

# Genome-enabled prediction for tick resistance in Hereford and Braford beef cattle via reaction norm models<sup>1</sup>

R. R. Mota,<sup>\*2</sup> P. S. Lopes,<sup>\*</sup> R. J. Tempelman,<sup>†</sup>  
F. F. Silva,<sup>\*</sup> I. Aguilar,<sup>‡</sup> C. C. G. Gomes,<sup>§</sup> and F. F. Cardoso,<sup>§#</sup>

<sup>\*</sup>Animal Science Department, Federal University of Viçosa, 36570-900 Viçosa, Brazil; <sup>†</sup>Animal Science Department, Michigan State University, East Lansing 48823; <sup>‡</sup>National Agricultural Research Institute, 90200 Las Brujas, Uruguay; <sup>§</sup>Embrapa South Livestock, Bagé, Brazil; and <sup>#</sup>Animal Science Department, Federal University of Pelotas, 96010-900 Pelotas, Brazil

**ABSTRACT:** Very few studies have been conducted to infer genotype × environment interaction (G×E) based in genomic prediction models using SNP markers. Therefore, our main objective was to compare a conventional genomic-based single-step model (HBLUP) with its reaction norm model extension (genomic 1-step linear reaction norm model [HLRNM]) to provide EBV for tick resistance as well as to compare predictive performance of these models with counterpart models that ignore SNP marker information, that is, a linear animal model (ABLUP) and its reaction norm extension (1-step linear reaction norm model [ALRNM]). Phenotypes included 10,673 tick counts on 4,363 Hereford and Braford animals, of which 3,591 were genotyped. Using the deviance information criterion for model choice, ABLUP and HBLUP seemed to be poorer fitting in comparison with their respective genomic model extensions. The HLRNM estimated lower average and reaction norm genetic variability compared with the ALRNM, whereas ABLUP and HBLUP seemed to be poorer

fitting in comparison with their respective genomic reaction norm model extensions. Heritability and repeatability estimates varied along the environmental gradient (EG) and the genetic correlations were remarkably low between high and low EG, indicating the presence of G×E for tick resistance in these populations. Based on 5-fold *K*-means partitioning, mean cross-validation estimates with their respective SE of predictive accuracy were 0.66 (SE 0.02), 0.67 (SE 0.02), 0.67 (SE 0.02), and 0.66 (SE 0.02) for ABLUP, HBLUP, HLRNM, and ALRNM, respectively. For 5-fold random partitioning, HLRNM (0.71 ± 0.01) was statistically different from ABLUP (0.67 ± 0.01). However, no statistical significance was reported when considering HBLUP (0.70 ± 0.01) and ALRNM (0.70 ± 0.01). Our results suggest that SNP marker information does not lead to higher prediction accuracies in reaction norm models. Furthermore, these accuracies decreased as the tick infestation level increased and as the relationship between animals in training and validation data sets decreased.

**Key words:** accuracy, cross-validation, genetic correlation, heritability

© 2016 American Society of Animal Science. All rights reserved.

J. Anim. Sci. 2016.94:1834–1843  
doi:10.2527/jas2015-0194

## INTRODUCTION

The cattle tick is of substantial concern in tropical areas, because it can greatly diminish animal performance. Furthermore, parasite resistance due to indiscriminate use of treatments with acaricides, risk of chemical residues in milk and beef, and also repeated failures of effective vaccine development has driven researchers to seek for alternative solutions. A potentially viable alternative to overcoming this problem is the selection of genetically superior animals for tick resistance based on genetic evaluation programs.

<sup>1</sup>Research supported by Embrapa – Brazilian Agricultural Research Corporation grants 02.09.07.004 and 01.11.07.002.07 and Agriculture and Food Research Initiative Competitive Grant number 2011-67015-30338 from the USDA National Institute of Food and Agriculture. The authors thank Delta G Connection by providing animals and data used in this research, CAPES – Coordination for the Improvement of Higher Level Personnel, and CNPq – National Council for Scientific and Technological Development by granting the scholarship.

<sup>2</sup>Corresponding author: rodrigo.mota@ufv.br

Received December 11, 2015.

Accepted March 10, 2016.

Studies have demonstrated that tick resistance may be heritable with heritability estimates ranging from 0.05 to 0.42 (Rechav et al., 1990; Burrow, 2001; Budeli et al., 2009; Oliveira et al., 2012; Ayres et al., 2013).

Reaction norms refer to phenotypic profiles as influenced by genotypes and by variation across environments. The use of linear reaction norm models (LRNM) specifically models genetic merit as a linear function of an environmental gradient (EG; Falconer and Mackay, 1996). Typically, these EG are based on the mean performance of the response variable for the various environments and are typically based on contemporary group (CG) assignments (Cardoso and Tempelman, 2012). Several studies have reported the importance of genotype  $\times$  environment interaction (G $\times$ E) in different traits in beef cattle that potentially translate into genetic rerankings of animals across different environments (Pégolo et al., 2009; Corrêa et al., 2010; Ambrosini et al., 2012). Mota et al. (2016) has suggested that G $\times$ E exists for tick resistance and may be captured using LRNM.

The use of genomic-wide selection (Meuwissen et al., 2001) is widely believed to enhance the accuracy of EBV such that it is conceivable to further infer G $\times$ E based on SNP marker information. Silva et al. (2014), studying total number born in pigs as their response, used LRNM in conjunction with SNP marker information, suggesting that as a promising approach to infer G $\times$ E.

However, this Silva et al. (2014) and other LRNM studies are typically conducted using a commonly used 2-step reaction norm approach (Calus et al., 2002; Kolmodin et al., 2002) whereby the EG covariate is first estimated as the CG mean or effect in pedigree-based linear animal model (ABLUP), that is, without a reaction norm specification. These CG estimates are then considered as known BLUP EG covariates in the subsequent LRNM analysis. Su et al. (2006) demonstrated that failing to take in account the uncertainty of these covariates could lead to biased inferences, including incorrect genetic rankings of animals. They subsequently proposed a 1-step Bayesian LRNM approach, which treats the EG covariate as having uncertainty. Mota et al. (2016) reported that models fitted using the 1-step approach demonstrated better fit than the 2-step approach for tick resistance in Hereford and Braford beef cattle populations. Cardoso (2013) developed software (Intergen) to fit 1-step LRNM that further allow for the specification of heterogeneous residual variances and genetic marker information. Cardoso and Tempelman (2012) reported that 1-step LRNM allowing for heterogeneous residual variances across environments were better fitting in comparison with specifications based on either 2-step or homogeneous residual variances across environments for postweaning BW gain in Angus cattle.

