CG effect being re-estimated; M_8 = one-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance; M_9 = Linear animal model with classification function grouped in discrete subclasses on heteroscedastic residual variance.

Inferences on variance components and genetic parameters

The one-step approach model (M_8) resulted in the highest (0.022±0.04) estimated reference or intercept genetic variance (σ^2_a) in comparison with all tested models (Table 4). In addition, M_8 showed the highest estimated significant slope genetic variance (σ^2_b) in comparison to two-step approach, except for M_3 which had the same estimate (0.046±0.022). The estimated reference permanent environment (PE) variance components (σ^2_c) were similar among tested models (range 0.007-

0.010). In agreement with σ^2_b, PE slopes (σ^2_d) were also significant (range 0.021-0.084). These results demonstrate the one-step approach to be best-fitting and also confirmed the presence of G*E. Biegelmeyer (2012) working with tick resistance for Hereford and Braford beef cattle reported similar estimates (0.012;0.022) for σ^2_a and σ^2_c, respectively.

Estimated correlations between intercept and slope for both sets of random effects (i.e., additive genetic and permanent environment effects) were characterized by great uncertainty as indicates by the widths of their respective 95% posterior probability interval (PPI; Table 4). These results differ from previous studies (Shariati et al., 2007; Mattar et al., 2011; Cardoso and Tempelman, 2012) whereby these correlations were estimated to be large and positive.

Residual variance estimates (σ^2_e) were similar among models, ranging from 0.062±0.001 to 0.074±0.010, being slightly higher in traditional animal models M_1 (0.072±0.001) and M_9 (0.074±0.010), thereby confirming the importance of considering G*E in genetic evaluations for Hereford and Braford beef cattle (Table 4). Cardoso and Tempelman (2012) also reported HBRNM having lower estimates of σ^2_e than that using AM, further corroborating with this study.

Despite the similarity of residual variance across the various reaction norm models, Figure 3 nevertheless illustrates the need to consider residual heteroscedasticity. Moreover, the first decile class was particularly deviant compared to other classes. The very large residual variance values reported at the low extreme CG effects boundary, i.e., lower tick infestation, may be a statistical artifact that often occurs for extreme covariate values in random regression models (Meyer, 1998). Similar situations were demonstrated by Cardoso and Tempelman (2012) in which residual variances remarkably decreased at the extremes of CG average performance. Furthermore, Figure 3 also explains the poor fit of models which modeled heteroscedastic residual variance in models M_5 and M_6 as an exponential function (i.e. S_1). This function forces a gradual monotonic change in the residual variances over the CG classes whereas in M_7, M_8 and M_9 no clear pattern was noticeable.

Heritability estimates (h^2) were generally higher for HBRNM and also for M_9 in comparison to M_1 (h^2= 0.19±0.04; Figure 4a). Similar estimates have

been reported in the literature using models such as M_1 and logarithmic transformation of the observed data (Budeli et al. 2009, Ayres et al., 2013). Moreover, M_8 had higher estimated average heritability values, indirectly also indicating better fit of one-step versus two-step models that consider residual heteroscedasticity. Other beef cattle studies also presented higher average heritability estimates using HBRNM compared to AM (Pégolo et al., 2009; Cardoso and Tempelman, 2012). Therefore, greater selection response is expected using reaction norm models that model heterogeneity of residual variances across CG.

Estimates of repeatability varied along the CG (range 0.18-0.45) being, in general, higher under high tick infestation levels (Figure 4b). These results demonstrate the particular importance of modeling permanent environmental effects in harsh environments such that more resistant animals are more likely to maintain a consistent performance even in harsher environments than in favorable environments.

The posterior means of the genetic correlations based on r_g ($a_j + b_j w_i$, $a_j + b_j w_{i'}$) along environmental gradients surface plot for Hereford and Braford LTTC obtained by the best-fitting model M_8 demonstrated a large plateau above 0.80 (Figure 5). Furthermore, genetic correlation estimates decreased between environments reporting negative correlations between extreme environmental levels, i.e., between more favorable and harsh environments. Similar results that demonstrate genetic correlations differences along environmental levels, mainly between high challenge and favorable environments have been reported in literature (Corrêa et al., 2009; Mattar et al., 2011; Cardoso and Tempelman, 2012). Nevertheless, Ambrosini et al. (2012) estimated small differences for Nellore yearling weight with genetic correlations ranging from 0.78 to 1.00. In addition, Figure 5 also shows if selection would occur in an average environment (CG level = 0), the responses in extreme environments, i.e., the poorest and best environments would be approximately 68% and 66% of the direct responses obtained under selection in these extreme environments. Therefore, it reiterates the importance to select animals considering G*E.

Table 4 Posterior means and 95% posterior probability intervals reported as [2.5th, 97.5th] posterior percentiles of dispersion parameters estimated for Hereford and Braford tick counts by different models

Parameter[1]	Model[2]								
	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9
σ^2_a	0.019 [0.011,0.026]	0.019 [0.012,0.025]	0.020 [0.013,0.026]	0.020 [0.013,0.026]	0.019 [0.013,0.025]	0.018 [0.012,0.025]	0.020 [0.014,0.027]	0.022 [0.014,0.028]	0.021 [0.013,0.028]
σ^2_b	N/A	0.032 [0.006,0.074]	0.046 [0.011,0.094]	0.036 [0.008,0.087]	0.036 [0.009,0.069]	0.038 [0.013,0.065]	0.031 [0.009,0.057]	0.046 [0.009,0.098]	N/A
σ^2_c	0.010 [0.003,0.018]	0.008 [0.003,0.014]	0.008 [0.003,0.015]	0.007 [0.002,0.014]	0.008 [0.003,0.015]	0.009 [0.004,0.015]	0.010 [0.004,0.016]	0.009 [0.003,0.016]	0.010 [0.004,0.017]
σ^2_d	N/A	0.063 [0.022,0.098]	0.053 [0.009,0.093]	0.084 [0.031,0.123]	0.040 [0.009,0.075]	0.023 [0.005,0.051]	0.021 [0.004,0.049]	0.055 [0.006,0.106]	N/A
r_{ab}	N/A	-0.23 [-0.69,0.42]	-0.19 [-0.65,0.43]	-0.14 [-0.65,0.55]	-0.28 [-0.73,0.40]	-0.16 [-0.60,0.26]	-0.28 [-0.67,0.29]	-0.07 [-0.57,0.58]	N/A
r_{cd}	N/A	-0.09 [-0.65,0.62]	-0.07 [-0.71,0.74]	-0.11 [-0.69,0.60]	0.33 [-0.47,0.88]	0.63 [-0.08,0.94]	0.53 [-0.36,0.93]	0.30 [-0.50,0.89]	N/A
σ^2_w	0.099 [0.079,0.126]	N/A	0.097 [0.076,0.123]	0.101 [0.080,0.129]	N/A	0.097 [0.076,0.123]	0.096 [0.076,0.123]	0.097 [0.076,0.124]	0.098 [0.077,0.125]
σ^2_e	0.072 [0.069,0.074]	0.063 [0.060,0.065]	0.064 [0.061,0.066]	0.062 [0.060,0.065]	0.064 [0.061,0.067]	0.066 [0.063,0.068]	0.068 [0.055,0.089]	0.065 [0.053,0.085]	0.074 [0.059,0.099]

[1] σ^2_a = reaction norm intercept genetic variance; σ^2_b = reaction norm slope genetic variance; σ^2_c = reaction norm intercept permanent environment variance; σ^2_d = reaction norm slope permanent environment variance; r_{ab} = genetic correlation between intercept and slope of the reaction norm; r_{cd} = permanent environment correlation between intercept and slope of the reaction norm; σ^2_w = contemporary group effect (environmental) variance; σ^2_e = residual variance.

[2] M_1 = Linear animal model; M_2 = 2-step linear reaction norm model; M_3 = 2-step linear reaction norm model with the random CG effect being re-estimated; M_4 = 1-step linear reaction norm model with homoscedastic residual variance; M_5 = 2-step linear reaction norm model with exponential function on heteroscedastic residual variance; M_6 = 2-step linear reaction norm model with exponential function on heteroscedastic residual variance and with the random CG effect being re-estimated; M_7= 2-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance and the random CG effect being re-estimated; M_8= 1-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance; M_9 = Linear animal model with classification function grouped in discrete subclasses on heteroscedastic residual variance; N/A = non-applicable.

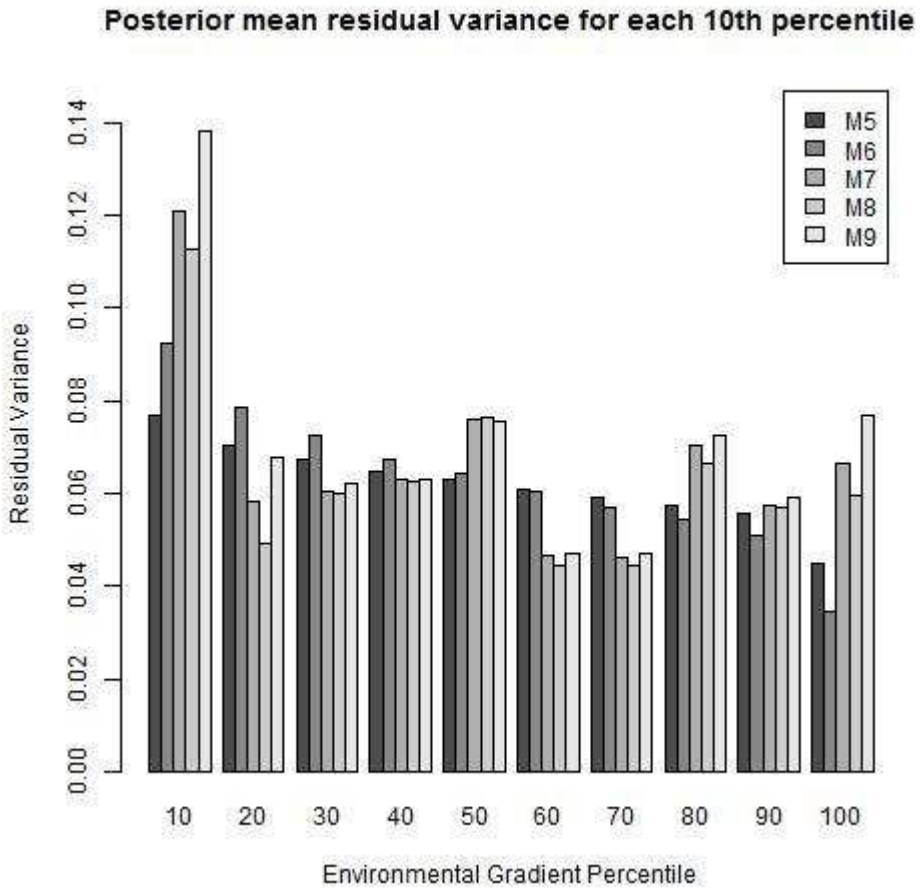


Figure 3 Posterior mean residual variance for each 10th percentile of the Hereford and Braford tick counts environmental gradient ranging from -0.849 to 0.805 obtained by two-step linear reaction norm model with exponential function on heteroscedastic residual variance (M_5); two-step linear reaction norm model with exponential function on heteroscedastic residual variance and with the random CG effect being re-estimated (M_6); two-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance and the random CG effect being re-estimated (M_7); one-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance (M_8); Linear animal model with classification function grouped in discrete subclasses on heteroscedastic residual variance (M_9).

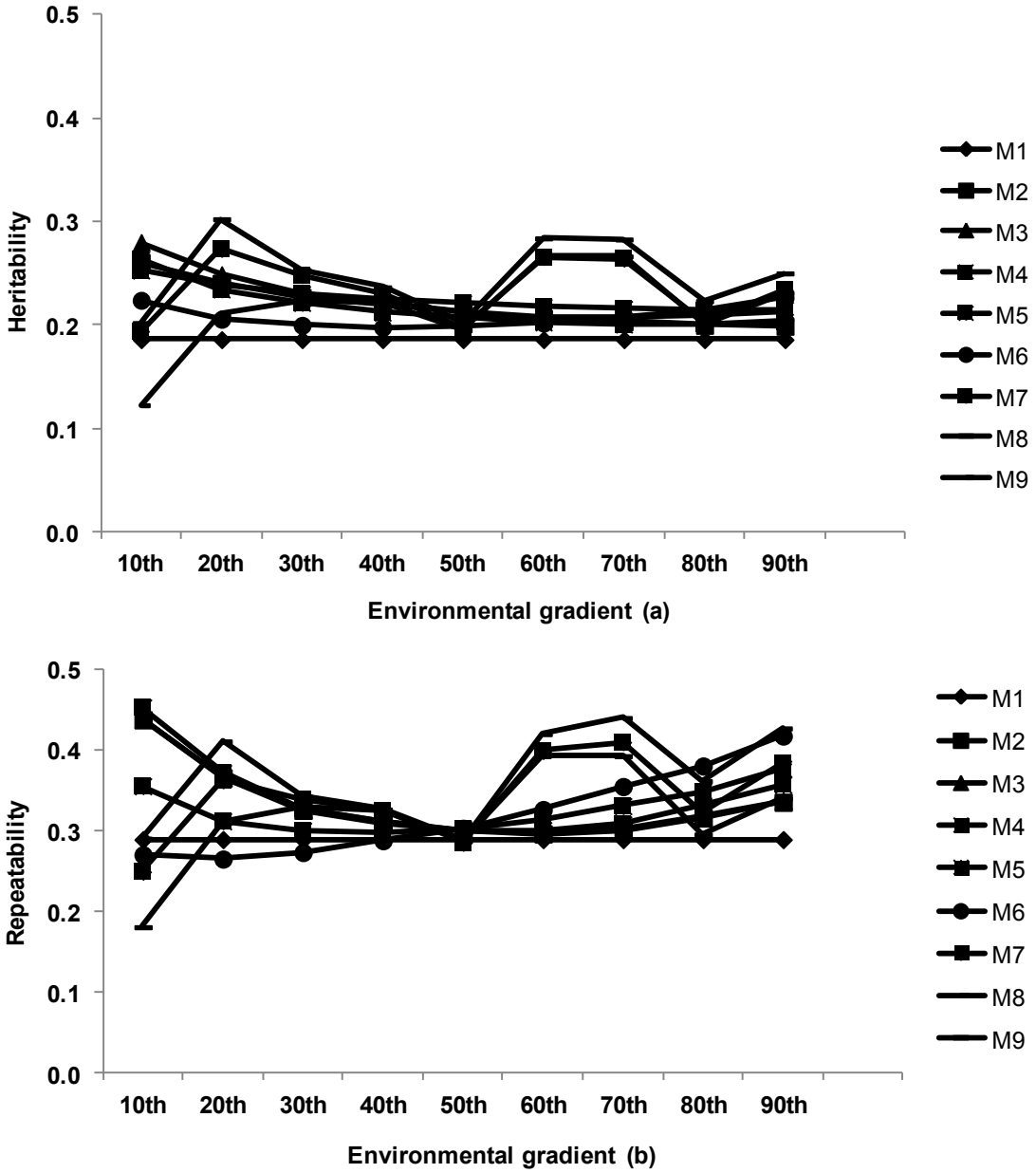


Figure 4 Posterior means heritabilities (a) and repeatabilities (b) for environments between the 10^{th} and 90^{th} percentiles of the environmental gradient obtained for Hereford and Braford tick counts by different models: M_1 = Linear animal model; M_2 = two-step linear reaction norm model; M_3 = two-step linear reaction norm model with the random contemporary group (CG) effect being re-estimated; M_4 = one-step linear reaction norm model with homoscedastic residual variance; M_5 = two-step linear reaction norm model with exponential function on heteroscedastic residual variance; M_6 = two-step linear reaction norm model with exponential function on heteroscedastic residual variance and with the random CG effect being re-estimated; M_7= two-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance and the random CG effect being re-estimated; M_8= one-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance; M_9 = Linear animal model with classification function grouped in discrete subclasses on heteroscedastic residual variance.

Inferences on genetic merit

Correlations were always above 0.87 indicating that rankings of animals for selection for a "typical" environment would be similar such that losses on selection precision by employing a traditional animal model might not at first seem to be substantial for all animals in data set or even the top 10% sires having large number of progeny (Table 5). However, Figure 6 demonstrates evidence of G*E provided by the extensive crossing over of sire-specific reaction norm lines, as observed for the 10% (60) most commonly used Hereford and Braford bulls; for example, animals that would be identified as top breedstock for tick infestation in a favorable environment could be identified as having particularly poor genetic merit in high challenge environments. These results are in agreement with other studies (Corrêa et al., 2010; Cardoso and Tempelman, 2012).

Table 5 Spearman rank correlations[a] among posterior means genetic values for tick counts of Hereford and Braford cattle at different tick infestation levels obtained by the linear animal models and the three better-fitting reaction norm models

Model[1]	**M1**	**M2**	**M2**	**M2**	**M4**	**M4**	**M4**	**M8**	**M8**	**M8**	**M9**
Environmental level[2,3]	(**Ov.**)	(**LTI**)	(**MTI**)	(**HTI**)	(**LTI**)	(**MTI**)	(**HTI**)	(**LTI**)	(**MTI**)	(**HTI**)	(**Ov.**l)
M_1 (**Ov.**)		0.97	0.99	0.97	0.97	0.98	0.96	0.94	0.97	0.94	0.98
M_2 (**LTI**)	0.97		0.98	0.92	0.99	0.96	0.90	0.98	0.95	0.89	0.94
M_2 (**MTI**)	0.98	0.97		0.98	0.97	0.99	0.97	0.95	0.99	0.95	0.99
M_2 (**HTI**)	0.96	0.91	0.98		0.92	0.98	0.99	0.90	0.98	0.98	0.99
M_4 (**LTI**)	0.94	0.97	0.94	0.89		0.97	0.91	0.98	0.96	0.89	0.94
M_4 (**MTI**)	0.96	0.95	0.97	0.95	0.97		0.98	0.94	0.99	0.96	0.99
M_4 (**HTI**)	0.94	0.89	0.95	0.96	0.92	0.98		0.88	0.97	0.99	0.98
M_8 (**LTI**)	0.94	0.98	0.94	0.88	0.95	0.91	0.86		0.95	0.88	0.93
M_8 (**MTI**)	0.97	0.96	0.99	0.97	0.92	0.95	0.93	0.95		0.98	0.99
M_8 (**HTI**)	0.94	0.89	0.96	0.99	0.86	0.92	0.94	0.88	0.98		0.98
M_9 (**Ov.**)	0.98	0.94	0.99	0.98	0.92	0.96	0.95	0.93	0.99	0.98	

[a] Correlations between all animal above the diagonal and between the most used sires (60) below the diagonal; [1] M_1 = Linear animal model; M_2 = two-step linear reaction norm model; M_4 = one-step linear reaction norm model with homoscedastic residual variance; M_8= one-step linear reaction norm model with classification function grouped in discrete subclasses on heteroscedastic residual variance; M_9 = Linear animal model with classification function grouped in discrete subclasses on heteroscedastic residual variance. [2] LTI = low tick infestation; MTI = medium tick infestation; HTI = high tick infestation; Ov. = overall. [3]LTI= 10^{th} (-0.424); MTI= 50^{th} (0.032); HTI= 90^{th} (0.316) percentiles of the contemporary group effects (environmental gradient).