The objectives of this study were 1) to compare conventional nongenomic- and genomic-based models with their reaction norm extensions on tick infestation data considering the CG effects as EG and 2) to compare breeding values obtained from using nongenomic and genomic approaches.

## MATERIALS AND METHODS

This work was developed using preexisting data sets. All experimental procedures that involved animals to generate the original data were approved by the Committee for Ethics in Animal Experimentation from the Federal University of Pelotas (Pelotas, RS, Brazil; Process Committee for Ethics in Animal Experimentation number 6849).

### *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 Rio Grande do Sul state, Brazil. Up to 3 TC were obtained on each animal from 326 to 729 d of age, ensuring that a minimum interval of 30 d has elapsed between counts. Tick counts were performed by recording the number of female ticks  $\geq 4.5$  mm of length present on one side of the animal (Wharton and Utech, 1970; Cardoso et al., 2015). The distribution of the number of measurements taken per animal was 241, 1,934, and 2,188 animals having 1, 2, and 3 TC measurements, respectively, for a total of 10,673 records. The average age during the evaluation period was  $524 \pm 65$  d and the overall mean TC was 34.99 with a SD of 42.15 (range 0–532). Because TC were not normally distributed, the log transformation of tick counts (LTTC) 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 0 (Biegelmeier, 2012; Ayres et al., 2013).

The CG were defined as groups of animals being within the same herd, year of birth, and 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 5 animals and with each LTTC record being within 3.5 SD from their respective CG means. Moreover, connectedness among CG was assessed by the AMC software (Roso and Schenkel, 2006), such that CG with fewer than 10 genetic links were removed. Finally, CG effects/means for LTTC were assumed to define the environmental covariates (i.e., EG) for a LRNM, as these effects are typically 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).

Genotypes based on 54,609 SNP markers from the BovineSNP50 Illumina BeadChip Technology (Illumina, San Diego, CA) were acquired on 3,591 of these Hereford and Braford beef cattle. Genotype quality control was implemented using the R/snpStats package (Clayton, 2012) to remove samples with call rates < 0.90, heterozygosity deviations > 3.0, mismatching sex, and duplicated records. Only SNP mapped to autosomes, with call rates greater than 0.98, minor allele frequencies > 0.03, or not highly significant deviations from Hardy–Weinberg equilibrium ( $P > 10^{-7}$ ), were used for further analyses. We considered the highest minor allele frequencies for SNP in the same position or highly correlated ( $r > 0.98$ ). 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 Braford; and 652 Hereford yearling bulls, steers, and heifers with TC 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 up to 10 generations of pedigree depth. Pedigree information recovered from historical breeding records comprised 11,967 animals and was highly incomplete because of multiple-sire mating. This resulted in 65% of the animals with TC having unknown paternity. For pairs of genotyped parent–progeny, mismatches on pedigree errors were checked by the percentage of Mendelian conflicts as proposed by Wiggans et al. (2009) by using seekparentf90 software (<http://nce.ads.uga.edu/wiki/doku.php?id=readme.seekparentf90>), accessed September 19, 2013), with maximum tolerance of 1% (threshold) to allow genotyping errors. If a parent–progeny pair conflict was observed or if one or neither parent had been genotyped, genotypes were compared with those of every other animal genotype to determine if there was a parent–progeny relationship. Unique putative parents of the appropriate sex with less than 1% Mendelian conflicts and suitable birthdates were designated as true parents.

We adopted the approach described by Fernandez and Toro (2006) that uses the simulated annealing algorithm in the MOL\_COANC software ((Instituto Nacional de Investigacion Agropecuaria - INIA, Madrid, Spain)) to reconstruct half-sibs families within multiple-sire groups based on observed genomic relationships and then set new half-sibling relationships, when a true sire was not identified with the above describe procedure. Pedigrees were reconstructed by creating a “virtual” ancestor for each identified half-sib family.

A total of 576 changes in relationship were made affecting 1,311 (12.28%) TC records; 96.52% (556) and 3.47% (20) were related to the sire and dam information,

**Table 1.** Pedigree structure as defined by parentage certainty

Parentage	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

respectively. It was observed 23.67% (196) and 13.33% (12) of conflicts, respectively, to genotyped sire–progeny and dam–progeny pairs. Virtual parents were assigned to 2,174 individuals, generating 704 half-sib families, and 12,754 animals remained after pedigree reconstruction and pruning, with a detailed breakdown provided in Table 1.

### Statistical Models

**A Conventional Genomic-Based Single-Step Model.** Consider the following conventional genomic-based single-step model (HBLUP):

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

Here,  $y_{ijk}$  is the  $k$ th phenotypic record of animal  $j$  recorded within CG  $i$ ;  $\boldsymbol{\beta}$  is the vector of fixed effects that includes an overall intercept and 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  $i$  ( $i = 1, \dots, 146$  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 random residual.

The following distributional assumptions were assumed:  $\mathbf{w} = \{w_i\} \sim N(\mathbf{0}, \mathbf{I}\sigma_w^2)$ ,  $\mathbf{a} = \{a_j\} \sim N(\mathbf{0}, \mathbf{H}\sigma_a^2)$ ,  $\mathbf{c} = \{c_j\} \sim N(\mathbf{0}, \mathbf{I}\sigma_c^2)$ , and  $\mathbf{e} = \{e_{ijk}\} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$ , in which  $\sigma_w^2$ ,  $\sigma_a^2$ ,  $\sigma_c^2$ , and  $\sigma_e^2$  represent variances due to CG, additive genetics, permanent environment, and residual terms, respectively. Here,  $\mathbf{I}$  is the identity matrix and  $\mathbf{H}$  represents a matrix that includes genomic information (Legarra et al., 2009; Misztal et al., 2009; Aguilar et al., 2010, 2011).

The inverse of  $\mathbf{H}$  ( $\mathbf{H}^{-1}$ ) was obtained using preGSf90 software (<http://nce.ads.uga.edu/wiki/doku.php?id=readme.pregsf90>), accessed January 6, 2014, and incorporated as a user-defined covariance structure  $\mathbf{H}^{-1}$  in the InterGen software (Cardoso, 2013) to combine genomic information with 1-step reaction norm models as described hereafter.