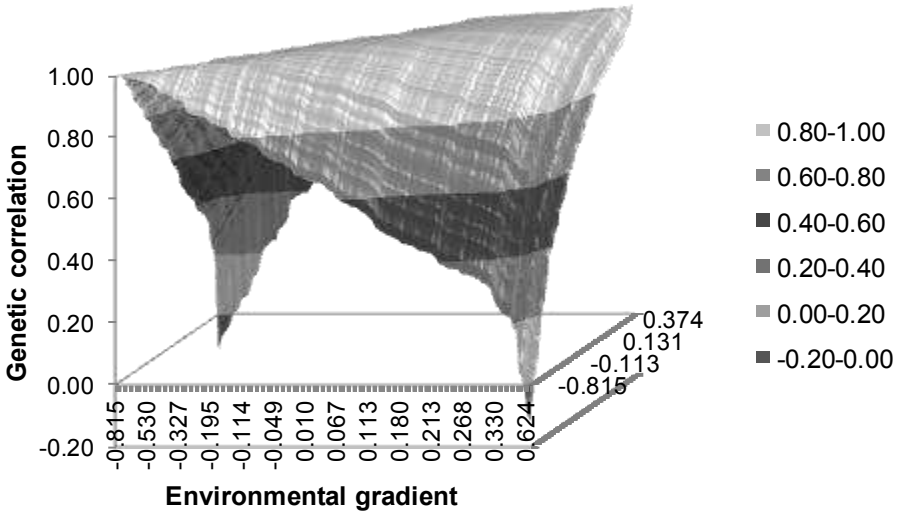


Figure 5 Genetic correlations for Hereford and Braford tick resistance performance on different environmental conditions obtained by the one-step linear reaction norm model with classification function on heteroscedastic residual variance (M_8)

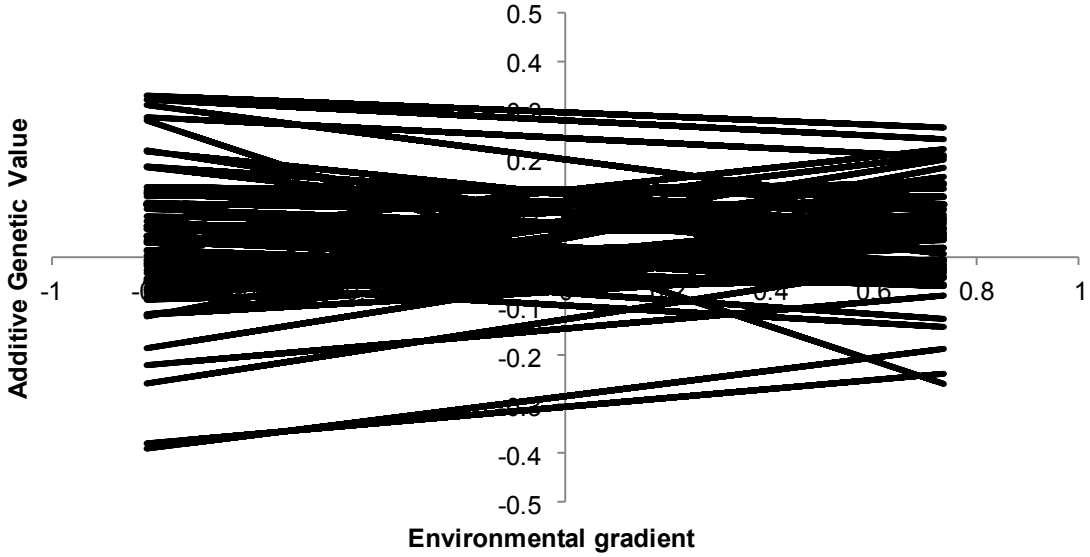


Figure 6 Genetic tick resistance reaction norms of 10% most used (60) Hereford and Braford sires obtained by the one-step linear reaction norm model with classification function on heteroscedastic residual variance (M_8).

CONCLUSIONS

Hierarchical Bayesian reaction norms models can be used to describe the presence of genotype by environment interaction on tick resistance in Hereford and Braford beef cattle. The model that best fitted tick counts in Brazilian climates was a one step model that considers heteroscedastic residual variance based on 10 discrete classes based on deciles of average CG performance (M_8) and hence should strongly be considered as a viable candidate model for genetic evaluation of this population. Reaction norm

models are a powerful tool to identify and quantify genotype by environment interaction and represent a promising alternative for genetic evaluation in tick resistance leading in selection efficiency and genetic progress.

ACKNOWLEDGMENTS

The authors thank Delta G Connection by providing the data used in this research; Embrapa South Livestock, Instituto Nacional de Investigacion Agropecuaria (INIA – Uruguay) and Michigan State University for theoretical and technical support; CAPES and CNPq by granting the scholarship. This project was supported by Agriculture and Food Research Initiative Competitive Grant no. 2011-67015-30338 from the USDA National Institute of Food and Agriculture.

REFERENCES

AMBROSINI, D.P.; CARNEIRO, P.L.S.; BRACCINI NETO, J. et al. Interação genótipo x ambiente para peso ao ano em bovinos Nelore Mocho no Nordeste do Brasil. **Pesquisa Agropecuária Brasileira**, v.47, p.1489-1495, 2012.

AYRES, D.R.; PEREIRA, J.R.; BOLIGON, A.A. et al. Linear and Poisson models for genetic evaluation of tick resistance in cross-bred Hereford x Nellore cattle. **Journal of Animal Breeding and Genetics**, v.130, p.417-424, 2013.

BIEGELMEYER, P. **Resistência genética à infestação natural e artificial por R*hipicephalus* (Boophilus) m*icroplus* em bovinos das raças Hereford e Braford**. 2012, 96p. Dissertação (Mestrado em Ciências) - Universidade Federal de Pelotas, Pelotas, RS.

BUDELI, M.A.; NEPHAWE, K.A.; NORRIS, D. et al. Genetic parameter estimates for tick resistance in Bonsmara cattle. **South African Journal of Animal Science**, v.39, p.321-327, 2009.

BURROW, H.M. Variances and covariances between productive and adaptive traits and temperament in a composite breed of tropical beef cattle. **Livestock Production Science**, v.70, p.213-233, 2001.

BROOKS, S.P.; ROBERTS, G.O. Convergence assessment techniques for Markov Chain Monte Carlo. **Statistics and Computing**, v.8, p.319-335, 1998.

CALUS, M.P.L.; GROEN, A.F.; DE JONG, G. Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. **Journal of Dairy Science**, v.85, p.3115-3123, 2002.

CARDOSO, F.F.; ROSA, G.J.M.; TEMPELMAN, R.J. Multiple-breed genetic inference using heavy-tailed structural models for heterogeneous residual variances. **Journal of Animal Science**, v.83, p.1766-1779, 2005.

CARDOSO, L.L.; BRACINI NETO, J.; CARDOSO, F.F. et al. Hierarchical Bayesian models for genotype x environment interaction estimation via

reaction norms applied to post-weaning gain of Hereford cattle. **Revista Brasileira de Zootecnia**, v.40, p.294-300, 2011.

CARDOSO, F.F.; TEMPELMAN, R.J. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. **Journal of Animal Science**, v.90, p.2130-2141, 2012.

CARDOSO, F.F. **Application of Bayesian inference in animal breeding using the Intergen program: Manual of version 1.22**. Documents 112, Embrapa South Livestock Bagé, RS, Brazil, 2013.

CARVALHEIRO, R.; FRIES, L.A.; SCHENKEL, F.S. et al. Efeitos da heterogeneidade de variância residual entre grupo de contemporâneos na avaliação genética de bovinos de corte. **Revista Brasileira de Zootecnia**, v.31, p.1680-1688, 2002.

CORRÊA, M.B.B.; DIONELLO, N.J.L.; CARDOSO, F.F. Caracterização da interação genótipo-ambiente e comparação entre modelos para ajuste do ganho do pós-desmama de bovinos Devon via normas de reação. **Revista Brasileira de Zootecnia**, v.38, p.1468-1477, 2009.

CORRÊA, M.B.B.; DIONELLO, N.J.L.; CARDOSO, F.F. Genetic evaluation of Devon cattle using a reaction norms model. **Revista Brasileira de Zootecnia**, v.39, p.128-133, 2010.

FALCONER, D.S. Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. **Genetics Research**, v.56, p.57-70, 1990.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to Quantitative Genetics**. 4 ed. Longman Group Ltd, Harlow, UK, 1996.

FRISCH, J.E.; O'NEILL, C.J. Comparative evaluation of beef cattle breeds of African, European and Indian origins. 2. Resistance to cattle ticks and gastrointestinal nematodes. **Animal Science**, v.67, p.39-48, 1998.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: BERNARDO, J.M.; BERGER, J.O.; DAWID, A.P. et al. (Eds.). **Bayesian Statistics 4**. New York: Oxford University, 1991. 526p.

IBELLI, A.M.G.; RIBEIRO, A.R.B.; GIGLIOTI, R. et al. Resistance of cattle of various genetic groups to the tick Rhipicephalus microplus and the relationship with coat traits. **Veterinary Parasitology**, v.186, p.425-430, 2012.

KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, p.979-993, 1990.

KIZILKAYA, K.; TEMPELMAN, R.J. A general approach to mixed effects modeling of residual variances in generalized linear mixed models. **Genetic Selection Evolution**, v.37, p.31-56, 2005.

KNAP, P.W.; SU, G. Genotype by environment interaction for litter size in pigs as quantified by reaction norm analysis. **Animal**, v.2, p.1742-1747, 2008.

KOLMODIN, R.; STRANDBERG, E.; MADSEN, P. et al. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. **Acta Agriculturae Scandinava Section A Animal Science**, v.52, p.11-24, 2002.

LYNCH, M.; WALSH, B. **Genetics and Analysis of Quantitative Traits**. 1st ed. Sinauer Associates, Inc., Sunderland, MA, 1998.

MATTAR, M.; SILVA, L.O.C.; ALENCAR, M.M. et al. Genotype by environment interaction for long-yearling weight in Canchim cattle quantified by reaction norm analysis. **Journal of Animal Science**, v.89, p.2349-2355, 2011.

MEYER, K. Estimating covariance functions for longitudinal data using a random regression model. **Genetic Selection Evolution**, v.30, p.221-240, 1998.

PÉGOLO, N.T.; OLIVEIRA, H.N.; ALBUQUERQUE, L.G. et al. Genotype by environment interaction for 450-day weight of Nelore cattle analyzed by reaction norm models. **Genetics and Molecular Biology**, v.32, p.281-287, 2009.

PRAYAGA, K.C.; HENSHALL, J.M. Adaptability in tropical beef cattle: genetic parameters of growth, adaptive and temperament traits in a crossbred population. **Australian Journal Agriculture Research**, v.45, p.971-983, 2005.

ROSO, V.M.; SCHENKEL, F.S. AMC – A computer program to assess the degree of connectedness among contemporary groups. ***Proceedings*** of 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil, 2006. (CD-ROM).

SHARIATI, M.M.; SU, G.; MADSEN, P.; et al. Analysis of milk production traits in early lactation using a reaction norm model with unknown covariates. **Journal of Dairy Science**, v.90, p.5759-5766, 2007.

SILVA, A.M.; ALENCAR, M.M.; REGITANO, L.C.A. et al. Infestação natural de fêmeas bovinas de corte por ectoparasitas na Região Sudeste do Brasil. **Revista Brasileira de Zootecnia**, v.39, p.1477-1482, 2010.

SPIEGELHALTER, D.J.; BEST, N.G.; CARLIN, B.P. et al. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society Series B-Statistical Methodology**, v.64, p.583-616, 2002.

SU, G.; MADSEN, P.; LUND, M.S. et al. Bayesian analysis of the linear reaction norm model with unknown covariates. **Journal of Animal Science**, v.84, p.1651-1657, 2006.

TORRES, R.A.; BERGMANN, J.A.G.; COSTA, C.N. et al. Heterogeneidade de variância e avaliação genética de bovinos da raça holandesa no Brasil. **Revista Brasileira de Zootecnia**, v.29. p.1050-1059, 2000.

WHARTON, R.H.; UTECH, K.B.W. The relation between engorgement and dropping of Boophilus microplus (Canestrini) (Ixodidae) to the assessment to tick numbers on cattle. **Journal Australian Entomologycal Society**, v.9, p.171-182, 1970.

Chapter 2

GENOME-ENABLED PREDICTION FOR TICK RESISTANCE IN HEREFORD AND BRAFORD BEEF CATTLE VIA REACTION NORM MODELS

Abstract: Studies to investigate genotype by environment interaction (G*E) and account for marker information are still scarce, especially using one-step reaction norm model approach. Therefore, the main objective of this study was to compare a conventional genomic-based single step BLUP model (GBLUP) with its one-step genomic reaction norm model extension (GLRNM) on tick infestation phenotypes, as well as to compare the performance of genomic estimated breeding values (GEBV) predictions obtained from using or not using marker information. Phenotypic data included 10,673 tick counts on 4,363 Hereford and Braford animals. The pedigree included 12,754 animals from which 3,591 were genotyped. Four different models were evaluated, being two animal models (BLUP, GBLUP) and two one-step reaction norm model (LRNM, GLRNM) with (GBLUP, GLRNM) and without (BLUP, LRNM) genomics. The deviance information criterion (DIC) was used for model choice. BLUP and GBLUP seemed to be poorer fitting in comparison with their one-step reaction norm extensions. The reaction norm model accounting for marker information (GLRNM) presented lower intercept and slope genetic variance compared to LRNM. Heritability and repeatability estimates varied along the environmental gradient (EG) and were, in general, similar for both models. Genetic correlations were remarkably low between extreme EG, indicating the presence of G*E for tick resistance in these populations. Cross validation estimates were in average 0.66±0.02, 0.67±0.02, 0.67±0.02 and 0.66±0.02 for BLUP, GBLUP, GLRNM and LRNM, respectively, based on K-means partitioning, whereas GLRNM was 0.71±0.01 and tend to better than BLUP (0.67±0.01), GBLUP (0.70±0.01) and LRNM (0.70±0.01) based on random partitioning. However, no statistical significance was reported between GLRNM and LRNM. Our results also suggest that marker information do not lead for higher prediction accuracies

which decreased as the tick infestation level increased and as the relationship between animals in training and validation datasets decreased.
Keywords: accuracy, cross validation, genetic correlation, heritability.

INTRODUCTION

The cattle tick is a substantial concern, especially in tropical areas, because it can greatly diminish animal performance. Furthermore, parasite resistance due to indiscriminate use of treatments with acaricides, chemical residues in milk and beef and also repeated failures of effective vaccines development had driven researchers to seek for alternative solutions. The best alternative to overcome this problem might be the identification of naturally resistant animals by selection and the encouragement of their intensive use for breeding. Studies have demonstrated that tick resistance may be heritable with different levels of cattle resistance to ticks, i.e., heritability estimates ranging from 0.05 to 0.42 (Rechav et al., 1990; Burrow et al., 2001; Budeli et al., 2009; Oliveira et al., 2012, Ayres et al., 2013).

Reaction norms are phenotype profiles influenced by genotypes varying across environments. The use of linear reaction norm models (RNM) allows the estimation of environmental sensitivities of different genotypes, which change gradually and continuously in an environmental gradient (EG; Falconer and Mackay, 1996). Several studies have reported the importance of genotype by environment interaction (G*E) in different traits in beef cattle leading to re-rankings of animals' performance in different environments (Pégolo et al., 2009; Corrêa et al., 2010; Ambrosini et al., 2012). Moreover, Mota et al. (2013) has suggested that G*E exists for tick resistance and may be captured using RNM.

Nowadays, the use of genomic-wide selection (GWS; Meuwissen et al., 2001) enhances the accuracy of genomic estimated breeding values (GEBV) allowing the investigation of G*E based on marker information. Silva et al. (2014) studying on total number born in pigs inferred the presence of G*E based on genomic RNM and suggested that RNM could be implemented for prediction of GEBVs as a promising approach to account for G*E. However, these studies were implemented using the most common reaction norm approach, the two-step approach (Calus et al., 2002; Kolmodin et al., 2002). This approach estimates environmental effects on the phenotype of interested not considering G*E by a conventional animal model (BLUP) as first step. In the second step, the environmental estimates from BLUP are considered as "known" covariates in a random regression model.

According to Su et al. (2006), prediction of two-step reaction norm breeding value as a function of EG could lead to bias inferences and consequently incorrect ranking of animals. These authors proposed a one-step Bayesian approach which treats the covariate associated with the reaction norm as unknown inferring for all unknowns parameters together. Mota et al. (2013) reported that models fitted by one-step approach showed better fit than the two-step for tick resistance in Hereford and Braford beef cattle populations as well as Cardoso (2013) developed software (INTERGEN) to fit these models considering heterogeneity residual variance and marker information. Cardoso and Tempelman (2012) reported that one-step models considering heterogeneity residual variance best-fit in comparison with two-step and/or homogeneity residual variance for post-weaning body weight gain in Angus cattle.

The objectives of this study were to compare a conventional genomic-based BLUP model (GBLUP) with its one-step genomic reaction norm model extension (GLRNM) on tick infestation data, to compare estimates obtained from using non genomic and genomic approaches as well as to investigate the efficiency of GEBV prediction through cross validation studies.

MATERIAL AND METHODS

Phenotypic and genotypic data

Phenotypic data used in this current study included records of tick counts (TC) on Hereford and Braford beef cattle from herds raised in the Rio Grande do Sul state, Brazil. TC were obtained on each animal from 326 to 729 days of age using the method described by Wharton and Utech (1970), whereby all female ticks ≥ 4.5 mm were counted on the entire left side of the animal. Up to 3 such counts were acquired from each animal with each such count separated by a minimum of 30 days similar as in other studies (Frisch and O'Neill, 1998; Budeli et al., 2009; Biegelmeyer, 2012). The distribution of the number of measurements taken per animal was 241, 1,934 and 2,188 animals having one, two and three TC measurements, respectively, for a total of 10,673 records. The average age during the evaluation period was 524 ± 65 days and the overall mean TC was 34.99 with a standard deviation of 42.15 (range 0- 532).

Because TC were not typically normally distributed, the log transformation was used such that LTTC = $\log_{10}$ (TC+1.001) was the response variable. The constant 1.001 was included in this transformation as some of the TC were equal to zero (Ayres et al., 2013; Biegelmeyer, 2012). Note that data transformation does not guarantee that the log transformed data are normal, but might improve properties of estimates, predictions and inferences (Budeli et al., 2009).

A CG was defined as a group of animals being within the same herd, year of birth, season of birth (April to July; August to November and December to March), of the same sex and from the same management group. Each CG was required to have at least five animals and with each LTTC record being within 3.5 Standard Deviation (SD) from their respective CG means. Moreover, connectedness among CG was assessed by the AMC software (Roso and Schenkel, 2006) such that CG with less than 10 genetic links, were removed. Finally, CG effects/means for LTTC were assumed to define the environmental covariates for a linear reaction norm model because they seemed to be the most appropriate entities to describe environmental conditions most important for beef cattle production (Cardoso et al., 2011; Mattar et al., 2011; Cardoso and Tempelman, 2012).