**A Genomic One-Step Linear Reaction Norm Model.** The model proposed by Su et al. (2006) is

purely Bayesian, because covariates associated with the reaction norm are treated as unknown, thereby allowing inference for all unknowns together within a HBLUP with its reaction norm model extension (genomic 1-step linear reaction norm model [HLRNM]):

$$y_{ijk} = \mathbf{x}'_j \boldsymbol{\beta} + w_i + a_j + b_j w_i + c_j + d_j w_i + e_{ijk} \quad [2]$$

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

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{P}\mathbf{w} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_b\mathbf{b} + \mathbf{Z}_c\mathbf{c} + \mathbf{Z}_d\mathbf{d} + \mathbf{e}, \quad [3]$$

in which  $\mathbf{y} = \{y_{ijk}\}$  is the  $n \times 1$  vector of observations,  $\boldsymbol{\beta}$  is the fixed effects vector of order  $p$ ,  $\mathbf{w} = \{w_i\}_{i=1}^n$  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  $n \times 1$  vector of residuals. Furthermore,  $\mathbf{X}$ ,  $\mathbf{P}$ ,  $\mathbf{Z}_a$ , and  $\mathbf{Z}_c$  are known incidence matrices, whereas the row address of matrices  $\mathbf{Z}_b$  and  $\mathbf{Z}_d$  has exactly 1 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, with all other elements in that row equal to 0 (Su et al., 2006), and  $\mathbf{e}$  is the vector of random residuals.

To infer environmental sensitivities by using a hierarchical Bayesian model, 3 stages are required. According to Su et al. (2006), the first stage defines the distribution of the phenotypic data conditional on all other parameters. The second stage is represented by the prior distributions of the location parameters ( $\boldsymbol{\beta}$ ,  $\mathbf{w}$ ,  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$ , and  $\mathbf{d}$ ).

Finally, the third stage is based on specifying prior distributions for the variance or covariance components. These stages were described in detail in Mota et al. (2016).

To compare the results obtained from pedigree-based and genomic relationship matrices and to investigate the efficiency of genomic EBV (GEBV) prediction across environments through cross-validation, we have used EBV from the ABLUP and its reaction norm extension (1-step linear reaction norm model [ALRNM]). These models were also described in Mota et al. (2016).

**Bayesian Inference and Model Comparison.** The conditional posterior distributions used in Monte Carlo Markov chain algorithms were described in detailed by Cardoso and Tempelman (2012). The Intergen software (Cardoso, 2013) was used considering a total of 1,000,000 cycles, after 100,000 cycles of burn-in, saving every 10th cycle. Global convergence was checked using the Geweke's  $Z$  criterion (Geweke, 1991).

To access the goodness of the fit, a deviance information criterion (DIC) was used (Spiegelhalter et al., 2002):

$$\text{DIC} = \bar{D}(\theta) + p_D = 2\bar{D}(\theta) - D(\bar{\theta}), \quad [4]$$

in which  $\bar{D}(\theta) = E_{\theta|y}[D(\theta)]$  is the posterior expectation of Bayesian deviance;  $p_D = \bar{D}(\theta) - D(\bar{\theta})$  corresponds the penalty for increasing model complexity, in which  $\theta$  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 a better-fitting model.

**Genetic Parameters Estimation and Genomic EBV over Environments.** The additive genetic variance for a specific environment  $i$  with effect  $w_i$  was obtained as follows:

$$\sigma_A^2 | w_i = \text{var}(a_j + b_j w_i) = \sigma_a^2 + w_i^2 \sigma_b^2 + 2w_i \sigma_{ab}. \quad [5]$$

Thus, the heritability, repeatability, and genetic covariance between 2 EG based on covariate values  $w_i$  and  $w_j$  was calculated following Mota et al. (2016).

The GEBV of sire  $j$  specific to a given environment  $i$  was obtained by  $\text{GEBV}_j | w_i + a_j + b_j w_i$  when using HBLUP and HLRNM. In addition, EBV from ABLUP and LRNM was obtained following Mota et al. (2016).

The sire GEBV or EBV were compared by the ranking of the animals obtained by all tested models for low, medium, and high tick infestation levels. Potential differences in reranking of sires for selection were determined via Spearman correlations as previously described by Mota et al. (2016).

### Cross-Validation Study

Cross-validation prediction accuracy was evaluated by 2 different 5-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 correlation between observed ( $y$ ) and predicted phenotypes ( $\hat{y}$ ) in the validation data sets, based on estimates derived from training data sets. The accuracy of the EBV was compared between all the tested models. Therefore, all models were fitted under the same cross-validation scheme, that is, considering the random intercept for ABLUP and HBLUP and random intercept and slope for ALRNM and HLRNM.

Finally, comparisons between models for cross-validation accuracy were based on a randomized complete block design analysis that has each treatment (models fitted) applied in each block (folds) of correlations as response variables. Letting  $y_{ij}$  denote the cross-validation prediction correlation for model

$i = 1, 2, 3, 4$  on cross-validation fold  $j = 1, 2, 3, 4, 5$ , the equation for the model is  $y_{ij} = \mu + \tau_i + b_j + e_{ij}$ , in which  $\mu$  is the overall mean,  $\tau_i$  is the effect of model  $i$ ,  $b_j$  is the random effect associated with fold  $j$ , and  $e_{ij}$  is the random error associated with the experimental unit in fold  $j$  that was analyzed using model  $i$ .

## RESULTS AND DISCUSSION

### Model Comparison via a Deviance Information Criterion

Deviance information criterion values were 3,647.73, 3,115.54, 2,443.07, and 2,055.09 for ABLUP, HBLUP, ALRNM, and HLRNM, respectively. Spiegelhalter et al. (2002) reported that models with differences in DIC values lower than 2 need to be considered as equally good, whereas models with values higher than 2 have been considered as having poorer fit.

Therefore, the models not modeling  $G \times E$  (i.e., ABLUP and HBLUP) appeared to be poorer fitting in comparison with their respective  $G \times E$  extensions (ALRNM and HLRNM), given the smaller DIC values of the latter. Hence, it would reinforce the importance of modeling  $G \times E$  for tick resistance in Hereford and Braford beef cattle. In addition, models using the genomic relationship matrices (HBLUP and HLRNM) yielded smaller DIC values compared with their respective animal model analogs that incorporated the pedigree based additive relationship matrix (ABLUP and ALRNM), confirming the importance of incorporating marker information in genetic evaluations.

### Variance Components and Genetic Parameters under Genotype $\times$ Environment Interactions

The HLRNM lead to lower intercept and slope genetic variance components posterior means but higher permanent variance component posterior means in comparison with ALRNM (Table 2). Furthermore, variance component estimates under HLRNM had lower posterior SD for both effects. Lower genetic variance component estimates based on the H matrix rather than the A matrix were also reported by Veerkamp et al. (2011). According to these authors, this difference may happen due to the scaling methods used to combine the pedigree-based matrix **A** and the genomic relationship matrix (**G**), which in our case was based on equaling the averages of diagonal and off-diagonal elements of **G** and  $A_{22}$  (Vitezica et al., 2011; Christensen et al., 2012). In contrast to this study, Silva et al. (2014) reported similar variance components for the intercept and slope using the **A**

**Table 2.** Posterior means (SD) for the variance components and genetic and permanent environment correlations of the 2-step reaction norm model using pedigree (**A**) and pedigree plus marker information (**H**) relationship matrices

Parameter <sup>1</sup>	Model <sup>2</sup>			
	ALRNM	HLRNM	ABLUP	HBLUP
$\sigma_a^2$	0.024 (0.002)	0.016 (0.002)	0.021 (0.004)	0.015 (0.002)
$\sigma_b^2$	0.037 (0.023)	0.030 (0.011)	N/A <sup>3</sup>	N/A
$\sigma_{ab}$	0.013 (0.006)	0.011 (0.004)	N/A	N/A
$\sigma_c^2$	0.007 (0.002)	0.012 (0.002)	0.010 (0.003)	0.015 (0.02)
$\sigma_d^2$	0.074 (0.027)	0.091 (0.018)	N/A	N/A
$\sigma_{cd}$	0.003 (0.006)	0.009 (0.005)	N/A	N/A
$\sigma_w^2$	0.097 (0.012)	0.093 (0.011)	0.098 (0.012)	0.097 (0.012)
$\sigma_{e1}^2$	0.122 (0.008)	0.129 (0.008)	1.912 (0.248)	1.806 (0.349)
$\sigma_{e2}^2$	0.056 (0.005)	0.057 (0.005)	0.938 (0.124)	0.894 (0.176)
$\sigma_{e3}^2$	0.062 (0.004)	0.065 (0.004)	0.859 (0.115)	0.809 (0.159)
$\sigma_{e4}^2$	0.047 (0.003)	0.049 (0.003)	0.874 (0.120)	0.621 (0.125)
$\sigma_{e5}^2$	0.099 (0.005)	0.102 (0.005)	1.046 (0.138)	1.318 (0.259)
$\sigma_{e6}^2$	0.023 (0.002)	0.022 (0.002)	0.650 (0.089)	0.395 (0.080)
$\sigma_{e7}^2$	0.053 (0.003)	0.052 (0.003)	0.652 (0.087)	0.765 (0.151)
$\sigma_{e8}^2$	0.062 (0.004)	0.062 (0.004)	1.006 (0.134)	0.952 (0.187)
$\sigma_{e9}^2$	0.057 (0.003)	0.056 (0.003)	0.817 (0.109)	0.765 (0.151)
$\sigma_{e10}^2$	0.058 (0.003)	0.056 (0.004)	1.062 (0.143)	0.978 (0.150)
$r_{ab}$	0.476 (0.227)	0.518 (0.184)	N/A	N/A
$r_{cd}$	0.157 (0.259)	0.265 (0.153)	N/A	N/A

<sup>1</sup> $\sigma_a^2$  = reaction norm intercept genetic variance;  $\sigma_b^2$  = reaction norm slope genetic variance;  $\sigma_{ab}$  = genetic covariance between intercept and slope;  $\sigma_c^2$  = reaction norm intercept permanent environment variance;  $\sigma_d^2$  = reaction norm slope permanent environment variance;  $\sigma_{cd}$  = permanent environment covariance between intercept and slope;  $\sigma_w^2$  = environmental variance;  $\sigma_{e1}^2$  = residual class 1;  $\sigma_{e2}^2$  = residual class 2;  $\sigma_{e3}^2$  = residual class 3;  $\sigma_{e4}^2$  = residual class 4;  $\sigma_{e5}^2$  = residual class 5;  $\sigma_{e6}^2$  = residual class 6;  $\sigma_{e7}^2$  = residual class 7;  $\sigma_{e8}^2$  = residual class 8;  $\sigma_{e9}^2$  = residual class 9;  $\sigma_{e10}^2$  = residual class 10;  $r_{ab}$  = genetic correlation between intercept and slope;  $r_{cd}$  = permanent environment correlation between intercept and slope.

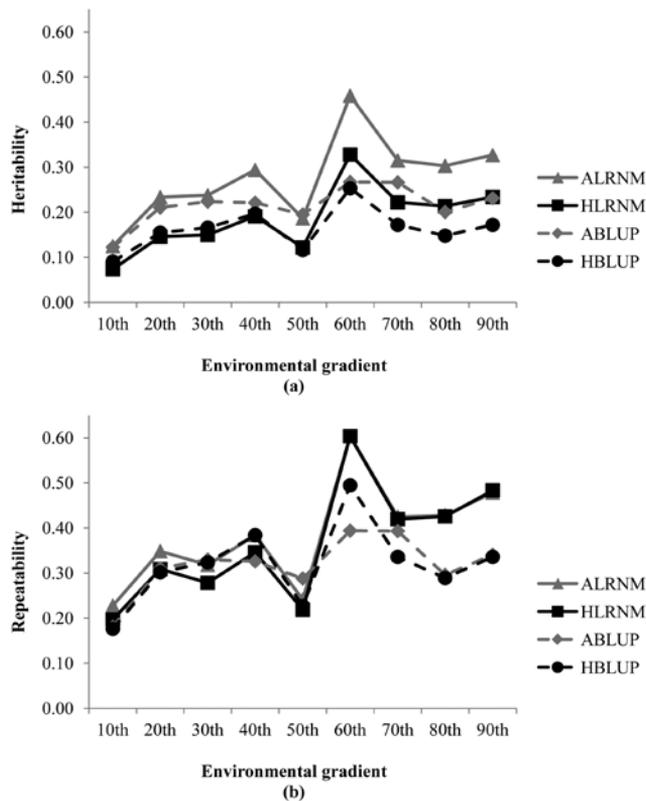
<sup>2</sup>ALRNM = 1-step linear reaction norm model; HLRNM = genomic 1-step linear reaction norm model; ABLUP = linear animal model; HBLUP = conventional genomic-based single-step model.