The DNA for each animal was obtained from a blood sample at the time of first tick count. In addition, semen samples from bulls used in artificial insemination (AI) were also obtained. Genotypes based on the BovineSNP50 Illumina v2 BeadChip were acquired on 3,591 of these Hereford and Braford beef cattle. It was provided a total of 54,609 highly informative SNPs uniformly distributed across the entire genome of major cattle breed types.

Genotype quality control was implemented using R/SNPStats package (Clayton, 2012) to remove samples with call rates <0.90, heterozygosity deviations >3.0, mismatching sex and duplicated records. Only SNPs mapped to autosomes, with call rates greater than 0.98, minor allele frequencies (MAF) >0.03 or not highly significant deviations from Hardy-Weinberg equilibrium ($P>10^{-7}$) were used for further analyses. We considered the highest MAF for SNPs in the same position or highly correlated (r>0.98). Finally, the HD panel was filtered to select only SNPs that were also present in the 50K panel as described by Sollero et al. (2014). Missing genotypes

were imputed for animals using FImpute (Sargolzaei et al., 2011) and after various quality control edits, 41,045 SNP markers (78%), including 136 sires, 2,803 BO and 652 HH yearling bulls, steers and heifers with tick count records, remained to estimate genomic relationship coefficients between animals.

The 4,363 animals having records were born between 2008 and 2011 and originated from 197 sires and 3,966 dams with 10 generations of pedigree depth. Pedigree information recovered from historical breeding records comprised 11,967 animals and was highly incomplete due to use of multiple-sire matings. This resulted in 65% of the animals with tick count records having unknown paternity. For genotyped parent-progeny pairs and trios (sire, dam and progeny) pedigree errors were checked by the percentage of Mendelian conflicts (Wiggans et al., 2009) using seekparentf90 software (http://nce.ads.uga.edu), with maximum tolerance of 1% to allow for genotyping errors. If a parent-progeny pair conflict was found or if one or neither parent had been genotyped, the animal's genotypes were compared with those of every other animal to determine if there was a parent-progeny relationship with an animal in the dataset. Unique putative parents of the appropriate sex with less than 1% Mendelian conflicts and suitable birthdates were designated as true parents, when identified.

When a true sire was not identified with the above describe procedure, the simulated annealing algorithm in the MOL_COANC software (Fernandez and Toro, 2006) was used to recreate half-sib families based on observed genomic relationships for genotyped animals born to multiple-sire matings. Pedigrees were reconstructed by creating a 'virtual' ancestor for each identified half-sib family. A total of 12,754 animals remained after pedigree reconstruction and pruning with a detailed breakdown provided in Table 1.

Table 1 Pedigree structure as defined by parentage certainty

	With tick counts	Without tick counts	Total
Both parents known	3,673	1,937	5,610
Both parents unknown	0	4,550	4,550
Only sire known	11	19	30
Only dam known	679	1,885	2,564
Total	4,363	8,391	**12,754**

Statistical modeling

Consider the following conventional single step genomic-based BLUP (GBLUP):

$$y_{ijk} = \mathbf{x'}_j \boldsymbol{\beta} + w_i + a_j + c_j + e_{ijk}. \quad [1]$$

Here y_{ijk} is the k^{th} phenotypic record of animal *j* recorded within CG *i*; **β** is the vector of fixed effects that includes an overall intercept; linear regression coefficients for Nellore breed proportion, heterozygosity and recombination loss as well as linear and quadratic regression coefficients on age of calf; $\mathbf{x'}_j$ is the known incidence row vector of covariates connecting **β** to y_{ijk}; w_i is the random effect of CG *i* (*i*=1,2,…,146 levels); a_j is the random additive genetic effect of animal *j*; c_j is the random permanent environment effect of animal *j* and e_{ijk} is the residual error.

Genomic information was combined together with pedigree information using single step procedure in which genetic relationships based on pedigree are adjusted for deviations due to genomic information (Legarra et al., 2009; Misztal et al., 2009; Aguilar et al., 2010a,b) and the following distributional assumptions were assumed: $\mathbf{w} = \{w_i\} \sim N(\mathbf{0}, \mathbf{I}\sigma^2_w)$, $\mathbf{a} = \{a_j\} \sim N(\mathbf{0}, \mathbf{H}\sigma^2_a)$ and $\mathbf{c} = \{c_i\} \sim N(\mathbf{0}, \mathbf{I}\sigma^2_c)$ and $\mathbf{e} = \{e_{ijk}\} \sim N(0, \mathbf{I}\sigma^2_e)$ by which $\sigma^2_w, \sigma^2_a, \sigma^2_c$ and σ^2_e represent, variances due to CG, additive genetics, permanent environment and residual terms, respectively. Here, **I** is the identity matrix and **H** represents the genetic relationships based on pedigree adjusted for deviations due to genomic information as Aguilar et al. (2010a):

$$\mathbf{H} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{G} - \mathbf{A}_{22} \end{bmatrix} = \mathbf{A} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G} - \mathbf{A}_{22} \end{bmatrix}, \quad [2]$$

where subscripts 1 and 2 represent ungenotyped and genotyped animals, respectively **A** is a numerator relationship matrix for all animals, **G** is a genomic relationship matrix constructed based on VanRaden (2008) which use the allele frequencies estimated from the genotyped individuals (Aguilar et al., 2010).

The inverse of **H** was obtained using preGSf90 software (http://nce.ads.uga.edu) and incorporate as a user defined covariance structure $\mathbf{H^{-1}}$ in INTERGEN software (Cardoso, 2013; adjustments shown in appendix) to combine genomic information with one-step reaction norm models as described hereafter.

Genomic single step linear reaction norm model (GLRNM)

The model proposed by Su et al. (2006) is purely Bayesian, since covariates associated with the reaction norm are treated as unknown, thereby allowing inference for all unknowns together within a one-step linear reaction norm model (GLRNM):

$$y_{ijk} = \mathbf{x'}_j \boldsymbol{\beta} + w_i + a_j + b_j w_i + c_j + d_j w_i + e_{ijk}. \quad [3]$$

This model can be rewritten in matrix notation as below (Su et al, 2006):

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Pw} + \mathbf{Z_a a} + \mathbf{Z_b b} + \mathbf{Z_c c} + \mathbf{Z_d d} + \mathbf{e}, \quad [4]$$

in which $\mathbf{y} = \{y_{ijk}\}$ is the *nx1* vector of observations, $\boldsymbol{\beta}$ is the fixed effects vector of order *p*, $\mathbf{w} = \{w_i\}_{i=1}^{n_w}$ is the vector of environmental effects, $\mathbf{a} = \{a_j\}_{j=1}^{q}$ is the vector of random genetic intercepts, $\mathbf{b} = \{b_j\}_{j=1}^{q}$ is the vector of random genetic slopes, $\mathbf{c} = \{c_j\}_{j=1}^{q}$ is the vector of random permanent environment intercepts, $\mathbf{d} = \{d_j\}_{j=1}^{q}$ is the vector of random permanent environment slopes and **e** is the *nx1* vector of residuals. Furthermore, **X**, **P**, $\mathbf{Z_a}$ and $\mathbf{Z_c}$ are known incidence matrices, whereas the row address of matrices $\mathbf{Z_b}$ and $\mathbf{Z_d}$ has exactly one element equal to the effect of the environmental covariate (w_i or an estimate of w_i) for that CG in the row address of the observation, will all other elements in that row equal to zero (Su et al., 2006) and **e** is the vector of random residual.

Prior distributional specifications

To infer environmental sensitivities using a hierarchical Bayesian model, three stages are required: the first stage defines the distribution of the observed data conditional on all other parameters (Su et al., 2006):

$$\mathbf{y}|\boldsymbol{\beta},\mathbf{w},\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{R} \sim \text{N}\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{P}\mathbf{w} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_b\mathbf{b} + \mathbf{Z}_c\mathbf{c} + \mathbf{Z}_d\mathbf{d}, \mathbf{R}\right). \quad [5]$$

A heteroskedastic residual alternative based on residual variances subclasses determined by a decile-based classification of $\hat{w}_i$ was implemented. That is, CG were ordered into one of 10 categories based on decile delimiters of $\hat{w}_i$ obtained from GBLUP, such that $\mathbf{R} = diag\left(\mathbf{I}_{n_k}\sigma_e^2\gamma_k\right)$ where order n_k denotes the number of records delimited by deciles k-1 and k, being respectively 1,157, 1,174, 1,047, 765, 1,188, 1,192, 1,208, 918, 1,150 and 874, respectively, for k = 1,2,…10. This strategy was also considered by Cardoso and Tempelman (2012).

The second stage is represented by the prior distributions of the location parameters, as follows:

$$\beta \sim \text{p}(\boldsymbol{\beta}) \quad [6]$$

$$\text{w}|\,\sigma_{\text{w}}^2 \sim \text{N}(0, \mathbf{I}\sigma_{\text{w}}^2) \quad [7]$$

$$\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} \sim \text{N}\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\sigma_a^2 & \sigma_{ab}\\ \sigma_{ab} & \sigma_b^2\end{bmatrix} \otimes \mathbf{H}\right) \quad [8]$$

$$\begin{bmatrix}\mathbf{c}\\ \mathbf{d}\end{bmatrix} \sim \text{N}\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\sigma_c^2 & \sigma_{cd}\\ \sigma_{cd} & \sigma_d^2\end{bmatrix} \otimes \mathbf{I}\right), \quad [9]$$

where $\text{p}(\boldsymbol{\beta}) \propto 1$, σ_{w}^2 is the environmental effect variance; σ_{a}^2 and σ_{b}^2 are, respectively, the additive genetic variances due to the reaction norm intercept and slope; σ_{c}^2 and σ_{d}^2 are, respectively, permanent environment variances due to reaction norm intercept and slope, σ_{ab} is the genetic covariance between reaction norm intercept and slope, σ_{cd} is the permanent environment covariance between reaction norm, intercept and slope.

Then $\mathbf{r}_{\mathbf{ab}} = \sigma_{\text{ab}} / \sqrt{\sigma_{\text{a}}^2 * \sigma_{\text{b}}^2}$ and $\mathbf{r}_{\mathbf{cd}} = \sigma_{\text{cd}} / \sqrt{\sigma_{\text{c}}^2 * \sigma_{\text{d}}^2}$ are the corresponding genetic and permanent environment correlations, respectively.

Finally, the third stage was based on specifying a scaled inverse (IG) distribution for the variance of the contemporary group effects; i.e. $\sigma^2_{\mathrm{w}}|\alpha_{\mathrm{w}},\beta_{\mathrm{w}} \sim \mathrm{IG}(\alpha_{\mathrm{w}}=1,\beta_{\mathrm{w}}=2800)$ where the mean of this distribution is:

$$\mathrm{E}\left(\mathrm{w}|\alpha_{\mathrm{w}},\beta_{\mathrm{w}}\right)=\frac{\Gamma(\alpha_{\mathrm{w}}+1)}{\beta_{\mathrm{w}}\Gamma(\alpha_{\mathrm{w}})}=\frac{\alpha_{\mathrm{w}}}{\beta_{\mathrm{w}}}. \quad [10]$$

Similarly, we specify $\sigma^2_{\mathrm{e}}|\alpha_{\mathrm{e}},\beta_{\mathrm{e}} \sim \mathrm{IG}(\alpha_{\mathrm{e}}=1,\beta_{\mathrm{e}}=600)$.

Likewise, an inverted Wishart distribution (IW) prior distribution was presumed, respectively, for the permanent environment and additive genetic covariance matrices:

$$\mathbf{U_0}=\begin{bmatrix}\sigma^2_{\mathrm{c}} & \sigma_{\mathrm{cd}}\\ \sigma_{\mathrm{cd}} & \sigma^2_{\mathrm{d}}\end{bmatrix} \sim \mathrm{IW}(\mathbf{T_0},\mathrm{v}) \quad [11]$$

$$\mathbf{G_0}=\begin{bmatrix}\sigma^2_{\mathrm{a}} & \sigma_{\mathrm{ab}}\\ \sigma_{\mathrm{ab}} & \sigma^2_{\mathrm{b}}\end{bmatrix} \sim \mathrm{IW}(\mathbf{T_0},\mathrm{v}). \quad [12]$$

Following Cardoso and Tempelman (2012), *v*=4 represents a presumed known degrees of freedom (DF), $\mathbf{T_0}=\begin{bmatrix}200 & 4\\ 4 & 0.1\end{bmatrix}^{-1}$ a presumed scale matrix, and $\mathrm{E}\left(\mathbf{G_0}\right)=\frac{\mathbf{T_0^{-1}}}{\mathrm{v}-\mathrm{p}-1}$ is the prior mean for *v>p+1*, where *p* is the number of parameters.

Furthermore, additional hierarchical specifications were required, i.e., $\gamma_{\mathrm{k}}|\alpha_\gamma \sim \mathrm{p}\left(\gamma_{\mathrm{k}}|\alpha_\gamma\right)=\mathrm{IG}\left(\alpha_\gamma,\alpha_\gamma-1\right)$, *k*=1,2,…,10 (Cardoso et al., 2005; Kizilkaya and Tempelman, 2005; Cardoso and Tempelman, 2012). In addition, the prior p(α_γ) on α_γ was a gamma with shape and scale hyperparameters values of 0.03 and 0.01, respectively (Cardoso and Tempelman, 2012). This assumption leads to a prior mean of α_γ equal to 3 ($\mathrm{E}(\alpha_\gamma)=3$) and a large prior variance ($\mathrm{var}(\alpha_\gamma)=300$) (Cardoso et al., 2005).

Bayesian inference and model comparison

Analyses were conducted using Bayesian analyses, specifically Monte Carlo Markov Chain (MCMC), to sample all parameters from their fully conditional posterior distributions. Gibbs sampling was generally used except

for the w_i's and α_γ. MCMC sampling of these parameters required a random walk Metropolis-Hastings steps due to all of their fully conditional posterior distributions being unrecognizable (see Cardoso and Tempelman, 2012 for details).

MCMC was implemented by INTERGEN software (Cardoso, 2013) by saving every 10 cycles from a total of 1,000,000 cycles, after 100,000 cycles of burn-in. Global convergence was checked using the Geweke's Z criterion (Geweke, 1991) applied to the conditional distribution of the data, as proposed by Brooks and Roberts (1998). In addition, visual inspection of trace plots was conducted and a minimum effective sample size of 100 for all unknown parameters was always obtained as well.

In order to access the goodness of the fit, Deviance Information Criterion (DIC) was used (Spiegelhalter et al., 2002):

$$DIC = \overline{D}(\theta) + p_D = 2\overline{D}(\theta) - D(\bar{\theta}), \quad [13]$$

where $\overline{D}(\theta) = E_{\theta|y}[D(\theta)]$ is the posterior expectation of Bayesian Deviance; $p_D = \overline{D}(\theta) - D(\bar{\theta})$ corresponds the penalty for increasing model complexity where θ is the model parameters vector and $D(\bar{\theta})$ is the Bayesian Deviance as a function of the posterior mean of the parameters. Smaller values of DIC indicate better-fitting model.

Genetic parameters estimation

The additive genetic variance for a specific environment *i* with effect w_i were obtained as follows:

$$\sigma^2_A \mid w_i = var\left(a_j + b_j w_i\right) = \sigma^2_a + w_i^2\sigma^2_b + 2w_i\sigma_{ab}. \quad [14]$$

Thus, the heritability and repeatability (*r*) for a specific environment was determined as:

$$h^2_A \mid w_i = \frac{\sigma^2_a \mid w_i}{\sigma^2_a \left| w_i + \sigma^2_c \right| w_i + \sigma^2_e \mid w_i} \quad [15]$$

and $r \mid w_i = \frac{\sigma^2_a \mid w_i + \sigma^2_c \mid w_i}{\sigma^2_a \left| w_i + \sigma^2_c \right| w_i + \sigma^2_e \mid w_i}$, [16]

respectively, where $\sigma_c^2|w_i$ and $\sigma_e^2|w_i$ are permanent environment and residual variances in environment i, respectively, and $\sigma_{e}^{2} \mid w_i = \sigma_{e}^{2} * \gamma_{k:i}$, where *k*:*i* denotes the decile-based classification *k* for CG *i*.

Furthermore, the genetic covariance between two environmental gradients based on covariate values w_i and $w_{i'}$ was calculated by:

$$\mathrm{cov}_A\left(a_j + b_j w_i, a_j + b_j w_{i'}\right) = \sigma_a^2 + \left(w_i + w_{i'}\right)\sigma_{ab} + w_i w_{i'}\sigma_b^2. \quad [17]$$

Then, the corresponding correlation between two specific environments was calculated as described below:

$$r_A\left(a_j + b_j w_i, a_j + b_j w_{i'}\right) = \frac{\mathrm{cov}_a\left(a_j + b_j w_i, a_j + b_j w_{i'}\right)}{\sqrt{(\sigma_a^2 + w_i^2\sigma_b^2 + 2w_i\sigma_{ab})(\sigma_a^2 + w_{i'}^2\sigma_b^2 + 2w_{i'}\sigma_{ab})}}. \quad [18]$$

Genomic estimated breeding values over environments

The genomic estimated breeding value (GEBV) of sire *j* specific to a given environment *i* was obtained by $\mathrm{GEBV}_j \mid w_i = a_j + b_j w_i$ (Cardoso and Tempelman, 2012). Values of b_j close to zero indicate that g_j is relatively constant across various environments (w_i) such that sire *j* has a robust genotype to environmental changes. On the other hand, a sensitive genotype has a large b_j and change substantially their performance on the environmental gradient (Falconer, 1990).

The sire breeding value estimates were compared by the ranking of the animals obtained by all tested models, for low, medium and high environmental levels. These values were respectively defined by the values representative of the 10th, 50th and 90th percentiles of w_i. Potential differences in re-ranking of sires for selection by these models were also determined by the Spearman correlation between breeding values obtained from those ranked among all animal as well as the top 10% (60) sires having large numbers (>12) of progeny for tick resistance between low, medium and high environmental levels.

Cross validation study

Cross validation prediction accuracy was evaluated by two different five-fold cross validation strategies: one strategy was based on the K-means procedure of Saatchi et al. (2011) that minimizes genetic ties between training and validation subsets. The other strategy was based on random partitioning of training and validation data sets for comparative purposes.

Cross validation accuracy ($r_{y,\hat{y}}$) was defined as the correlations between observed (y) and predicted phenotypes ($\hat{y}$) in the validation datasets, based on estimates derived from training datasets. The accuracy of the GEBVs was compared between the all tested models. Therefore, all models were fitted under the same cross validation scheme, i.e., considering the random intercept for GBLUP and random intercept and slope for GLRNM.

To implement a cross validation study, analysis using a traditional relationship matrix (**A**) was also performed: the phenotypes of the individuals comprising validation data set were assumed to be missing and the accuracy of estimated breeding values (EBV) were calculated using the same formula described before.