<sup>3</sup>N/A = not applicable.

and **G** matrices fitting a 2-step random regression reaction norm model in pigs.

In general, the residual class variances were slightly higher for HLRNM compared with ALRNM, with no clear pattern being noticeable for both models (Table 2). Cardoso and Tempelman (2012), working with  $G \times E$  models in postweaning gains in Angus cattle, also observed that the residual variance did not monotonically increase over the EG.

Under the nongenomic 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 a great deal of uncertainty, as indicated by the posterior SD for the permanent environment correlation (Table 2). However, the HLRNM analyses lead to es-

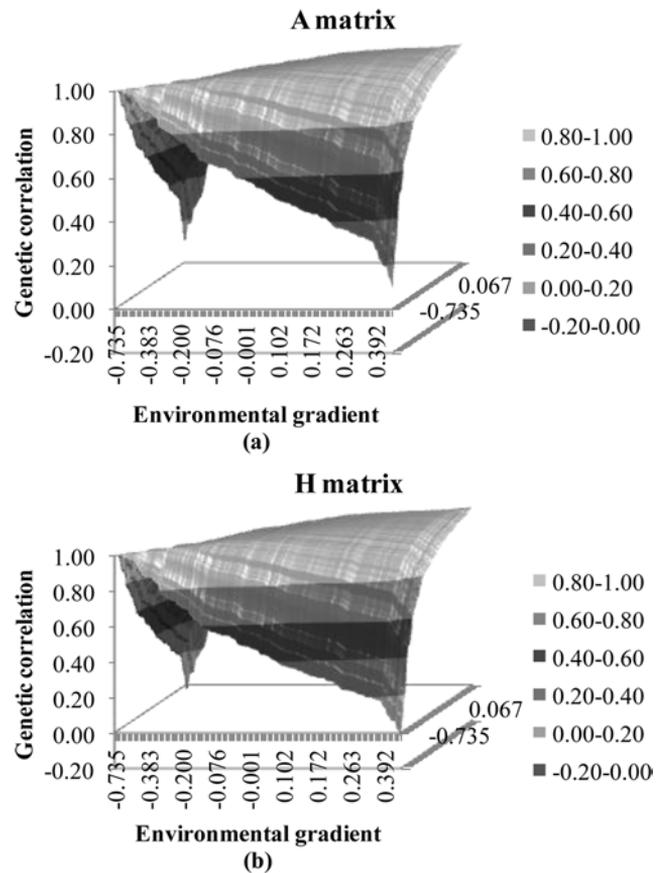


**Figure 1.** Posterior means heritabilities (a) and repeatabilities (b) for environments between the 10th and 90th percentiles of the environmental gradient obtained for Hereford and Braford tick counts by a 1-step linear reaction norm model based on pedigree (1-step linear reaction norm model [ALRNM]), a genomic 1-step linear reaction norm model (HLRNM), a traditional animal model based on pedigree (linear animal model [ABLUP]), and a conventional genomic-based single-step model (HBLUP).

timated positive correlations with high and moderate magnitude for additive genetic and permanent environment effects, respectively. 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 components were slightly higher in ALRNM compared with HLRNM (Table 2). The heritability estimates ( $h^2$ ) over the 10th and 90th percentiles are shown in Fig. 1a. 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).

Heritability estimates were higher for reaction norm models compared with conventional animal models regardless of genomic information (Fig. 1a). This reinforces the use of reaction norm models as a powerful alternative in genetic evaluation of this population. Additionally,  $h^2$  estimates were higher for ALRNM in all tick infestation levels (Fig. 1a), although  $h^2$  estimates from HLRNM had smaller posterior SD. The larger uncertainty about estimated parameters in the nongenomic approach (ALRNM) may



**Figure 2.** Genetic correlations for Hereford and Braford tick resistance performance on different environmental conditions obtained by the 1-step linear reaction norm models considering pedigree (a) and pedigree + genomics (b).

be due to incomplete pedigree information from multiple-sire matings. Veerkamp et al. (2011) also found smaller SE of  $h^2$  using a SNP plus pedigree-based as opposed to pedigree-based relationships, even though  $h^2$  estimates were smaller in the latter. These authors have reported that may happen due to tight management and/or relatively homogeneous groups of animals, which may be also a reason for the present study. Our 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 reaction norm model, respectively.

The repeatability estimates varied along the EG (range 0.18–0.60) and were, in general, similar under approaches but higher for reaction norm extensions (Fig. 1b). These results demonstrate the importance of considering permanent environment effects; the higher estimated repeatabilities in harsh environments indicate that more resistant animals are more likely to maintain a consistent performance in their resistance in harsher environments than in favorable environments (i.e., low tick infestation).

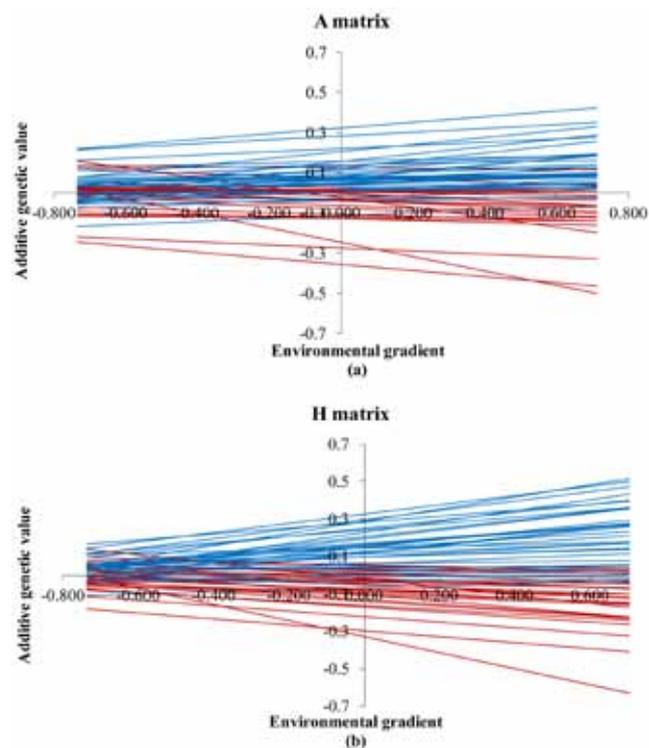
Genetic correlations along the EG were remarkably low between the extreme EG, whereas EG with

very similar values had high genetic correlations regardless whether ALRNM (Fig. 2a) or HLRNM was implemented (Fig. 2b). In addition, negative correlations could be observed between extreme EG under a genomic approach, which may indicate substantial  $G \times E$  for tick resistance. However, low genetic correlations within models could be artifacts created by few records or skewed distributions in extreme environments using reaction norm models. Cardoso and Tempelman (2012) also reported low genetic correlations between extreme EG under a nongenomic approach for postweaning gains in Angus cattle.