Finally, comparisons between models were based on randomized completes block design (RCBD) analysis that has each treatment (models fitted) applied in each block (folds) of correlations as response variables. Letting y_{ij} denote the correlations that received m fitted models and f folds, the equation for the model is $y_{fm} = \mu + \tau_m + b_f + e_{fm}$, where m=4, f=5 μ and τ_i are fixed parameters such that the mean for the m^{th} model is $\mu_m = \mu + \tau_m$, b_f is the random effect associated with the f^{th} fold, e_{fm} is the random error associated with the experimental unit in fold f that received model m.

In addition, block (fold) and error effects are assumed to be normally and independently distributed with mean 0 and variance σ_b^2 and σ_e^2, respectively, i.e., $b_j \sim$ iid N(0, σ_b^2) and $e_{fm} \sim$ iid N(0, σ_e^2).

RESULTS AND DISCUSSION

Model comparison via DIC

The non reaction norm models (BLUP and GBLUP) seemed to be poorer fitting in comparison with its one-step extension LRNM or GLRNM based on a smaller DIC value (Table 2). Hence, it would reinforce the importance of consider G*E for genetic evaluations of tick resistance in Hereford and Braford beef cattle. In addition, models under genomic approach presented smaller DIC values compared to their respective model based on pedigree information confirming the importance to incorporate marker information in genetic evaluations (Table 2).

Table 2 Deviance Criterion Information (DIC) value and model ranking of all models

Model[1]	DIC Value	Ranking
BLUP	3647.73	4º
GBLUP	3115.54	3º
LRNM	2443.07	2º
GLRNM	2055.09	1º

[1] BLUP = Linear animal model; GBLUP = Genomic linear animal model; LRNM= one-step linear reaction norm model; GLRNM = Genomic one-step linear reaction norm model.

Variance components and genetic parameters under G*E interactions

The genomic reaction norm model GLRNM presented lower intercept and slope genetic variance estimates in comparison with the models that included the pedigree-based relationship matrix **A** (LRNM). However, for the permanent environment effect, GLRNM showed higher values (Table 3).

Despite lower genetic values for GLRNM, estimates from this model seem to be more accurate due to lower standard deviations for both effects. Lower genetic variance estimates were also reported by Veerkamp et al. (2011) using a simulated data and multi-trait animal model fitted to milk yields, body and dry matter to compare **A**, **G** and **H** matrices. According to these authors, there are several reasons to understand why estimated genetic variances differ between models using **A** or **H** matrices, such as due to **A** and **H** are simply different in scale, the sum of eigenvalues and also because the relationships in **A** are based on identity by descent and **G** based on identity by state probabilities. On the other hand, in the context of reaction norm models, Silva et al. (2014) reported similar variance components for the

intercept and slope using **A** and **G** matrices fitting a two-step random regression reaction norm model in pigs.

In general, the residual class variances were slightly higher for GLRNM compared to LRNM and no clear pattern was noticeable for both models (Table 3). Cardoso and Tempelman (2012) working with G*E models in post-weaning gains in Angus Cattle also observed that the residual variance did not increase categorically over the environmental gradient.

Under the non genomic approach (**A** matrix), estimated correlations between intercept and slope for both sets of random effects (i.e., additive genetic and permanent environment effects) were positive but characterized by great uncertainty as indicated by the standard deviations for the permanent environment effect (Table 3). However, the combined analyses using all phenotypes and genomic relationships for genotyped animals presented positive correlations with high and moderate magnitude for additive genetic and permanent environment effects, respectively. It would indicate that animals with small values for genetic merit for a specific tick infestation level tend to have better performance in less changeable environments. These results are in agreement with previous studies under pedigree-based (Shariati et al., 2007; Mattar et al., 2011; Cardoso and Tempelman, 2012) and genomic approaches (Silva et al., 2014).

The environmental variance were slightly higher in LRNM (Table 3) and the estimated EG posterior means ($\hat{w}_i$) were similar ranging from -0.761 to 0.728 and from -0.735 to 0.712 for LRNM and GLRNM, respectively. Once $\hat{w}_i$ GLRNM seemed to be more accurate, corresponding values of $\hat{w}_i$, -0.396, -0.241, -0.112, -0.057, 0.020, 0.109, 0.170, 0.258 and 0.320, were assumed as the cutoff points for the ten residual heteroscedastic discrete subclasses, respectively, in order from the 10^{th} to 90^{th} percentiles to become comparable the genetic parameters estimates from both models.

The heritability estimates (h^2) between 10^{th} and 90^{th} percentiles as a function of CG levels are shown in Figure 1a (range 0.07 - 0.46). Similar heritability estimates have been reported in literature using logarithmic transformation of the observed data (Budeli et al., 2009; Oliveira et al., 2012; Ayres et al., 2013). Additionally, h^2 were higher for LRNM in all tick infestation levels (Figure 1a). However, h^2 from GLRNM seem to more

accurate due to smaller standard deviations. Veerkamp et al. (2011) found a clearly improvement in accuracy of h^2 using a SNP-based relationships even though h^2 decreased. In addition, It might happen because some the environmental effects confounded with the pedigree relationships are not included in the genetic variance using genomic relationships (Veerkamp et al., 2011).

This remarkable difference between **A** and **H** matrices approaches might be due to differences in variance components estimates and/or due to heterogeneity of the residual variance assumed in the fitted reaction norm models. These results diverged from those reported by Forni et al. (2011) and Silva et al. (2014) in which h^2 were similar for **A** and **G** matrices approaches in pigs using an animal model and a two-step random regression reaction norm model, respectively.

The repeatability estimates varied along the EG (range 0.20-0.60) and were, in general, similar under both approaches (Figure 1b). These results showed the importance in consider the environment permanent effect once that effect presented difference among animal's performance. These results demonstrate the particular importance of modeling permanent environmental effects in harsh environments such that more resistant animals are more likely to maintain a consistent performance even in harsher environments than in favorable environments.

Genetic correlations along the EG were remarkably low between extreme EG, whereas closest EG were high regardless whether **A** or **H** relationship matrices was implemented (Figure 2). In addition, negative correlations could be observed between extreme EG under genomic approach. These values indicate substantial G*E for tick resistance. Cardoso and Tempelman (2012) also reported low genetic correlations between extreme EG under non genomic approach for post-weaning in Angus cattle. Moreover, Figure 2 also shows if selection would occur in an average environment (CG level = 0), the responses in extreme environments, i.e., the poorest and best environments would be approximately 59% and 86% for **A**, as well as 48% and 86% for **H** of the direct responses obtained under selection in these extreme environments. Therefore, it reiterates the importance to select animals considering G*E.

Differences in environmental sensitivity resulted in re-rakings of top 10% most used sires (>12 progeny) at different CG levels for both approaches, but especially for a non genomic approach. Under non genomic approach the sires presented more stable breeding values behavior across the EG (Figure 3). Nevertheless, Figure 3 also demonstrates that genetic merit also depends upon EG under genomic approach and that differences in GEBV between animals decrease with a low EG. It further indicates the difficulty to identify superior breedstock in better (i.e., low tick infestation) environments.

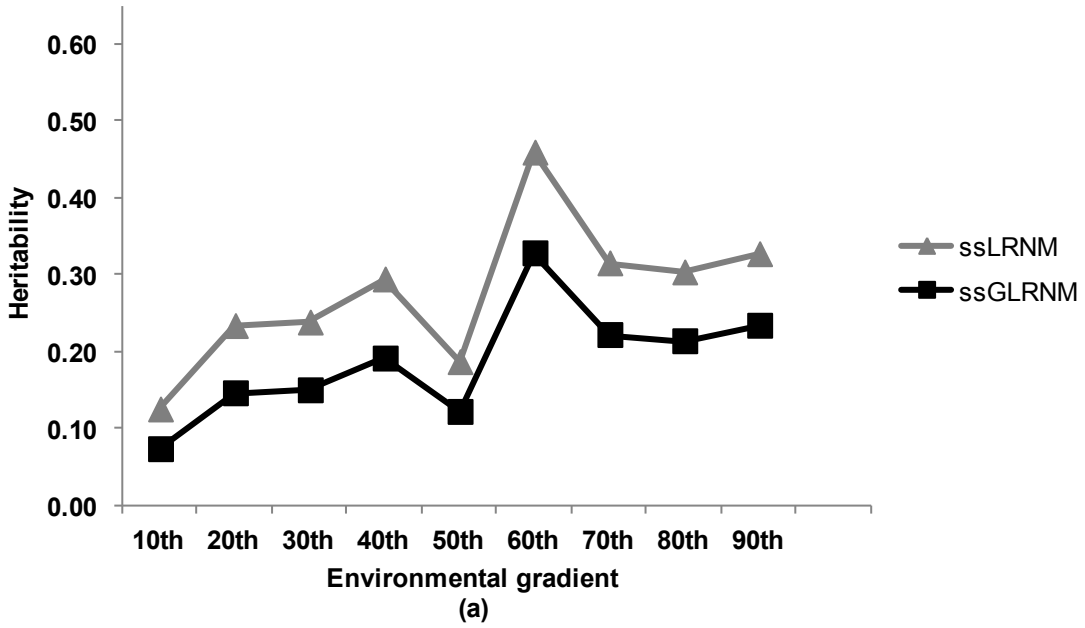
Rank correlations among posterior means of the genetic merit predictions (a_j) under both approaches, obtained by the animal models (BLUP and GBLUP) with those ($g_j|w_i$) obtained by their extensions (LRNM and GLRNM) are shown in Table 4. Those values were above 0.60 with lower values across **A** and **H** relationship matrices. It indicates that rankings of animals for selection would be similar under different environmental conditions within and different across approaches.

Furthermore, once correlations across **A** and **H** matrices were lower than within approaches, the impact on rankings of introducing marker information is relevant because correlations within genomic were higher than non genomic approach (Table 5). Finally, losses on selection precision by employing a traditional animal model would not be expected to be substantial in both approaches due to similarity correlation of magnitude in low, medium and also high EG. Cardoso and Tempelman (2012) reported similar results under non genomic models where would be expected losses by employing conventional models such as BLUP in this study across EG. Although we observed slight correlation values when we filtered sires with large number of progeny (top 10%; Table 4) the same pattern was observed as mentioned before for the total number of animals in data set.

Table 3 Posterior means and standard deviations for the variance components and genetic and permanent environment correlations of the one-step reaction norm model using pedigree (**A** matrix) and pedigree plus marker information (**H** matrix) relationship matrices.

Parameter[1]	Model[2]	
	A matrix	**H matrix**
σ^2_a	0.024±0.002	0.016±0.002
σ^2_b	0.037±0.023	0.030±0.011
σ_{ab}	0.013±0.006	0.011±0.004
σ^2_c	0.007±0.002	0.012±0.002
σ^2_d	0.074±0.027	0.091±0.018
σ_{cd}	0.003±0.006	0.009±0.005
σ^2_w	0.097±0.012	0.093±0.011
σ^2_{e1}	0.122±0.008	0.129±0.008
σ^2_{e2}	0.056±0.005	0.057±0.005
σ^2_{e3}	0.062±0.004	0.065±0.004
σ^2_{e4}	0.047±0.003	0.049±0.003
σ^2_{e5}	0.099±0.005	0.102±0.005
σ^2_{e6}	0.023±0.002	0.022±0.002
σ^2_{e7}	0.053±0.003	0.052±0.003
σ^2_{e8}	0.062±0.004	0.062±0.004
σ^2_{e9}	0.057±0.003	0.056±0.003
σ^2_{e10}	0.058±0.003	0.056±0.004
r_{ab}	0.476±0.227	0.518±0.184
r_{cd}	0.157±0.259	0.265±0.153

[1] σ^2_a = reaction norm intercept genetic variance; σ^2_b = reaction norm slope genetic variance; σ_{ab} = genetic covariance between intercept and slope; σ^2_c = reaction norm intercept permanent environment variance; σ^2_d = reaction norm slope permanent environment variance; σ_{cd} = permanent environment covariance between intercept and slope; σ^2_w = environmental variance; σ^2_{e1} = residual class 1; σ^2_{e2} = residual class 2; σ^2_{e3} = residual class 3; σ^2_{e4} = residual class 4; σ^2_{e5} = residual class 5; σ^2_{e6} = residual class 6; σ^2_{e7} = residual class 7; σ^2_{e8} = residual class 8; σ^2_{e9} = residual class 9; σ^2_{e10} = residual class 10; r_{ab} genetic correlation between intercept and slope; r_{cd} permanent environment correlation between intercept and slope.



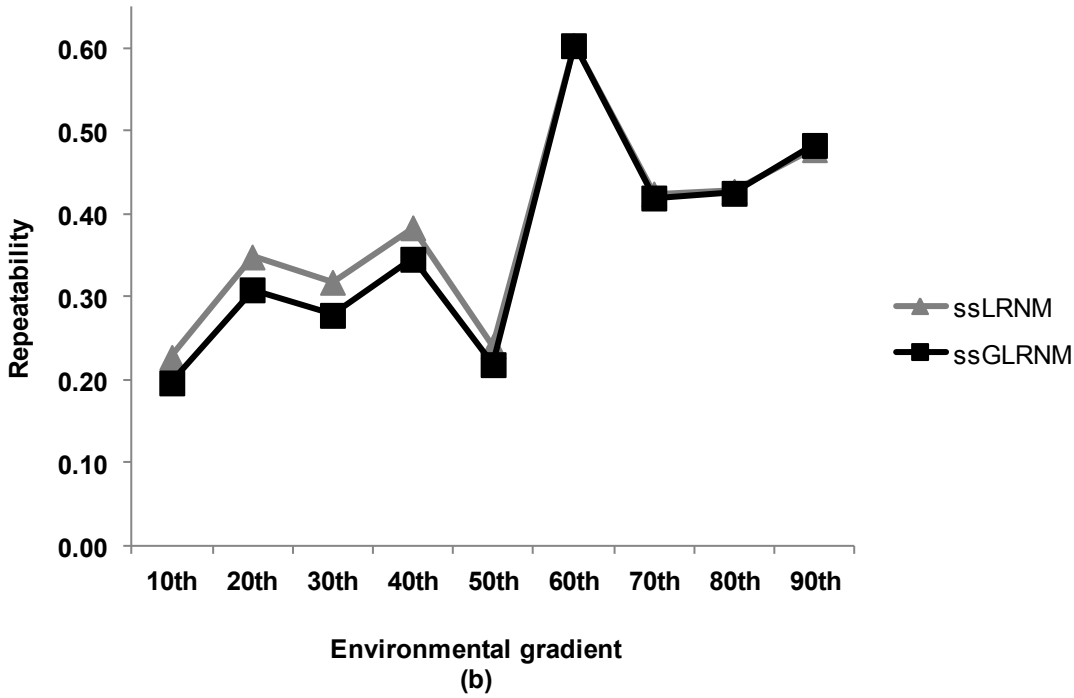


Figure 1 Posterior means heritabilities (a) and repeatabilities (b) for environments between the 10^{th} and 90^{th} percentiles of the environmental gradient obtained for Hereford and Braford tick counts by LRNM = one-step linear reaction norm model; GLRNM = genomic one-step linear reaction norm model.

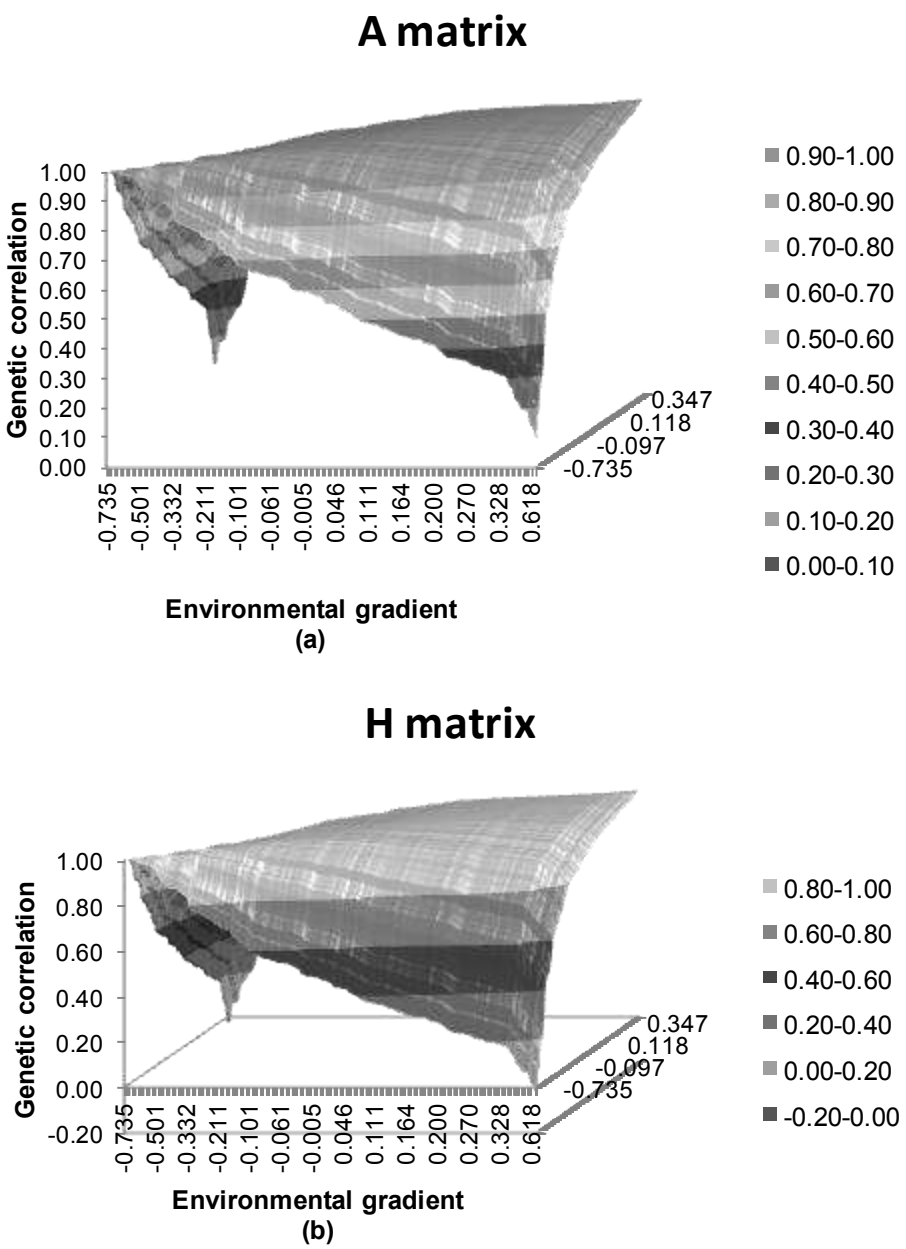


Figure 2 Genetic correlations for Hereford and Braford tick resistance performance on different environmental conditions obtained by the one-step linear reaction norm models considering pedigree (a) and pedigree + genomics (b).

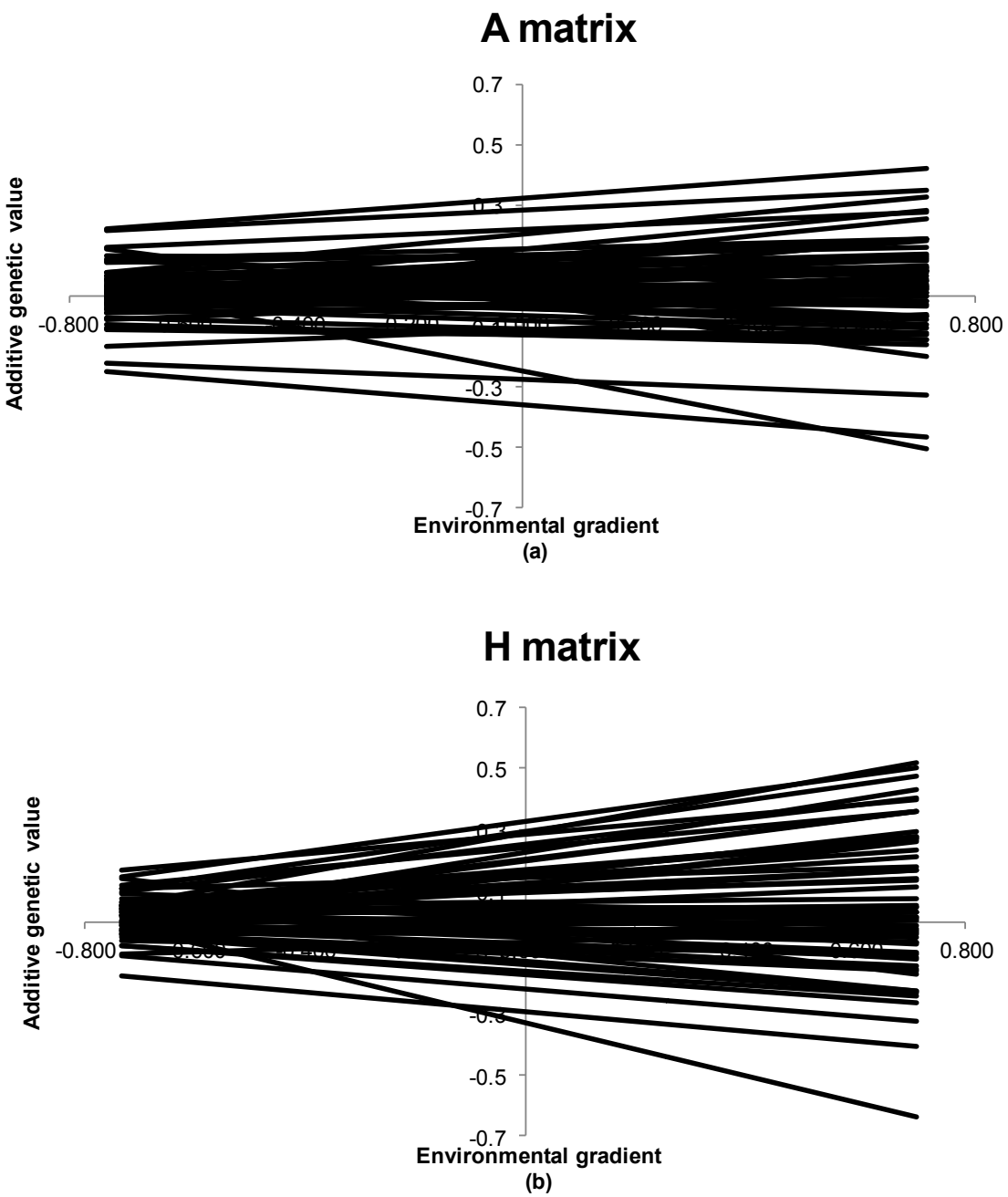


Figure 3 Genetic tick resistance reaction norms of 10% most used (large number of progeny; >12) Hereford and Braford sires obtained by the one-step linear reaction norm models considering pedigree (a) and pedigree + genomics (b).

Table 4 Spearman rank correlations[a] among posterior means genetic values for tick counts of Hereford and Braford beef cattle at different environmental (tick infestations) levels obtained by the linear conventional animal and reaction norm models

Model[1] Environmental level[2,3]	BLUP (Ov.)	LRNM (LTI)	LRNM (MTI)	LRNM (HTI)	GLRNM (LTI)	GLRNM (MTI)	GLRNM (HTI)	GBLUP (Ov.)
BLUP (Ov.)		0.96	0.97	0.96	0.69	0.64	0.62	0.68
LRNM (LTI)	0.90		0.97	0.94	0.72	0.65	0.62	0.67
LRNM (MTI)	0.93	0.94		0.99	0.71	0.68	0.66	0.68
LRNM (HTI)	0.91	0.88	0.99		0.70	0.68	0.67	0.68
GLRNM (LTI)	0.66	0.72	0.72	0.70		0.96	0.93	0.95
GLRNM (MTI)	0.61	0.62	0.68	0.69	0.96		0.99	0.97
GLRNM (HTI)	0.59	0.58	0.66	0.69	0.93	1.00		0.95
GBLUP (Ov.)	0.68	0.61	0.66	0.67	0.93	0.96	0.95	

[a] Correlations between all animal above the diagonal and between the most used sires (larger number of progeny) below the diagonal; [1] BLUP = Linear animal model; LRNM= one-step linear reaction norm model; GLRNM = Genomic one-step linear reaction norm model; GBLUP = Genomic linear animal model; [2] LTI = low tick infestation; MTI = medium tick infestation; HTI = high tick infestation; Ov.= overall; [3]LTI= 10^{th} (-0.396); MTI= 50^{th} (0.020); High= 90^{th} (0.320) percentiles of the environmental gradient.

Prediction ability via cross validation

Cross-validation prediction accuracies ($r_{y,\hat{y}}$) within each of tested models (BLUP, GBLUP, GLRNM and LRNM) were effective, being higher than 0.55 in K-means and 0.59 in random partitioning strategies (Figure 4).

Cross validation estimates were in average 0.66±0.02, 0.67±0.02, 0.67±0.02 and 0.66±0.02 for BLUP, GBLUP, GLRNM and LRNM, respectively, based on K-means partitioning, whereas GLRNM was 0.71±0.01 and tend to better compared to BLUP (0.67±0.01), GBLUP (0.70±0.01) and LRNM (0.70±0.01) based on random partitioning (Figure 4). However, no statistical significance was reported between GLRNM and LRNM (Table 5).

Perhaps a reason for no statistical difference between models via cross validation could be due to over-fitting, i.e., sometimes a better fitting does not necessarily mean better prediction. Bishop (2006) working in a simple regression problem, data was generated from the function sin(2πX), fitted polynomials having orders equal to 0, 1, 3 and 9 to solve curve fitting problem, and noticed that even with the ninth order presented an excellent fit to the training data, passing exactly through each data point, the fitted curve oscillated wildly and gave a very poor representation of the function sin(2πX). Therefore, we can imply that even GLRNM presented a smaller DIC value,

this model does not necessarily predict like it fitted. Furthermore, in contrast with this study, Silva et al. (2014) found higher genomic prediction accuracies and genetic correlations for reaction norm models compared to a standard sire model in pigs.

Finally, smaller values for $r_{(y,\hat{y})}$ presented by K-means compared to random partitioning were expected and thereby reinforce that prediction accuracy deteriorate as the relationship between animals decrease (Saatchi et al., 2011). These authors reported lower accuracies for K-means partitioning compared to random partitioning for all 16 traits analyzed in American Angus beef cattle.

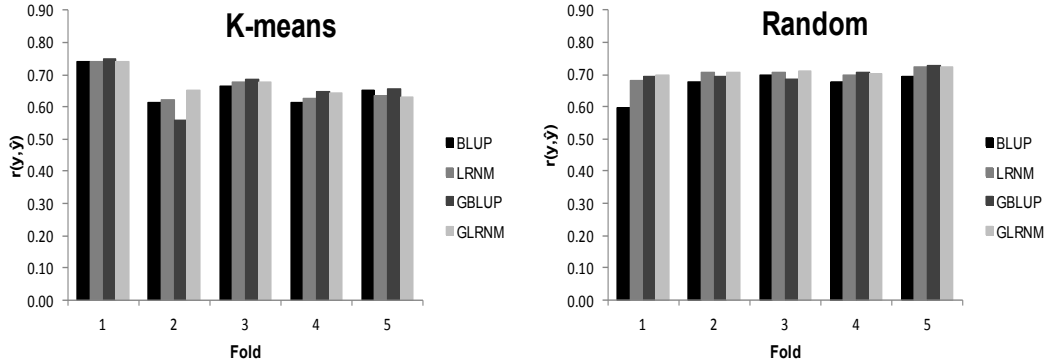


Figure 4 Accuracy of genomic selection obtained by K-means and random partitioning based on a five-fold cross validation study using genotype by environment interaction (LRNM; GLRM) and conventional animal models (BLUP; GBLUP)**.**

Table 5 *P*-values (Pr > |t|) in a t test for contrast of each difference of least square means for the accuracy $r_{(y,\hat{y})}$ between models for K-means and random partitioning using randomized completes block design (RCBD) analysis

Model[1]/Strategy	K-means	Random
BLUP vs GBLUP	0.2762	0.0139
BLUP vs GLRNM	0.5725	0.0010
BLUP vs LRNM	0.8273	0.0056
GBLUP vs GLRNM	0.5853	0.2241
GBLUP vs LRNM	0.3767	0.6701
GLRNM vs LRNM	0.7271	0.4183

[1] BLUP = Linear animal model; GBLUP = Genomic linear animal model; LRNM= one-step linear reaction norm model; GLRNM = Genomic one-step linear reaction norm model.

CONCLUSION

We inferred the presence of genotype by environment interaction for tick resistance in Hereford and Braford beef cattle based on genomic reaction norm models as well as these models appear to be an important and valid tool to account for genotype by environment interaction in genetic evaluation for tick resistance in Hereford and Braford beef cattle.

The results also suggest that marker information do not lead to higher accuracies of prediction which decreased as the tick infestation level increased and as the relationship between animals in training and validation datasets decreased.

ACKNOWLEDGMENTS

The authors thank Delta G Connection by providing the data used in this research; Embrapa South Livestock, Instituto Nacional de Investigacion Agropecuaria (INIA – Uruguay) and Michigan State University for theoretical and technical support; CAPES and CNPq by granting the scholarship. This project was supported by Agriculture and Food Research Initiative Competitive Grant no. 2011-67015-30338 from the USDA National Institute of Food and Agriculture.

REFERENCES

AGUILAR, I.; MISZTAL, I.; JOHNSON, D.L. et al. 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, 2010a.

AGUILAR, I.; MISZTAL I.; LEGARRA, A. et al. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. **Journal Animal Breeding and Genetics**, v.128, p.422-428, 2010b.

AMBROSINI, D.P.; CARNEIRO, P.L.S.; BRACCINI NETO, J. et al. Genotype × environment interaction for yearling weight in Polled Nellore cattle in Northeast Brazil. **Pesquisa Agropecuária Brasileira**, v.47, p.1489-1495, 2012.

AYRES, D.R.; PEREIRA, J.R.; BOLIGON, A.A. et al. Linear and Poisson models for genetic evaluation of tick resistance in cross-bred Hereford x Nellore cattle. **Journal of Animal Breeding and Genetics**, v.130, p.417-424, 2013.

BIEGELMEYER, P. **Genetic resistance to *Rhipicephalus* (Boophilus) *microplus* natural and artificial infestation in Hereford and**

Braford cattle. 2012, 96p. Dissertation (Master in Science) - Universidade Federal de Pelotas, Pelotas, RS, Brazil.

BISHOP, C.M. **Pattern Recognition and Machine Learning**. Vol. 1. Springer, New York, NY, USA, 2006.

BROOKS, S.P.; ROBERTS, G.O. Convergence assessment techniques for Markov Chain Monte Carlo. **Statistics and Computing**, v.8, p.319-335, 1998.

BUDELI, M.A.; NEPHAWE, K.A.; NORRIS, D. et al. Genetic parameter estimates for tick resistance in Bonsmara cattle. **South African Journal of Animal Science**, v.39, p.321-327, 2009.

BURROW, H.M. Variances and covariances between productive and adaptive traits and temperament in a composite breed of tropical beef cattle. **Livestock Production Science**, v.70, p.213-233, 2001.

CALUS, M.P.L.; GROEN A.F.; DE JONG G. Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. **Journal of Dairy Science**, v.85, p.3115-3123, 2002.

CARDOSO, F.F.; ROSA, G.J.M.; TEMPELMAN, R.J. Multiple-breed genetic inference using heavy-tailed structural models for heterogeneous residual variances. **Journal of Animal Science**, v.83, p.1766-1779, 2005.

CARDOSO, L.L.; BRACINI NETO, J.; CARDOSO F.F. et al. Hierarchical Bayesian models for genotype x environment interaction estimation via reaction norms applied to post-weaning gain of Hereford cattle. **Revista Brasileira de Zootecnia**, v.40, p.294-300, 2011.

CARDOSO, F.F.; TEMPELMAN, R.J. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. **Journal of Animal Science**, v.90, p.2130-2141, 2012.

CARDOSO, F.F. **Application of Bayesian inference in animal breeding using the Intergen program: Manual of version 1.22**. Documents 112, Embrapa South Livestock Bagé, RS, Brazil, 2013.

Clayton, D. snpStats: SnpMatrix and XSnpMatrix classes and methods. http://www-gene.cimr.cam.ac.uk/clayton. (Accessed 30 October 2012.), 2012.

CORRÊA, M.B.B.; DIONELLO, N.J.L.; CARDOSO, F.F. Genetic evaluation of Devon cattle using a reaction norms model. **Revista Brasileira de Zootecnia**, v.39, p.128-133, 2010.

FALCONER, D.S. Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. **Genetics Research**, v.56, p.57-70, 1990.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to Quantitative Genetics**. 4 ed. Longman Group Ltd, Harlow, UK, 1996.

FERNANDEZ, J.; TORO, M.A. A new method to estimate relatedness from molecular markers. **Molecular Ecology**, v.15, p.1657-1667, 2006.

FORNI, S.; AGUILAR, I.; MISZTAL, I. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. **Genetics Selection Evolution**, v.43:1, 2011.

FRISCH, J.E.; O'NEILL, C.J. Comparative evaluation of beef cattle breeds of African, European and Indian origins. 2. Resistance to cattle ticks and gastrointestinal nematodes. **Animal Science**, v.67, p.39-48, 1998.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: BERNARDO, J.M.; BERGER, J.O.; DAWID, A.P. et al. (Eds.). **Bayesian Statistics 4**. New York: Oxford University, 1991. 526p.

KIZILKAYA, K.; TEMPELMAN, R.J. A general approach to mixed effects modeling of residual variances in generalized linear mixed models. **Genetic Selection Evolution**, v.37, p.31-56, 2005.

KOLMODIN, R.; STRANDBERG, E.; MADSEN, P. et al. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. **Acta Agriculturae Scandinava Section A Animal Science**, v.52, p.11-24, 2002.

LEGARRA, A.; AGUILAR, I.; MISZTAL, I. A relationship matrix including full pedigree and genomic information. **Journal of Dairy Science**, v.92, p.4656-4663, 2009.

MATTAR, M.; SILVA, L.O.C.; ALENCAR, M.M. et al. Genotype by environment interaction for long-yearling weight in Canchim cattle quantified by reaction norm analysis. **Journal of Animal Science**, v.89, p.2349-2355, 2011.

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.; 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.

MOTA, R.R.; CARDOSO, F.F.; LOPES, P.S. et al. Hierarchical Bayesian models for genotype x environment estimates in tick resistance of Hereford and Braford beef cattle via reaction norms. ***Proceedings*** of 10th Simpósio Brasileiro de Melhoramento Animal, Uberaba, Brazil: SBMA, 2013.

OLIVEIRA, M.M.; GOMES, C.C.G.; ROSO, V.M. et al. Genomic selection for tick resistance in Braford and Hereford cattle ***Proceedings*** of 9th Simpósio Brasileiro de Melhoramento Animal, João Pessoa, Brazil: SBMA, 2012.

PÉGOLO, N.T.; OLIVEIRA, H.N.; ALBUQUERQUE, L.G. et al. Genotype by environment interaction for 450-day weight of Nelore cattle analyzed by reaction norm models. **Genetics and Molecular Biology**, v.32, p.281-287, 2009.

RECHAV, Y.; DAUTH, J.; ELS, D.A. Resistance of Brahman and Simmental cattle to southern African ticks. **Onderstepoort Journal of Veterinary Research**, v.57, p.7-12, 1990.

ROSO, V.M.; SCHENKEL, F.S. AMC – A computer program to assess the degree of connectedness among contemporary groups. ***Proceedings*** of 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil, 2006. (CD-ROM).

SAATCHI, M.; MCCLURE, M.C.; MCKAY, S.D. et al. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. **Genetics Selection Evolution**, v.43, 2011.

SARGOLZAEI, M.; CHESNAIS, J. SCHENKEL, F.; FImpute-An efficient imputation algorithm for dairy cattle populations. **Journal of Animal Science**, v.94 (E-Suppl. 1, p.421(abstr. 333), 2011.

SHARIATI, M.M.; SU, G.; MADSEN, et al. Analysis of milk production traits in early lactation using a reaction norm model with unknown covariates. **Journal of Dairy Science**, v.90, p.5759-5766, 2007.

SILVA, F.F., MULDER, H.A., KNOL, E.F. et al. Sire evaluation for total number born in pigs using a genomic reaction norms approach. **Journal of Animal Science**, v.92, p.3825-3834, 2014.

SOLLERO, B.P.; GOMES, C.C.G.; ROSO, V.M. et al. Goodness of Fit Comparisons among Five Bayesian Models in Genome-Wide Association of Tick Resistance in Brazilian Hereford and Braford Beef Cattle. ***Proceedings*** of 10th World Congress on Genetics Applied to Livestock Production, Vancouver, Canada, 2014.

SPIEGELHALTER, D.J.; BEST, N.G.; CARLIN, B.P. et al. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society Series B-Statistical Methodology**, v.64, p.583-616, 2002.

SU, G.; MADSEN, P.; LUND, M.S. et al. Bayesian analysis of the linear reaction norm model with unknown covariates. **Journal of Animal Science**, v.84, p.1651-1657, 2006.

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

VEERKAMP, R.F.; MULDER, H.A.; THOMPSON, R. et al. Genomic and pedigree-based genetic parameters for scarcely recorded traits when some animals are genotyped. **Journal of Dairy Science**, v.94, p.4189-4197, 2011.

WHARTON, R.H.; UTECH, K.B.W. The relation between engorgement and dropping of *Boophilus microplus* (*Canestrini*) (*Ixodidae*) to the assessment to tick numbers on cattle. **Journal Australian Entomologycal Society**, v.9, p.171-182, 1970.

WIGGANS, G.R.; SONSTEGARD, T.S.; VANRADEN, P.M. et al. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. **Journal of Dairy Science**, v.92, p.3431-3436, 2009.