Rank correlations among posterior means of the genetic merit predictions ( $a_j$ ) under both approaches, obtained by the animal models (ABLUP and HBLUP) with those ( $g_j|w_j$ ) obtained by their extensions (ALRNM and HLRNM), are shown in Table 3. 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 quite different between all tested models.

However, differences in environmental sensitivity did not result in many rerankings of the top 10% most-used sires (>12 progeny) at different CG levels for both approaches. Under a genomic approach, the sires presented more outstanding breeding value variation across the EG (Fig. 3). Nevertheless, Fig. 3 also demonstrates that genetic merit also depends on EG under a genomic approach. Genomic EBV differences between animals decrease with a low EG, not revealing a complex  $G \times E$ . It further indicates the difficulty in identifying superior breeding stock in low tick infestation environments.

Furthermore, once correlations across **A** and **H** matrices were lower than within approaches, the impact



**Figure 3.** Genetic tick resistance reaction norms of 10% most used (large number of progeny; >12) Hereford and Braford sires obtained by the 1-step linear reaction norm models considering pedigree (a) and pedigree + genomics (b). Animals with positive slopes are represented by the color blue. Animals with negative slopes are represented by the color red. **A** matrix = the numerator relationship matrix based on pedigree; **H** matrix = a matrix that includes genomic information.

on rankings of introducing marker information is relevant because correlations within a genomic approach were higher than those within a nongenomic approach (Table 4). Finally, losses on selection precision by using a traditional animal model would not be expected

**Table 3.** Spearman rank correlations<sup>1</sup> 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 <sup>2</sup>							
	ABLUP	ALRNM	ALRNM	ALRNM	HLRNM	HLRNM	HLRNM	HBLUP
	Environmental level <sup>3,4</sup>							
	Ov.	LTI	MTI	HTI	LTI	MTI	HTI	Ov.
ABLUP (Ov.)		0.96	0.97	0.96	0.69	0.64	0.62	0.68
ALRNM (LTI)	0.90		0.97	0.94	0.72	0.65	0.62	0.67
ALRNM (MTI)	0.93	0.94		0.99	0.71	0.68	0.66	0.68
ALRNM (HTI)	0.91	0.88	0.99		0.70	0.68	0.67	0.68
HLRNM (LTI)	0.66	0.72	0.72	0.70		0.96	0.93	0.95
HLRNM (MTI)	0.61	0.62	0.68	0.69	0.96		0.99	0.97
HLRNM (HTI)	0.59	0.58	0.66	0.69	0.93	1.00		0.95
HBLUP (Ov.)	0.68	0.61	0.66	0.67	0.93	0.96	0.95	

<sup>1</sup>Correlations between all animals above the diagonal and between the most used sires (larger number of progeny) below the diagonal.

<sup>2</sup>ABLUP = linear animal model; ALRNM = 1-step linear reaction norm model; HLRNM = genomic 1-step linear reaction norm model; HBLUP = conventional genomic-based single-step model.

<sup>3</sup>LTI = low tick infestation; MTI = medium tick infestation; HTI = high tick infestation; Ov. = overall.

<sup>4</sup>LTI represents the 10th (−0.396), MTI represents the 50th (0.020), and HTI represents the 90th (0.320) percentiles of the environmental gradient.

**Table 4.** The  $P$ -values ( $P > |t|$ ) in a  $t$  test for contrast of each difference of least squares means for the accuracy  $r_{y,\hat{y}}$  between models for  $K$ -means and random partitioning using randomized complete block design (RCBD) analysis

Model/strategy <sup>1</sup>	$K$ -means	Random
ABLUP vs. HBLUP	0.2762	0.0139
ABLUP vs. HLRNM	0.5725	0.0010
ABLUP vs. ALRNM	0.8273	0.0056
HBLUP vs. HLRNM	0.5853	0.2241
HBLUP vs. ALRNM	0.3767	0.6701
HLRNM vs. ALRNM	0.7271	0.4183

<sup>1</sup>ABLUP = linear animal model; ALRNM = 1-step linear reaction norm model; HBLUP = conventional genomic-based single-step model; HLRNM = genomic 1-step linear reaction norm model.

to be substantial in both approaches due to similarity correlation of magnitude in low, medium, and also high EG. Cardoso and Tempelman (2012) reported different results under nongenomic models where one would expect losses by using conventional models such as ABLUP in this study across EG. Although we observed slight correlation values when we filtered sires with large numbers of progeny (top 10%; Table 3), the same pattern was observed, as mentioned before, for the total number of animals in the data set.

### Prediction Ability via Cross-Validation

Cross-validation prediction accuracies ( $r_{y,\hat{y}}$ ) within each of the tested models (ABLUP, HBLUP, HLRNM, and ALRNM) were effective, being higher than 0.55 in  $K$ -means and 0.59 in random partitioning strategies (Fig. 4).

Cross-validation estimates were, on average,  $0.66 \pm 0.02$ ,  $0.67 \pm 0.02$ ,  $0.67 \pm 0.02$ , and  $0.66 \pm 0.02$  for ABLUP, HBLUP, HLRNM, and ALRNM, respectively, based on  $K$ -means partitioning. For 5-fold random partitioning, HLRNM ( $0.71 \pm 0.01$ ) was statisti-

cally different from ABLUP ( $0.67 \pm 0.01$ ). However, no statistical significance was reported when considering HBLUP ( $0.70 \pm 0.01$ ) and ALRNM ( $0.70 \pm 0.01$ ; Fig. 4; Table 4). In this context, although HLRNM presented the smallest DIC, prediction accuracies from this model were similar to other tested models. In part, this can be explained by the number of genotyped animals, which may not have been enough to increase prediction accuracy. In contrast with this study, Silva et al. (2014) found higher genomic prediction accuracies and genetic correlations for reaction norm models compared with a standard sire model in pigs.