Appendix A

Subroutine code to incorporate the inverse of the matrix inverse that combines pedigree and genomic relationships ($\mathbf{H}^{-1}$; Aguilar et al., 2010a) as a tested file in INTERGEN software (Cardoso, 2013).

```
subroutine add_g_usr(eff,text)
! generates contributions user defined matrix; if text is 'inv',
! this matrix is inverted
! Originally code from remlf90 by M.P.E.
!
integer :: eff,i,j,k,m,n,row,col,t1,t2,type=0,maxrow=0,maxcol=0,maxdiag=0
integer,parameter::upper=1,lower=2
real (rh) ::val,x,ident(nlev(eff)),sol(nlev(eff))
type (sparse_hashm)::a_usr
type (sparse_ija)::a_usr_ija
character (*)::text
!
if (round == start) then
!call init(a_usr)
call zerom(a_usr,nlev(eff))
n=0

do
read(io_off+eff,*,iostat=io)row,col,x
if (io /= 0) exit
n=n+1

! Only upper or lower half stored matrices are accepted
if (row /=col) then
if (type == 0) then
if (row < col) type = upper
if (row > col) type = lower
endif

if ((type == upper .and. row > col) .or. &
(type == lower .and. row < col)) then
print*,'Mixed upper and lower diagonals in g_usr_inv:',row,col,x
stop
endif
endif

call addm(x,row,col,a_usr)
if (row /=col) call addm(x,col,row,a_usr)
```

```
maxrow=max(maxrow,row)
maxcol=max(maxcol,col)
if (row==col) maxdiag=max(maxdiag,row)
enddo

df_random(eff)=nlev(eff)
print*,'largest row, column,diagonal:',maxrow,maxcol,maxdiag,df_random(eff)
!print*,'a_usr'
!call printm(a_usr)
!call init(ainv(eff))
call zerom(ainv(eff),nlev(eff))

if (nlev(eff) <= 20) then
print*,'user defined matrix'
call printm(ainv(eff))
endif
call reset(a_usr)
endif

do i=0,randomnumb(eff) - 1 !loops for correlated effects
do j=0,randomnumb(eff) - 1
do t1=1,ntrait
do t2=1,ntrait
if (pos_eff(eff+i,t1)/=0 .and. pos_eff(eff+j,t2)/=0) then
do k=1,ainv(eff)%nel
row=ainv(eff)%x(1,k)
if (row /=0) then
col=ainv(eff)%x(2,k)
x=ainv(eff)%x(3,k)
m=address1(eff+i,row,t1)
n=address1(eff+j,col,t2)
val=ginv(eff,t1+i*ntrait,t2+j*ntrait)*x
call addm(val,m,n,xx)
!if (m /=n) call addm(val,n,m,xx)
if (row /= col) call addm(val,n,m,xx)
endif
enddo
endif
enddo
enddo
enddo
enddo
end subroutine
```

Appendix B

In the following, we present *parameterfiles* incorporating genomic information ($\mathbf{H}^{-1}$) to run a classical animal model, frequently used quantitative genetic evaluation and also a one-step reaction norm model based on Bayesian methods. Examples including data, the $\mathbf{H}^{-1}$ or $\mathbf{A}^{-1}$ and parameter files are also available at the Embrapa Pecuária Sul website, at (http://www.cppsul.embrapa.br/unidade/servicos/intergen).

```
#Example 1– Animal Model using the matrix inverse that combines pedigree and genomic relationships (H inverse)
MCMC_CHAIN: TOTAL_CYCLES BURN_IN THINNING_INTERVAL
11000 1000 10
SEED
2
RESTART: Y/N? [CYCLE_TO_RESTART]
n
DATAFILE NAME N_RECORDS
animodel.dat 240
NUMBER_OF_TRAITS
1
NUMBER_OF_EFFECTS
2
OBSERVATION(S)
4
WEIGHT(S)

EFFECTS: POSITIONS_IN_DATAFILE NUMBER_OF_LEVELS TYPE_OF_EFFECT
SAVE_SAMPLES? [EFFECT NESTED]
1 300 cross n      #animal effect
#Practicing – Animal Model
3 2 cross n        #gender
RANDOM_RESIDUAL: TYPE PRIOR_DEGREES_OF_BELIEF
homogeneous 1
RESIDUAL_PRIOR_(CO)VARIANCES
150
RANDOM_GROUP
1
RANDOM_TYPE PRIOR_DEGREES_OF_BELIEF
user_file 1
PEDIGREEFILE: NAME N_ANIMAL N_GENETIC_GROUPS [N_BREEDS]
Hinv.txt
(CO)VARIANCES
100
```

```
# Example 2- One-step reaction norm model considering heterogeneous residual variance and using the matrix inverse that combines pedigree and genomic relationships (H inverse)
MCMC_CHAIN: TOTAL_CYCLES BURN_IN THINNING_INTERVAL
1000000 100000 10
SEED
123
RESTART: Y/N? [CYCLE_TO_RESTART]
n
DATAFILE NAME N_RECORDS
animodel.dat 10673
NUMBER_OF_TRAITS
1
NUMBER_OF_EFFECTS
12
OBSERVATION(S)
1
```

```
WEIGHT(S)

EFFECTS:        POSITIONS_IN_DATAFILE        NUMBER_OF_LEVELS        TYPE_OF_EFFECT
SAVE_SAMPLES [EFFECT NESTED]
4 1 cov n
5 146 unknowncov n
4 11965 cov n 2
5 11965 rnorm n 2
4 4363 cov n 3
5 4363 rnorm n 3
12 10 cross n 2
6 1 cov n
7 1 cov n
8 1 cov n
9 1 cov n
10 1 cov n
RANDOM_RESIDUAL: TYPE PRIOR_DEGREES_OF_BELIEF
structural 1
METROPOLIS_STEP_OF_STRUCTURAL_EFFECTS:
 ROUNDS_WITHIN_CYCLE TUNING_SKIP
1 10
NUMBER_OF_STRUCTURAL_EFFECTS
1
STRUCTURAL_EFFECTS: LINE_FROM_EFFECTS_SECTION SAVE_SAMPLES?
7 y
RESIDUAL_PRIOR_(CO)VARIANCES
0.0632
RANDOM_GROUP
2
RANDOM_TYPE PRIOR_DEGREES_OF_BELIEF
diagonal 1
PEDIGREEFILE: NAME N_ANIMAL N_GENETIC_GROUPS [N_BREEDS]

(CO)VARIANCES
0.1102
RANDOM_GROUP
3 4
RANDOM_TYPE PRIOR_DEGREES_OF_BELIEF
user_file 1
PEDIGREEFILE: NAME N_ANIMAL N_GENETIC_GROUPS [N_BREEDS]
Hinv.txt
(CO)VARIANCES
0.0127 0.0131
0.0131 0.0102
RANDOM_GROUP
5 6
RANDOM_TYPE PRIOR_DEGREES_OF_BELIEF
diagonal 1
PEDIGREEFILE: NAME N_ANIMAL N_GENETIC_GROUPS [N_BREEDS]

(CO)VARIANCES
0.0136 -0.0101
-0.0101 0.0878
RANDOM_GROUP
7
RANDOM_TYPE PRIOR_DEGREES_OF_BELIEF
diagonal 1000000000
PEDIGREEFILE: NAME N_ANIMAL N_GENETIC_GROUPS [N_BREEDS]

(CO)VARIANCES
0.00000001
```

Chapter 3

POST-GENOMIC REACTION NORMS ANALYSES REVEALED NEW CHROMOSOME REGIONS AFFECTING TICK RESISTANCE IN BEEF CATTLE

Abstract: Traditionally, markers effects are estimated for each population in a specific environment. However, it would not be necessarily the same across environments. Under a genome wide selection (GWS) approach, the genotype by environment interactions (G*E) can be regarded based on marker information, thus enabling to study the marker by environment interactions. Although genetic variation in response to environmental changes has been reported for economically important traits in beef cattle, for tick resistance these studies are still scarce. Therefore, this study was aimed to perform genome-enabled predictions for tick resistance in Hereford and Braford cattle by using single step genomic BLUP methodology (ssGBLUP), to estimate marker effects from reaction norms associated with tick resistance as well as to identify candidate genes derived from the most relevant SNP markers. Phenotypic data included 10,673 tick counts on 4,363 Hereford and Braford cattle, being the pedigree composed by 12,754 animals from which 3,591 were genotyped. A one-step reaction norm model was fitted to estimate the (co)variance components and genetic parameters. To study SNP effects across different tick infestation (TI) levels, the top 1% of SNPs were identified in each TI and pointed out to the similarity between these markers across the levels. The additive genetic and permanent environment effects showed significant slope confirming the presence of G*E. Correlations between intercept and slope were positive with high (0.52±0.18) and moderate (0.26±0.15) magnitudes for genetic and permanent environment effects, respectively. From the top 1% SNPs (410), 75 were consistently relevant across TI and indicated SNP by environment interaction. The most relevant SNPs were located on chromosomes 1, 2, 6, 7, 9, 11, 14, 21 and 23 and the annotated genes closest these markers showed functions related to energy metabolism, pigment epithelium, maintenance and integrity of the photoreceptor cells, and cell differentiation.

Key-words: correlations, environmental gradient, gene function, single-step.

INTRODUCTION

The knowledge of genotype by environment interaction (G*E) can be relevant for breeders and failing to consider G*E in genetic evaluations can affect genetic gains since breedstock comparisons will be environment-dependent (Cardoso and Tempelman, 2012; Silva et al, 2014). That is, animals identified as top breeders in a certain environment may not be best for other environmental conditions; this issue is further exacerbated if progeny are managed in an environment different from their selected parents (Cardoso and Tempelman, 2012). In addition, although genetic variation in response to environmental changes has been reported for economically important traits in beef cattle populations (Corrêa et al., 2010; Cardoso et al., 2011; Cardoso and Tempelman, 2012), for tick resistance these studies are still scarce.

Genotype by environment interactions can be described by different methodologies such as multi-trait models and reaction norm models (RNM). RNM approach is based on the use of covariance functions (Kirkpatrick et al., 1990), allowing the prediction of reaction norms of animals in gradual linear changes of the environment. Fitting models such as RNM allow understanding how genotypes are influenced by changes in the environmental gradient and could also improve the accuracy of the prediction of breeding values and selection applications. Several studies have been used reaction norm models to describe G*E in beef cattle populations (Corrêa et al., 2010; Ambrosini et al., 2012; Cardoso et al., 2011; Cardoso and Tempelman, 2012).

Traditionally, markers effects are estimated for each population in a specific environment. However, marker effects would not necessarily be the same in different environments. In genome wide selection (GWS; Meuwissen et al., 2001), G*E interactions can be verified based on marker information, to identify and make use of marker by environment interactions (Silva et al., 2014). In beef cattle, the main interest of genomic-enabled predictions is using markers to improve the accuracy of GEBVs for economically important traits (Pryce et al., 2010). Therefore, we expect that models including G*E would increase the accuracy prediction on GEBV over different environments. Silva et al. (2014) working with total number born in pigs

concluded that genomic reaction norm models outperformed standard genomic models in terms of GEBV accuracy over the environmental levels. However, these authors have implemented a two-step reaction norm model, which is the simplest approach of RNM. According to Su et al. (2006) and Cardoso and Tempelman (2012), the hierarchical Bayesian approach, which treats the covariate associated with the reaction norm as unknown, allows inferring for all unknowns parameters simultaneously, and therefore, seems to be more appropriate than standard two-step model.

Finally, another point that deserves to be highlighted is the genetic dissection of complex phenotypes such as tick counts through the identification of candidate genes derived from the most relevant markers. Relevant studies involving this method were applied for other complex traits such as reproduction traits in beef cattle (Fortes et al., 2010; Hawken et al., 2012), but not for tick resistance. Therefore, this study was aimed to perform genome-enabled predictions for tick resistance in Hereford and Braford cattle by using single step genomic BLUP methodology (ssGBLUP), to estimate marker effects from reaction norms associated with tick resistance as well as to identify candidate genes derived from the most relevant SNP markers.

MATERIAL AND METHODS

Phenotypic and genotypic data

Phenotypic data used in this current study included records of tick counts (TC) on Hereford and Braford beef cattle from herds raised in the Rio Grande do Sul state, Brazil. TC were obtained on each animal from 326 to 729 days of age using the method described by Wharton and Utech (1970), whereby all female ticks ≥ 4.5 mm were counted on the entire left side of the animal. Up to 3 such counts were acquired from each animal with each such count separated by a minimum of 30 days similar as in other studies (Frisch and O'Neill, 1998; Budeli et al., 2009; Biegelmeyer, 2012). The distribution of the number of measurements taken per animal was 241, 1,934 and 2,188 animals having one, two and three TC measurements, respectively, for a total of 10,673 records. The average age during the evaluation period was 524 ± 65 days and the overall mean TC was 34.99 with a standard deviation of 42.15 (range 0- 532).

Because TC were not typically normally distributed, the log transformation was used such that LTTC = $\log_{10}$ (TC+1.001) was the response variable. The constant 1.001 was included in this transformation as some of the TC were equal to zero (Ayres et al., 2013; Biegelmeyer, 2012). Note that data transformation does not guarantee that the log transformed data are normal, but might improve properties of estimates, predictions and inferences (Budeli et al., 2009).

A CG was defined as a group of animals being within the same herd, year of birth, season of birth (April to July; August to November and December to March), of the same sex and from the same management group. Each CG was required to have at least five animals and with each LTTC record being within 3.5 Standard Deviation (SD) from their respective CG means. Moreover, connectedness among CG was assessed by the AMC software (Roso and Schenkel, 2006) such that CG with less than 10 genetic links, were removed. Finally, CG effects/means for LTTC were assumed to define the environmental covariates for a linear reaction norm model because they seemed to be the most appropriate entities to describe environmental conditions most important for beef cattle production (Cardoso et al., 2011; Mattar et al., 2011; Cardoso and Tempelman, 2012).

The DNA for each animal was obtained from a blood sample at the time of first tick count. In addition, semen samples from bulls used in artificial insemination (AI) were also obtained. Genotypes based on the BovineSNP50 Illumina v2 BeadChip were acquired on 3,591 of these Hereford and Braford beef cattle. It was provided a total of 54,609 highly informative SNPs uniformly distributed across the entire genome of major cattle breed types.

Genotype quality control was implemented using R/SNPStats package (Clayton, 2012) to remove samples with call rates <0.90, heterozygosity deviations >3.0, mismatching sex and duplicated records. Only SNPs mapped to autosomes, with call rates greater than 0.98, minor allele frequencies (MAF) >0.03 or not highly significant deviations from Hardy-Weinberg equilibrium ($P>10^{-7}$) were used for further analyses. We considered the highest MAF for SNPs in the same position or highly correlated (r>0.98). Finally, the HD panel was filtered to select only SNPs that were also present in the 50K panel as described by Sollero et al. (2014). Missing genotypes

were imputed using FImpute (Sargolzaei et al., 2011) and after various quality control edits, 41,045 SNP markers (78%), including 136 sires, 2,803 BO and 652 HH yearling bulls, steers and heifers with tick count records, remained to estimate genomic relationship coefficients between animals.

The 4,363 animals having records were born between 2008 and 2011 and originated from 197 sires and 3,966 dams with 10 generations of pedigree depth. Pedigree information recovered from historical breeding records comprised 11,967 animals and was highly incomplete due to use of multiple-sire matings. This resulted in 65% of the animals with tick count records having unknown paternity. For genotyped parent-progeny pairs and trios (sire, dam and progeny) pedigree errors were checked by the percentage of Mendelian conflicts (Wiggans et al., 2009) using seekparentf90 software (http://nce.ads.uga.edu), with maximum tolerance of 1% to allow for genotyping errors. If a parent-progeny pair conflict was found or if one or neither parent had been genotyped, the animal's genotypes were compared with those of every other animal to determine if there was a parent-progeny relationship with an animal in the dataset. Unique putative parents of the appropriate sex with less than 1% Mendelian conflicts and suitable birthdates were designated as true parents, when identified.

When a true sire was not identified with the above describe procedure, the simulated annealing algorithm in the MOL_COANC software (Fernandez and Toro, 2006) was used to recreate half-sib families based on observed genomic relationships for genotyped animals born to multiple-sire matings. Pedigrees were reconstructed by creating a 'virtual' ancestor for each identified half-sib family. A total of 12,754 animals remained after pedigree reconstruction and pruning.

Statistical modeling

We fitted a genomic one-step reaction norm model (GLRNM) as following:

$$y_{ijk} = \mathbf{x}'_j \boldsymbol{\beta} + w_i + a_j + b_j w_i + c_j + d_j w_i + e_{ijk}. \quad [1]$$

whose matrix notation is given by:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{P}\mathbf{w} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_b\mathbf{b} + \mathbf{Z}_c\mathbf{c} + \mathbf{Z}_d\mathbf{d} + \mathbf{e}, \quad [2]$$

in which **y** = {y_{ijk}} is the *nx1* vector of observations, **β** is the fixed effects vector of order *p*, $\mathbf{w} = \{w_i\}_{i=1}^{n_w}$ is the vector of environmental effects, $\mathbf{a} = \{a_j\}_{j=1}^{q}$ is the vector of random genetic intercepts, $\mathbf{b} = \{b_j\}_{j=1}^{q}$ is the vector of random genetic slopes, $\mathbf{c} = \{c_j\}_{j=1}^{q}$ is the vector of random permanent environment intercepts, $\mathbf{d} = \{d_j\}_{j=1}^{q}$ is the vector of random permanent environment slopes and **e** is the *nx1* vector of residuals. Furthermore, **X**, **P**, $\mathbf{Z_a}$ and $\mathbf{Z_c}$ are known incidence matrices, whereas the row address of matrices $\mathbf{Z_b}$ and $\mathbf{Z_d}$ has exactly one element equal to the environmental covariate (w_i or an estimate of w_i) for that CG in the row address of the observation, will all other elements in that row equal to zero, and **e** is the vector of random residual.

Prior specifications

To infer environmental sensitivities using a hierarchical Bayesian model, three stages are required: the first stage defines the distribution of the observed data conditional on all other parameters (Su et al., 2006):

$$\mathbf{y}|\boldsymbol{\beta},\mathbf{w},\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d},\mathbf{R} \sim N(\mathbf{X\beta} + \mathbf{Pw} + \mathbf{Z_a a} + \mathbf{Z_b b} + \mathbf{Z_c c} + \mathbf{Z_d d}, \mathbf{R}). \quad [3]$$

A heteroskedastic residual alternative based on residual variances subclasses determined by a decile-based classification was implemented. That is, CG were ordered into one of 10 categories based on decile delimiters of contemporary group estimates ($\hat{w}_i$) from a simple linear animal model, such that $\mathbf{R} = diag\left(\mathbf{I}_{n_k}\sigma_e^2\gamma_k\right)$ where order n_k denotes the number of records delimited by deciles *k*-1 and *k*, being respectively 1,157, 1,174, 1,047, 765, 1,188, 1,192, 1,208, 918, 1,150 and 874, respectively, for *k* = 1,2,…10. This strategy was also considered by Cardoso and Tempelman (2012).

The second stage is represented by the prior distributions of the location parameters, as follows:

$$\beta \sim p(\boldsymbol{\beta}) \quad [4]$$

$$w|\sigma_w^2 \sim N(0, \mathbf{I}\sigma_w^2) \quad [5]$$

$$\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} \sim N\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\sigma_a^2 & \sigma_{ab}\\ \sigma_{ab} & \sigma_b^2\end{bmatrix} \otimes \mathbf{H}\right) \quad [6]$$

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_c^2 & \sigma_{cd} \\ \sigma_{cd} & \sigma_d^2 \end{bmatrix} \otimes \mathbf{I} \right), \quad [7]$$

where $\mathrm{p}(\boldsymbol{\beta}) \propto 1$, σ_w^2 is the environmental effect variance; σ_a^2 and σ_b^2 are, respectively, the additive genetic variances due to the reaction norm intercept and slope; σ_c^2 and σ_d^2 are, respectively, permanent environment variances due to reaction norm intercept and slope, σ_{ab} is the genetic covariance between reaction norm, intercept and slope, σ_{cd} is the permanent environment covariance between reaction norm, intercept and slope.

Then

$\mathbf{r_{ab}} = \sigma_{ab} / \sqrt{\sigma_a^2 * \sigma_b^2}$ and $\mathbf{r_{cd}} = \sigma_{cd} / \sqrt{\sigma_c^2 * \sigma_d^2}$ are the corresponding genetic and permanent environment correlations, respectively, as well as **I** is an identity matrix and **H** represents the genetic relationships based on pedigree adjusted for deviations due to genomic information as Aguilar et al. (2010) as follow:

$$\mathrm{H} = \begin{bmatrix} \mathbf{A_{11}} & \mathbf{A_{12}} \\ \mathbf{A_{21}} & \mathbf{G - A_{22}} \end{bmatrix} = \mathbf{A} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G - A_{22}} \end{bmatrix}, \quad [8]$$

where subscripts 1 and 2 represent ungenotyped and genotyped animals, respectively **A** is a numerator (traditional pedigree) relationship matrix for all animals, **G** is a genomic relationship matrix constructed based on VanRaden (2008) which use the allele frequencies estimated from the genotyped individuals (see Aguilar et al., 2010).

Finally, the third stage was based on specifying a scaled inverse (IG) distribution for the variance of the contemporary group effects; i.e. $\sigma_w^2 | \alpha_w, \beta_w \sim \mathrm{IG}(\alpha_w = 1, \beta_w = 2800)$ where the mean of this distribution is:

$$\mathrm{E}\left(\mathrm{w} | \alpha_w, \beta_w\right) = \frac{\Gamma(\alpha_w + 1)}{\beta_w \Gamma(\alpha_w)} = \frac{\alpha_w}{\beta_w}. \quad [9]$$

Similarly, we specified $\sigma_e^2 | \alpha_e, \beta_e \sim \mathrm{IG}(\alpha_e = 1, \beta_e = 600)$.

Likewise, an inverted Wishart distribution (IW) prior distribution was presumed, respectively, for the permanent environment and additive genetic covariance matrices:

$$\mathbf{U}_0 = \begin{bmatrix} \sigma_c^2 & \sigma_{cd} \\ \sigma_{cd} & \sigma_d^2 \end{bmatrix} \sim \mathrm{IW}(\mathbf{T}_0, \mathrm{v}) \quad [10]$$

$$\mathbf{G}_0 = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix} \sim \mathrm{IW}(\mathbf{T}_0, \mathrm{v}). \quad [11]$$

Following Cardoso and Tempelman (2012), *v*=4 represents a presumed known degrees of freedom (DF), $\mathbf{T}_0 = \begin{bmatrix} 200 & 4 \\ 4 & 0.1 \end{bmatrix}^{-1}$ a presumed scale matrix, and $\mathrm{E}(\mathbf{G}_0) = \frac{\mathbf{T}_0^{-1}}{\mathrm{v} - \mathrm{p} - 1}$ is the prior mean for *v>p+1*, where ***p*** is the number of parameters.

Furthermore, additional hierarchical specifications were required, i.e., $\gamma_k | \alpha_\gamma \sim \mathrm{p}(\gamma_k | \alpha_\gamma) = \mathrm{IG}(\alpha_\gamma, \alpha_\gamma - 1)$, *k*=1,2,…,10 (Cardoso et al., 2005; Kizilkaya and Tempelman, 2005; Cardoso and Tempelman, 2012). In addition, the prior p(α_γ) on α_γ was a gamma with shape and scale hyperparameters values of 0.03 and 0.01, respectively (Cardoso and Tempelman, 2012). This assumption leads to a prior mean of α_γ equal to 3 ($\mathrm{E}(\alpha_\gamma) = 3$) and a large prior variance ($\mathrm{var}(\alpha_\gamma) = 300$) (Cardoso et al., 2005).

Bayesian inference

Analyses were conducted using Bayesian analyses, specifically Monte Carlo Markov Chain (MCMC), to sample all parameters from their fully conditional posterior distributions. Gibbs sampling was generally used except for the w_i's and α_γ. MCMC sampling of these parameters required a random walk Metropolis-Hastings steps due to all of their fully conditional posterior distributions being unrecognizable (see Cardoso and Tempelman, 2012 for details).

MCMC was implemented by INTERGEN software (Cardoso, 2013) by saving every 10 cycles from a total of 1,000,000 cycles, after 100,000 cycles of burn-in. Global convergence was checked using the Geweke's Z criterion (Geweke, 1991) applied to the conditional distribution of the data, as proposed by Brooks and Roberts (1998). In addition, visual inspection of

trace plots was conducted and a minimum effective sample size of 100 for all unknown parameters was always obtained as well.

Derivation of SNP effects from predicted breeding values

According to Wang et al. (2012), for a simple GBLUP model, if we decompose the animal effects into genotyped (a_g) and ungenotyped (a_u) animals, the effect of animals *j* (a_j) are a function of SNP effects as follow:

$$a_j = \mathbf{Z}u, \quad [12]$$

where **Z** is a matrix relating genotypes of each locus and **u** is a vector of SNP effects:

Therefore, the equation for predicting SNP effects is:

$$\hat{u} = \mathbf{Z}'[\mathbf{Z}\mathbf{Z}']^{-1}\hat{a}_g. \quad [13]$$

This is the best predictor of SNP effects given animal effects. On the other hand, for a reaction norm model, it was possible to determine a vector of SNP effects ($\hat{u}_i$) for each environment level **i** (**i**=CG levels= 146 levels). According to Van Randen (2008), the sum over all marker loci is assumed to equal the vector of breeding values. Therefore, we used the normal equation to calculate the SNP effects using rrBlupM6 function from rrBlupMethod6 package (Schulz-Streeck et al., 2012) as given bellow:

$$\begin{aligned}
\hat{u}_i &= (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\hat{u}_i) \\
\hat{u}_i &= (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'(\hat{a}+\hat{b}\hat{w}_i)) \\
\hat{u}_i &= (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\hat{a}+\mathbf{Z}'\hat{b}\hat{w}_i) \\
\hat{u}_i &= (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\hat{a})+(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\hat{b})\hat{w}_i
\end{aligned} \quad [14]$$

Thus, a general linear prediction equation can be obtained:

$$\hat{u}_i = \hat{a}_{SNP} + \hat{b}_{SNP}\hat{w}_i, \quad [15]$$

which allows the estimation of a vector of marker effects for each contemporary group level of interest (Silva et al., 2014). In addition, following Silva et al. (2014), to verify if SNPs have the same effect across CG levels, we identified the top 1% of SNPs (410) that presented the greatest absolute values in each environmental level, and it was used as a simple statistic to reflect the similarity of SNP associations across environments.

Gene function annotation

To provide information regarding gene identity and function mapped from SNP markers, a list of genes closest at maximum distance of 2.5 Mb between SNPs from the most relevant SNPs were extracted in low (LTI), medium (MTI) and high (HTI) tick infestation environments, represented by 10^{th}, 50^{th} and 90^{th} environmental gradient percentiles, respectively (Verardo et al., 2014).

Putative genes identified for Hereford and Braford breeds were established by BLAST Homology search knowing human gene transcripts were downloaded from the NCBI RefSeqsequence database (http://www.ncbi.nlm.nih.gov). The biological function of these genes and possible relation to tick resistance were investigated. Thus, it was possible to identify the biological mechanisms and functions involving the identified genes as well as highlight those associated to different tick infestation levels in Hereford and Braford beef cattle.

RESULTS AND DISCUSSION

Variance components and correlations

The main results of the variance component estimation are presented in Table 1. The additive genetic and permanent environment effects showed significant slope confirming the presence of G*E and the importance to consider the environment permanent effect. The residual class variances ranged between 0.02 and 0.13 and showed no clear pattern (Table 1). Cardoso and Tempelman (2012) working with G*E models in post-weaning gains in Angus Cattle also observed that the residual variance did not increase categorically over the environmental gradient.

The correlations between intercept and slope were positive with high (0.52±0.18) and moderate (0.26±0.15) magnitude for additive genetic and permanent environment effects, respectively. It would indicate that animals with larger genetic values tend to show better greater performance in more favorable environments. Cardoso et al. (2011) working with post weaning gain in Hereford beef cattle also observed a high and positive correlation between genetic intercept and slope and also reported that animals of higher

average breeding value were the ones which presented the best response to environmental improvement.

Table 1 Posterior means (standard deviations) for the one-step reaction norm model parameters

Variance components[1]			
σ^2_a	0.016 (0.002)	σ^2_{e1}	0.129 (0.008)
σ^2_b	0.030 (0.011)	σ^2_{e2}	0.057 (0.005)
σ_{ab}	0.011 (0.004)	σ^2_{e3}	0.065 (0.004)
σ^2_c	0.012 (0.002)	σ^2_{e4}	0.049 (0.003)
σ^2_d	0.091 (0.018)	σ^2_{e5}	0.102 (0.005)
σ_{cd}	0.009 (0.005)	σ^2_{e6}	0.022 (0.002)
σ^2_w	0.093 (0.011)	σ^2_{e7}	0.052 (0.003)
r_{ab}	0.518 (0.184)	σ^2_{e8}	0.062 (0.004)
r_{cd}	0.265 (0.153)	σ^2_{e9}	0.056 (0.003)
σ^2_e	0.073 (0.015)	σ^2_{e10}	0.056 (0.004)

[1]σ^2_a = reaction norm intercept genetic variance; σ^2_b = reaction norm slope genetic variance; σ_{ab} = genetic covariance between intercept and slope; σ^2_c = reaction norm intercept permanent environment variance; σ^2_d = reaction norm slope permanent environment variance; σ_{cd} = permanent environment covariance between intercept and slope; σ^2_w = environmental variance; σ^2_e = residual term; σ^2_{e1} = residual class 1; σ^2_{e2} = residual class 2; σ^2_{e3} = residual class 3; σ^2_{e4} = residual class 4; σ^2_{e5} = residual class 5; σ^2_{e6} = residual class 6; σ^2_{e7} = residual class 7; σ^2_{e8} = residual class 8; σ^2_{e9} = residual class 9; σ^2_{e10} = residual class 10; r_{ab} genetic correlation between intercept and slope; r_{cd} permanent environment correlation between intercept and slope.

SNP by environment interaction

The number of top 1% of SNPs that presented the greatest absolute values in each environmental level was 410, but only 75 were consistently present across all tick infestation (TI) levels. The SNP effects, in absolute values, estimated for these 75 consistently markers were plotted along the environmental gradient (Figure 1a), which indicated SNP by environment interaction due to remarkable changes in the magnitude of the SNP effects across TI levels, especially in low tick infestation environments. It can be clearly observed in Figure 1b where was plotted the 10 most relevant SNP markers along the environmental gradient.

In addition, we also observed that most of these 75 markers presented an increasing effect with increasing TI levels, i.e., none of the SNP showed smaller effect at higher TI levels. Figure 1 also demonstrates that differences in SNP effects estimates decrease with a low TI level, i.e., the variation of the SNP effects was greater at high TI levels. It could further indicate the difficulty to identify superior breedstock in better (low TI) environments.

Chromosome regions affecting tick resistance

The Manhattan plots show the top 0.01% (4; dark gray line) and top 0.02% (8; light gray line) for three different environment levels associated with tick infestation across environments (Figure 2). From the top 0.02% (8) most relevant SNPs for the three different environments of tick infestation levels (low, medium and high), six were detected in both, and seven between low and medium and also between medium and high TI level. The most relevant SNPs were detected on chromosomes 1, 2, 6, 7, 9, 11, 14, 21 and 23 across all environments tested (Table 2). In addition, studies have already reported QTLs in chromosomes 2, 11, 23 (Machado et al., 2010), 7 (Gasparin et al., 2007) and 14 (Gasparin et al., 2007; Porto Neto et al., 2012) for tick resistance as detected in this study.

However, it was also detected relevant SNP markers in chromosomes (1, 6, 9 and 21) that have not been reported in literature. In addition, relevant SNPs number and position, for the three different environment levels associated with TI across environments, LTI, MTI and HTI, are presented in Table 2. The most relevant SNP in all TI levels was detected on BTA9. The fifth highest value SNP marker in all TI levels was located on BTA21. Furthermore, a SNP marker on BTA1 was also verified for LTI, MTI and HTI. At last, a SNP marker affecting LTI and MTI was also reported on BTA6.

QTLs in BTA1, BTA6, BTA9 and BTA21 have already been reported in other beef cattle studies. Mcclure et al. (2010) performed a quantitative trait locus (QTL) scan for 14 economically important traits including calving ease (direct and maternal), height (yearling and mature), marbling score, rib eye area and weaning weight in two commercial Angus populations using 390 microsatellites, 11 single nucleotide polymorphisms (SNPs) and one duplication loci in two different populations of Angus cattle (Hu et al., 2013). In addition, studies have also been reported in dairy cattle; on BTA6 for several milk traits (Kucerova et al., 2007; Khatib et al., 2007), on BTA9 for milk fat yield (Wiener et al., 2000) and clinical mastitis (Lund et al., 2008) and BTA21 for abomasums displacement (Momke et al., 2008), gastrointestinal nematode burden (Coppieters et al., 2009) and stillbirth (Seidenspinner et al., 2009).

The polymorphism with higher effect located on BTA9 is close to *MYO6* and *IMPG1* genes. Majewski et al. (2010) indicated that *MYO6* is present in secretory adrenal medulla cells and serves as a motor engaged in chromaffin granule transport within the cytoplasm. Those authors also opened the discussion on its possible involvement of *MYO6* in nucleo-cytoplasmic trafficking and have suggested further investigation. On the other hand, the *IMPG1* gene (interphotoreceptor matrix proteoglycan 1), marked by the same SNP as for *MYO6*, is a unique extracellular matrix occupying the space between the photoreceptors and the retinal pigment epithelium. This gene is related in maintenance and integrity of photoreceptor cells and also retinal adhesion (Felbor et al., 1998; Kothapalli et al., 2007). In a tick resistance point of view, this gene could be important due to different epithelium pigmentation is a factor causing differing resistance to ticks in beef cattle, for example, animals with dark skin tend to be less susceptible to tick infestation than animals with light skin.

The peroxisome proliferator-activated receptor gamma coactivator 1 alpha protein, encoded by *PPARGC1A* gene, was related to a SNP marker located on BTA6 and is a metabolic switch, which transcriptionally activates a complex pathway of mitochondrial biogenesis and energy (lipid and glucose) metabolism. Despite *PPARGC1A* was postulated as the most plausible gene underlying a QTL for fat yield (Streit et al., 2013), Komisarek et al. (2012) reported that *PPARGC1A* had a significant effect on lengths of calving interval and calving to conception interval. Although this gene is not part of the immune system, it would be important to further investigate if energy metabolism is related to tick resistance, i.e., for example, if animals presenting different fat deposition would be more or less susceptible to ticks or none. Breed differences and associated effects of maturity or growth potential on the subcutaneous or intramuscular fatty acid composition of beef have been reported in literature (Mills et al., 1992; Siebert et al., 1996; Pitchford et al., 2002).

The *SETD3* gene (SET domain containing 3) is a protein-coding gene, expressed abundantly in muscular tissues, that was marked as near from a significant SNP marker on BTA21 for tick resistance. This gene has been reported to be related to histone methyltransferase and transcription

coactivator activity, i.e., *SETD3*, when over expressed, activates transcipiton of muscle-related genes, myogenin, muscle creatine kinase (MCK), and myogenic factor 6 (Myf6), thereby inducing muscle cell differentiation. Otherwise, knockdown of SETD3 by short hairpin RNA (shRNA) significantly retards muscle cell differentiation (Eom et al., 2011). Further investigation on this gene would be important because it may be related to immune system and consequently tick resistance.

The results from this study confirmed the presence of SNP by environment interaction and provided functional information on the most relevant identified loci (individual significant SNP markers) that affect different levels of tick infestation. Therefore, a complementary study associated with the validation of GWAS candidate genes such as reported by Streit et al. (2013) and/or expression analyses may be considered in the future. In addition, SNP-derived gene function annotation additional analysis for all significant SNPs for both terms (intercept and slope) is needed to identify the candidate genes and their role in the tick resistance of Hereford and Braford beef cattle.

Table 2 SNP number (SNP), Chromosome (Chr.) and Position from the highest to lowest SNP effect value in three different tick infestation environments

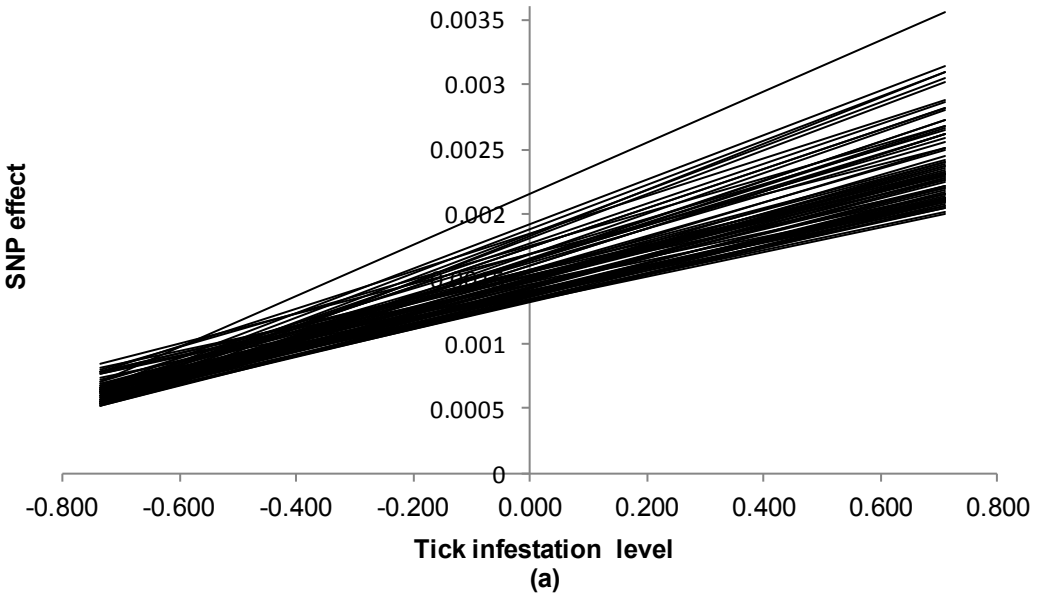
LTI			MTI			HTI		
SNP	**Chr.**	**Position**	**SNP**	**Chr.**	**Position**	**SNP**	**Chr.**	**Position**
16590	9	15767136	16590	9	15767136	16590	9	15767136
25340	14	51396430	25340	14	51396430	35442	23	12997137
2081	1	125283942	35442	23	12997137	25340	14	51396430
14167	7	96192503	14167	7	96192503	14167	7	96192503
34159	21	66064177	34159	21	66064177	34159	21	66064177
11321	6	45216251	21270	11	86751976	21270	11	86751976
21270	11	86751976	2081	1	125283942	274	1	18508517
4300	2	104905603	11321	6	45216251	2081	1	125283942

LTI= low tick infestation; MTI= medium tick infestation; HTI= high tick infestation.