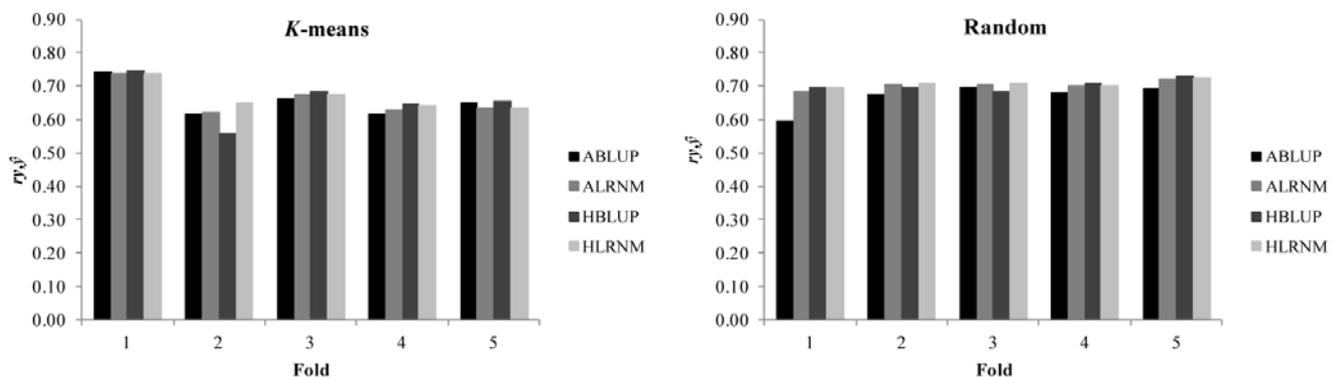
Finally, smaller values for  $r_{y,\hat{y}}$  presented by  $K$ -means compared with random partitioning were expected and thereby reinforce that prediction accuracy deteriorate as the relationship between animals decreases (Saatchi et al., 2011). These authors reported lower accuracies for  $K$ -means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

In general, one of the main contributions of the present study was to draw interest toward genome-wide selection models that consider  $G \times E$ . Despite the fact that the EG considered in the present study is related to CG effects in one region of Brazil, the proposed 1-step reaction norms methodology can easily accommodate other traits and environments such as across country or multiple across country evaluations.

### Conclusions

Although the complex genotype  $\times$  environment interaction was not reported, reaction norm models might be used 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



**Figure 4.** Accuracy of genomic selection obtained by  $K$ -means and random partitioning based on a 5-fold cross-validation study using genotype  $\times$  environment interaction (1-step linear reaction norm model [ALRNM] and a genomic 1-step linear reaction norm model [HLRNM]) and conventional animal models (linear animal model [ABLUP] and a conventional genomic-based single-step model [HBLUP]).  $r_{y,\hat{y}}$  = cross-validation accuracy.

the relationship between animals in training and validation data sets decreased.

## LITERATURE CITED

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743–752. doi:10.3168/jds.2009-2730
- Aguilar, I., I. Misztal, A. Legarra, and S. Tsuruta. 2011. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. *J. Anim. Breed. Genet.* 128:422–428. doi:10.1111/j.1439-0388.2010.00912.x
- Ambrosini, D. P., P. L. S. Carneiro, J. Braccini Neto, C. H. M. Malhado, R. Martins Filho, and F. F. Cardoso. 2012. Genotype  $\times$  environment interaction for yearling weight in polled Nellore cattle in northeast Brazil. *Pesq. Agropecu. Bras.* 47:1489–1495. doi:10.1590/S0100-204X2012001000011
- Ayres, D. R., J. R. Pereira, A. A. Boligon, F. F. Silva, F. S. Schenkel, V. M. Roso, and L. G. Albuquerque. 2013. Linear and Poisson models for genetic evaluation of tick resistance in cross-bred Hereford  $\times$  Nellore cattle. *J. Anim. Breed. Genet.* 130:417–424. doi:10.1111/jbg.12036
- Biegelmeier, P. 2012. Genetic resistance to *Rhipicephalus* (Boophilus) *microplus* natural and artificial infestation in Hereford and Braford cattle. MS Diss., Federal University of Pelotas, Pelotas, Brazil.
- Budeli, M. A., K. A. Nephawe, D. Norris, N. W. Selepa, L. Bergh, and A. Maiwashe. 2009. Genetic parameter estimates for tick resistance in Bonsmara cattle. *S. Afr. J. Anim. Sci.* 39:321–327.
- Burrow, H. M. 2001. Variances and covariances between productive and adaptive traits and temperament in a composite breed of tropical beef cattle. *Livest. Prod. Sci.* 70:213–233. doi:10.1016/S0301-6226(01)00178-6
- Calus, M. P. L., A. F. Groen, and G. de Jong. 2002. Genotype  $\times$  environment interaction for protein yield in Dutch dairy cattle as quantified by different models. *J. Dairy Sci.* 85:3115–3123. doi:10.3168/jds.S0022-0302(02)74399-3
- Cardoso, F. F. 2013. Application of Bayesian inference in animal breeding using the InterGen program: Manual of version 1.22. Documents 112. Embrapa South Livestock, Bagé, Brazil.
- Cardoso, L. L., J. Braccini Neto, F. F. Cardoso, J. A. Cobuci, I. O. Biassus, and J. O. J. Barcellos. 2011. Hierarchical Bayesian models for genotype  $\times$  environment interaction estimation via reaction norms applied to post-weaning gain of Hereford cattle. *R. Bras. Zootec.* 40:294–300. doi:10.1590/S1516-35982011000200009
- Cardoso F. F., C. C. G. Gomes, B. P. Sollero, M. M. Oliveira, V. M. Roso, M. L. Piccoli, R. H. Higa, M. J. Yokoo, A. R. Caetano, and I. Aguilar. 2015. Genomic prediction for tick resistance in Braford and Hereford cattle. *J. Anim. Sci.* 93:1-13.
- Cardoso, F. F., and R. J. Tempelman. 2012. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. *J. Anim. Sci.* 90:2130–2141. doi:10.2527/jas.2011-4333
- Christensen, O. F., P. Madsen, B. Nielsen, T. Ostensen, and G. Su. 2012. Single-step methods for genomic evaluation in pigs. *Animal* 6:1565–1571. doi:10.1017/S1751731112000742
- Clayton, D. 2012. snpStats: SnpMatrix and XSnMatrix classes and methods. <https://www-gene.cimr.cam.ac.uk> (Accessed 30 October 2012.).
- Corrêa, M. B. B., N. J. L. Dionello, and F. F. Cardoso. 2010. Genetic evaluation of Devon cattle using a reaction norms model. *R. Bras. Zootec.* 39:128–133. doi:10.1590/S1516-35982010000100017
- Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman Group Ltd., Harlow, UK.
- Fernandez, J., and M. A. Toro. 2006. A new method to estimate relatedness from molecular markers. *Mol. Ecol.* 15:1657–1667. doi:10.1111/j.1365-294X.2006.02873.x
- Forni, S., I. Aguilar, and I. Misztal. 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genet. Sel. Evol.* 43:1. doi:10.1186/1297-9686-43-1
- Geweke, J. 1991. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: J. O. Bernardo, J. M. Berger, A. P. Dawid, and A. F. M. Smith, editors, Bayesian statistics 4. Oxford Univ. Press, Oxford, UK. p. 164–193.
- Kolmodin, R., E. Strandberg, P. Madsen, J. Jensen, and H. Jorjani. 2002. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. *Acta Agric. Scand., Sect. A* 52:11–24.
- Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92:4656–4663. doi:10.3168/jds.2009-2061
- Mattar, M., L. O. C. Silva, M. M. Alencar, and F. F. Cardoso. 2011. Genotype by environment interaction for long-yearling weight in Canchim cattle quantified by reaction norm analysis. *J. Anim. Sci.* 89:2349–2355. doi:10.2527/jas.2010-3770
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome wide dense marker maps. *Genetics* 157:1819–1829.
- Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J. Dairy Sci.* 92:4648–4655. doi:10.3168/jds.2009-2064
- Mota, R. R., R. J. Tempelman, P. S. Lopes, I. Aguilar, F. F. Silva, and F. F. Cardoso. 2016. Genotype by environment interaction for tick resistance of Hereford and Braford beef cattle using reaction norm models. *Genet. Sel. Evol.* 48:3. doi:10.1186/s12711-015-0178-5
- Oliveira, M. M., C. C. G. Gomes, V. M. Roso, A. R. Caetano, I. Aguilar, and F. F. Cardoso. 2012. Genomic selection for tick resistance in Braford and Hereford cattle. In: Proc. 9th Simpósio Brasileiro de Melhoramento Animal, João Pessoa, Brazil. p. 1-3.
- Pégolo, N. T., H. N. Oliveira, L. G. Albuquerque, L. A. F. Bezerra, and R. B. Lôbo. 2009. Genotype by environment interaction for 450-day weight of Nelore cattle analyzed by reaction norm models. *Genet. Mol. Biol.* 32:281–287. doi:10.1590/S1415-4752009005000027
- Rechav, Y., J. Dauth, and D. A. Els. 1990. Resistance of Brahman and Simmental cattle to southern African ticks. *Onderstepoort J. Vet. Res.* 57:7–12.
- Roso, V. M., and F. S. Schenkel. 2006. AMC – A computer program to assess the degree of connectedness among contemporary groups. In: Proc. 8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, Brazil. p. 26-27.
- Saatchi, M., M. C. McClure, S. D. McKay, M. M. Rolf, J. Kim, J. E. Decker, T. M. Taxis, R. H. Chapple, H. R. Ramey, S. L. Northcutt, S. Bauck, B. Woodward, J. C. M. Dekkers, R. L. Fernando, R. D. Schnabel, D. J. Garrick, and J. F. Taylor. 2011. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. *Genet. Sel. Evol.* 43:40. doi:10.1186/1297-9686-43-40
- Sargolzaei, M., J. Chesnais, and F. S. Schenkel. 2011. FImpute – An efficient imputation algorithm for dairy cattle populations. *J. Anim. Sci.* 89(E-Suppl. 1):333.

- Shariati, M. M., G. Su, P. Madsen, and D. Sorensen. 2007. Analysis of milk production traits in early lactation using a reaction norm model with unknown covariates. *J. Dairy Sci.* 90:5759–5766. doi:10.3168/jds.2007-0048
- Silva, F. F., H. A. Mulder, E. F. Knol, M. S. Lopes, S. E. F. Guimarães, P. S. Lopes, P. K. Mathur, J. M. S. Viana, and J. W. M. Bastiaansen. 2014. Sire evaluation for total number born in pigs using a genomic reaction norms approach. *J. Anim. Sci.* 92:3825–3834. doi:10.2527/jas.2013-6486
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc., B* 64:583–616. doi:10.1111/1467-9868.00353
- Su, G., P. Madsen, M. S. Lund, D. Sorensen, I. R. Korsgaard, and J. Jensen. 2006. Bayesian analysis of the linear reaction norm model with unknown covariates. *J. Anim. Sci.* 84:1651–1657. doi:10.2527/jas.2005-517
- Veerkamp, R. F., H. A. Mulder, R. Thompson, and M. P. L. Calus. 2011. Genomic and pedigree-based genetic parameters for scarcely recorded traits when some animals are genotyped. *J. Dairy Sci.* 94:4189–4197. doi:10.3168/jds.2011-4223
- Vitezica, Z., I. Aguilar, I. Misztal, and A. Legarra. 2011. Bias in genomic predictions for 860 populations under selection. *Genet. Res.* 93:357–366. doi:10.1017/S001667231100022X
- Wharton, R. H., and K. B. W. Utech. 1970. The relation between engorgement and dropping of *Boophilus microplus* (Canestrini) (Ixodidae) to the assessment to tick numbers on cattle. *J. Aust. Entomol. Soc.* 9:171–182. doi:10.1111/j.1440-6055.1970.tb00788.x
- Wiggans, G. R., T. S. Sonstegard, P. M. VanRaden, L. K. Matukumalli, R. D. Schnabel, J. F. Taylor, F. S. Schenkel, and C. P. Van Tassell. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. *J. Dairy Sci.* 92:3431–3436. doi:10.3168/jds.2008-1758