CONCLUSIONS

The presence of SNP by environment interactions was confirmed for tick resistance via one-step reaction norm models and SNP markers that affect intercept and slope and consequently different tick infestation environments could be identified.

We reported the most relevant SNP markers associated with tick infestation across environments closest to genes with different direct functions such as energy metabolism, retinal pigment epithelium, maintenance and integrity of the photoreceptor cells as well as cell differentiation which might be related to tick resistance.



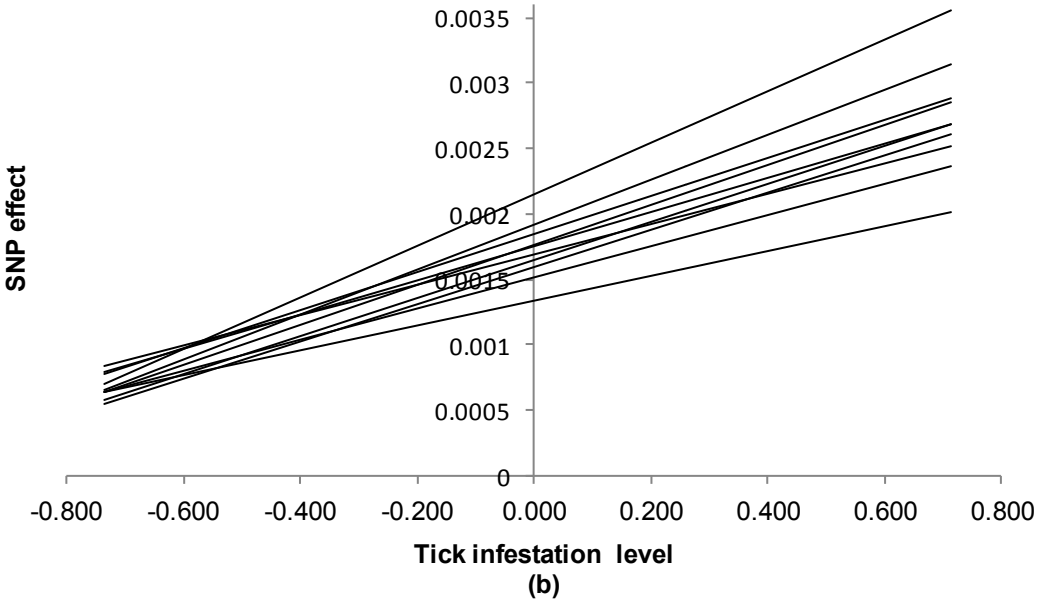


Figure 1 SNP effects estimates (absolute values) for the 75 (a) and 10 (b) most relevant SNPs at all tick infestation levels.

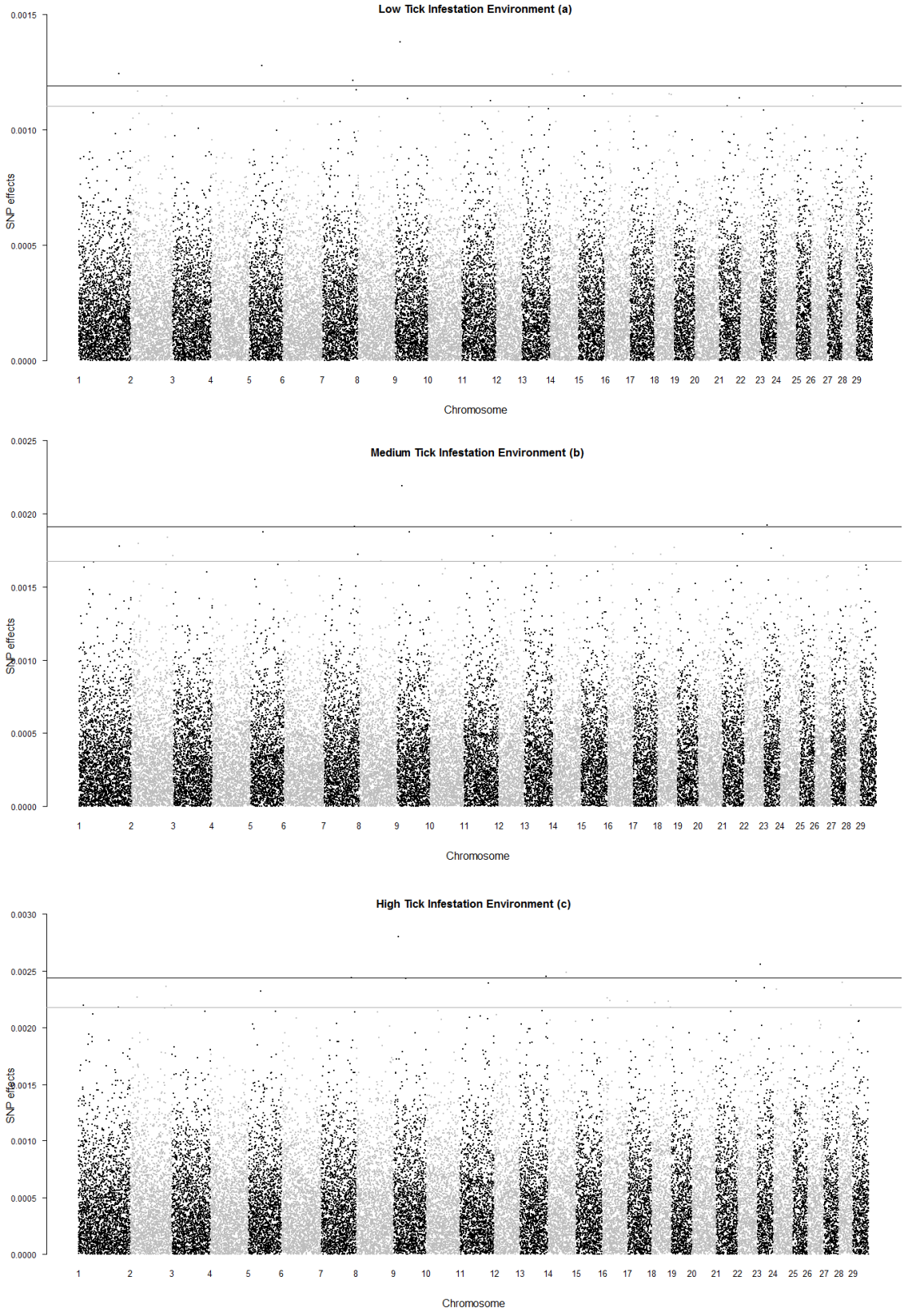


Figure 2 Manhattan plots in low (a), medium (b) and high (c) tick infestation levels. Chromosomes 1 to 29 are shown separated by alternating colors. The corresponding horizontal line indicates the four (dark gray) and eight (light gray) most relevant SNP markers.

ACKNOWLEDGMENTS

The authors thank Delta G Connection by providing the data used in this research; Embrapa South Livestock, Instituto Nacional de Investigacion Agropecuaria (INIA – Uruguay) and Michigan State University for theoretical and technical support; CAPES and CNPq by granting the scholarship. This project was supported by Agriculture and Food Research Initiative Competitive Grant no. 2011-67015-30338 from the USDA National Institute of Food and Agriculture.

REFERENCES

AGUILAR, I.; MISZTAL, I.; JOHNSON, D.L. et al. 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.

AMBROSINI, D.P.; CARNEIRO, P.L.S.; BRACCINI NETO, J. et al. Genotype × environment interaction for yearling weight in Polled Nellore cattle in Northeast Brazil. **Pesquisa Agropecuária Brasileira**, v.47, p.1489-1495, 2012.

AYRES, D.R.; PEREIRA, J.R.; BOLIGON, A.A. et al. Linear and Poisson models for genetic evaluation of tick resistance in cross-bred Hereford x Nellore cattle. **Journal of Animal Breeding and Genetics**, v.130, p.417-424, 2013.

BIEGELMEYER, P. **Genetic resistance to *Rhipicephalus* (Boophilus) *microplus* natural and artificial infestation in Hereford and Braford cattle**. 2012, 96p. Dissertation (Master in Science) - Universidade Federal de Pelotas, Pelotas, RS .

BROOKS, S.P.; ROBERTS, G.O. Convergence assessment techniques for Markov Chain Monte Carlo. **Statistics and Computing**, v.8, p.319-335, 1998.

BUDELI, M.A.; NEPHAWE, K.A.; NORRIS, D. et al. Genetic parameter estimates for tick resistance in Bonsmara cattle. **South African Journal of Animal Science**, v.39, p.321-327, 2009.

CARDOSO, F.F.; ROSA, G.J.M.; TEMPELMAN, R.J. Multiple-breed genetic inference using heavy-tailed structural models for heterogeneous residual variances. **Journal of Animal Science**, v.83, p.1766-1779, 2005.

CARDOSO, L.L.; BRACINI NETO, J.; CARDOSO F.F. et al. Hierarchical Bayesian models for genotype x environment interaction estimation via reaction norms applied to post-weaning gain of Hereford cattle. **Revista Brasileira de Zootecnia**, v.40, p.294-300, 2011.

CARDOSO, F.F.; TEMPELMAN, R.J. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. **Journal of Animal Science**, v.90, p.2130-2141, 2012.

CARDOSO, F.F. **Application of Bayesian inference in animal breeding using the Intergen program: Manual of version 1.22**. Documents 112, Embrapa South Livestock Bagé, RS, Brazil, 2013.

Clayton, D. snpStats: SnpMatrix and XSnpMatrix classes and methods. http://www-gene.cimr.cam.ac.uk/clayton. (Accessed 30 October 2012.), 2012.

COPPIETERS, W.; MES, T.H.; DRUET, T. et al. Mapping QTL influencing gastrointestinal nematode burden in Dutch Holstein-Friesian dairy cattle. **BMC Genomics**, v.10, 2009.

CORRÊA, M.B.B.; DIONELLO, N.J.L.; CARDOSO, F.F. Genetic evaluation of Devon cattle using a reaction norms model. **Revista Brasileira de Zootecnia**, v.39, p.128-133, 2010.

EOM, G.H.; KIM, K.; KIM, J.H. et al. Histone Methyltransferase SETD3 Regulates Muscle Differentiation. **The Journal of Biological Chemistry**, v.286, p.34733-34742, 2011.

FELBOR, U.; GEHRIG, A.; SAUER, G. et al. Genomic organization and chromosomal localization of the interphotoreceptor matrix proteoglycan-1 (IMPG1) gene: a candidate for 6q-linked retinopathies. **Cytogenetics and Cell Genetics**, v.81, p.12-17, 1998.

FERNANDEZ, J.; TORO, M.A. A new method to estimate relatedness from molecular markers. **Molecular Ecology**, v.15, p.1657-1667, 2006.

FORTES, M.R.S.; REVERTER A.; ZHANG Y. et al. Association weight matrix for the genetic dissection of puberty in beef cattle. ***Proceedings*** of the National Academic of Sciences USA, v.107, p.13642-13647, 2010.

FRISCH, J.E.; O'NEILL, C.J. Comparative evaluation of beef cattle breeds of African, European and Indian origins. 2. Resistance to cattle ticks and gastrointestinal nematodes. **Animal Science**, v.67, p.39-48, 1998.

GASPARIN, G.; MIYATA, M.; COUTINHO, L.L. et al. Mapping of quantitative trait loci controlling tick [*Riphicephalus* (*Boophilus*) *microplus*] resistance on bovine chromosomes 5, 7 and 14. **Animal Genetics**, v.38, p.453-459, 2007.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: BERNARDO, J.M.; BERGER, J.O.; DAWID, A.P. et al. (Eds.). **Bayesian Statistics 4**. New York: Oxford University, 1991. 526p.

HAWKEN, R.J.; ZHANG, Y.D.; FORTES, M.R.S. et al. Genome - wide association studies of female reproduction in tropically adapted beef cattle. **Journal of Animal Science**, v.90, p.1398-1410, 2012.

HU Z.L.; PARK C.A.; WU X.L. et al. Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. **Nucleic Acids Research**, v.41, p.871–879, 2013.

KHATIB, H.; ZAITOUN, I.; WIEBELHAUS-FINGER, J. et al. The association of bovine PPARGC1A and OPN genes with milk composition in two independent Holstein cattle populations. **Journal of Dairy Science**, v.90, p.2966-2970, 2007.

KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, p.979-993, 1990.

KIZILKAYA, K.; TEMPELMAN, R.J. A general approach to mixed effects modeling of residual variances in generalized linear mixed models. **Genetic Selection Evolution**, v.37, p.31-56, 2005.

KOMISAREK, J.; WALENDOWSKA, A. Analysis of the PPARGC1A gene as a potential marker for productive and reproductive traits in cattle. **Folia Biologica**, v.60, p.171-174, 2012.

KOTHAPALLI, K.S.D.; ANTHONY, J.C.; PAN, B.S et al. Differential Cerebral Cortex Transcriptomes of Baboon Neonates Consuming Moderate and High Docosahexaenoic Acid Formulas. **Plos one**, v.4, 2007.

KUCEROVA, J.; LUND, M.S.; SORENSEN, P. et al. Multitrait quantitative trait Loci mapping for milk production traits in Danish Holstein cattle. **Journal of Dairy Science**, v.89, p. 2245-2256.

LUND, M.S.; GULDBRANDTSEN, A.J.; THOMSEN, B. et al. Detection of Quantitative Trait Loci in Danish Holstein Cattle Affecting Clinical Mastitis, Somatic Cell Score, Udder Conformation Traits, and Assessment of Associated Effects on Milk Yield. **Journal of Dairy Science**, v.91, p.4028-4036, 2008.

MACHADO, M.A.; AZEVEDO, A.L.S.; TEODORO, R.L. et al. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (*Bos taurus* × *Bos indicus*). **BMC Genomics**, v.11, 2010.

MAJEWSKI, L.; SOBCZAK, M.; REDOWICZ, M.J. Myosin VI is associated with secretory granules and is present in the nucleus in adrenal medulla chromaffin cells. **Acta Biochimica Polonica**, v.57, p.109-114, 2010.

MATTAR, M.; SILVA, L.O.C.; ALENCAR, M.M. et al. Genotype by environment interaction for long-yearling weight in Canchim cattle quantified by reaction norm analysis. **Journal of Animal Science**, v.89, p.2349-2355, 2011.

MCCLURE, M.C.; MORSCI, N.S.; SCHNABEL, R.D. et al. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. **Animal Genetics**, v.41, p.597-606, 2010.

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.

MILLS, E.W.; COMERFORD, J.W.; HOLLENDER R. et al. Meat composition and palatability of Holstein and beef steers as influenced by forage type and protein source, **Journal of Animal Science**, v.70, p. 2446–2451, 1992.

MOMKE, S.; SCHOLZ, H.; DOLL, K. et al. Mapping quantitative trait Loci for left-sided displacement of the abomasum in German Holstein dairy cows. **Jounal of Dairy Science**, v.91, p.4383-4392, 2008.

NCBI - Ncbi.nlm.nih.gov. National Center for Biotechnology Information. Avaiable from: <http://www.ncbi.nlm.nih.gov/>. Acessed in: February 3rd, 2015.

PITCHFORD, W.S.; DELAND, M.P.B.; SIEBERT, B.D. et al. Genetic variation in fatness and fatty acid composition of crossbred cattle, **Journal of Animal Science**, v.80, p.2825–2832, 2002.

PORTO NETO, L.R.; JONSSON, N.N.; INGHAM, A. et al. The RIPK2 gene: a positional candidate for tick burden supported by genetic associations in

cattle and immunological response of knockout mouse. **Immunogenetics**, v.64, p.379-388, 2012.

PRYCE, J.E.; BOLORMAA, S.; CHAMBERLAIN, A.J. et al. A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. **Journal of Dairy Science**, v.93, p.3331-3345, 2010.

ROSO, V.M.; SCHENKEL, F.S. AMC – A computer program to assess the degree of connectedness among contemporary groups. ***Proceedings*** of 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil, 2006. (CD-ROM).

SARGOLZAEI, M.; CHESNAIS, J. SCHENKEL, F.; FImpute-An efficient imputation algorithm for dairy cattle populations. **Journal of Animal Science**, v.94 (E-Suppl. 1, p.421(abstr. 333), 2011.

SCHULZ-STREECK, T.; ESTAGHVIROU, B.; TECHNOW, F. rrBlupMethod6. R package version 1.3, 2012.

SIEBERT, B.D.; DELAND, M.P.; PITCHFORD W.S. Breed differences in the fatty acid composition of subcutaneous and intramuscular lipid of early and late maturing, grain-finished cattle. **Australian Journal of Agricultural Research**, v.47, p.943–952, 1996.

SEIDENSPINNER, T.; BENNEWITZ, J.; REINHARDT, F. et al. Need for sharp phenotypes in QTL detection for calving traits in dairy cattle. **Journal of Animal Breeding and Genetics**, v.126, p.455-462, 2009.

SILVA, F.F., MULDER, H.A., KNOL, E.F. et al. Sire evaluation for total number born in pigs using a genomic reaction norms approach. **Journal of Animal Science**, v.92, p.3825-3834, 2014.

SOLLERO, B.P.; GOMES, C.C.G.; ROSO, V.M. et al. Goodness of Fit Comparisons among Five Bayesian Models in Genome-Wide Association of Tick Resistance in Brazilian Hereford and Braford Beef Cattle. ***Proceedings*** of 10th World Congress on Genetics Applied to Livestock Production, Vancouver, Canada, 2014.

STREIT, M.; WELLMANN, R.; REINHARDT, F. et al. Using genome-wide association analysis to characterize environmental sensitivity of milk traits in dairy cattle. **Genetics**, v.3, p.1085-1093, 2013.

SU, G.; MADSEN, P.; LUND, M.S. et al. Bayesian analysis of the linear reaction norm model with unknown covariates. **Journal of Animal Science**, v.84, p.1651-1657, 2006.

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

VERARDO, L.L; SILVA, F.F.; VARONA, L. et al. Bayesian GWAS and network analysis revealed new candidate genes for number of teats in pigs. **Journal of Applied Genetics**, v.56, p.123-132, 2014.

WANG, H.; MISZTAL, I.; AGUILAR, I. et al. Genome-wide association mapping including phenotypes from relatives without genotypes. **Genetics Research**, v.94, p.73-83, 2012.

WHARTON, R.H.; UTECH, K.B.W. The relation between engorgement and dropping of Boophilus microplus (Canestrini) (Ixodidae) to the assessment to tick numbers on cattle. **Journal Australian Entomologycal Society**, v.9, p.171-182, 1970.

WIENER, P.; MACLEAN, I.; WILLIAMS, J.L. et al. Testing for the presence of previously identified QTL for milk production traits in new populations. **Animal Genetics**, v.31, p.385-395, 2000.

WIGGANS, G.R.; SONSTEGARD, T.S.; VANRADEN, P.M. et al. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. **Journal of Dairy Science**, v.92, p.3431-3436, 2009.