

**Genomic Wide-Selection for Tick Resistance
in Hereford and Braford Cattle via Reaction Norm Models**

R.R. Mota^{*,†}, R.J. Tempelman^{*}, F.F. Cardoso[‡], I.G. Aguilar[§], P.S. Lopes[†]
^{*}Michigan State University, [†]Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil, [‡]Embrapa
Southern Region Animal Husbandry, Bagé, Rio Grande do Sul, Brazil, [§]Instituto Nacional de
Investigación Agropecuaria - INIA Las Brujas- Canelones, Uruguay.

ABSTRACT: The objective of this study was to compare a conventional genomic model (GBLUP) and its extension to a linear reaction norm model (GLRNM) specifying genotype by environment interaction (G*E) for tick resistance in Brazilian cattle. Tick counts (TC) from 4,363 Hereford and Braford cattle from 146 contemporary groups (CG) were available of which 3,591 animals had BovineSNP50 Illumina v2 BeadChip genotypes. The reaction norm covariate was based on CG estimates of TC from a first-step model. Analysis was conducted based on adapting the single step GBLUP/REML procedure. Five-fold cross validation based on K-means and random partitioning was used to compare the fit of the two models. Cross validation correlations were strong and not significantly different between models for either partitioning strategy. Nevertheless, it seems apparent that G*E for tick infestation exists and can be captured by GLRNM models.

Keywords: cross validation; single-step; tick counts

Introduction

In beef cattle production the bovine tick *Rhipicephalus (Boophilus) microplus* can greatly diminish animal performance beyond parasite resistance due to repeated use of treatments with acaricides. Furthermore, repeated failures of effective vaccines development had driven researchers to seek for alternative solutions.

Previous work have demonstrated the existence of genetic variability for tick resistance (Budeli et al. (2009); Oliveira et al. (2012)) with estimated heritabilities ranging from 0.05 to 0.42 thereby implying the feasibility of genetic improvement for this trait. Mota et al. (2013) has recently suggested genotype by environment interaction (G*E) exists for this trait of a nature that can be captured using linear reaction norm models (LRNM).

Genome wide selection (GWS) is a powerful tool that could be further used to enhance the accuracy of genomic prediction on tick resistance. So far, there has been little work incorporating GWS information in LRNM which might be necessary to fine tune GWS for specific environments. The objective of this study was to compare a conventional genomic-based BLUP (GBLUP) model (GBLUP) to its genomic LRNM extension (GLRNM) using tick infestation data.

Material and methods

Phenotypes were obtained from genetic breeding program conducted by Delta G Connection and consisted of tick counts (TC) from 4,363 Hereford and Braford cattle

born between 2008 and 2011. TC were obtained between 326 and 729 days of age, consisting of the count of all engorged female ticks ≥ 4.5 mm on one side of the animal. Tick infestations were acquired naturally and from one to three TC were collected on each animal with a minimum interval of 30 days between each TC. The variable used for analysis was log transformed such that, $LTTC = \log_{10}(TC+1.001)$.

Genotypes based on the BovineSNP50 Illumina v2 BeadChip were acquired on 3,591 of these cattle. After various quality control edits, 41,045 SNPs remained and were used to estimate genomic relationship coefficients between animals. Genomic information was combined together with pedigree information using the single-step procedure of Aguilar et al. (2010) where the genetic relationships based on pedigree are adjusted for deviations due to genomic information. Cross validation prediction accuracy was evaluated by two 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 whereas the other strategy was based on random partitions of training and validation data sets.

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.

$$r_{y,\hat{y}} = \frac{\sigma_{y,\hat{y}}}{\sqrt{\sigma_y^2 * \sigma_{\hat{y}}^2}}$$

The fit of the GBLUP and GLRNM were compared using Akaike Information Criterion (AIC). In order to implement GLRNM, solutions to the 146 random CG effects from GBLUP were used as “known” covariates for the reaction norm specification in the GLRNM (Kolmodin et al. (2002)). Both GBLUP and GLRNM fitted additive genetic, permanent environment and CG level (herd-year-season-sex-management) as random intercept effects and Nellore percentage, heterozygosity, epistatic loss and linear and quadratic effects of age as covariates whereas GLRNM additionally included random genetic and permanent environmental slope effects on the reaction norm, allowing for covariances between these slopes and the corresponding random intercepts.

The heritability ($h^2|w_i$) and repeatability ($r|w_i$) for a specific environment were calculated, respectively as follow:

$$h^2 | w_i = \frac{\sigma_{ge}^2 | w_i}{\sigma_{ge}^2 | w_i + \sigma_e^2}$$

$$r | w_i = \frac{\sigma_{ge}^2 | w_i + \sigma_{pe}^2 | w_i}{\sigma_{ge}^2 | w_i + \sigma_{pe}^2 | w_i + \sigma_e^2},$$

where $\sigma_{ge}^2 | w_i$ and $\sigma_{pe}^2 | w_i$ are the additive genetic and permanent environment variances, respectively, for a specific environment i , whereas σ_e^2 denotes the residual variance. Here, w_i was defined as contemporary groups (CG) estimates (\hat{w}_i) obtained by GBLUP analysis. The environmental burdens (EB) were divided as low, medium and high tick infestation. The strategy chosen was based on values less than the 10th (i.e., $\hat{w}_i \leq -0.408$ LTTC), between the 10th and 50th (i.e., $-0.408 \leq \hat{w}_i \leq 0.035$ LTTC) and greater than 90th (i.e., $\hat{w}_i \geq 0.320$ LTTC) percentiles as low, medium and high tick infestation, respectively. In addition, the genetic merit and producing abilities (sum of additive and PE effects) of TC for each animal were also predicted.

The estimation of (co)variance components and genetic parameters were performed by restricted maximum likelihood (REML; Patterson and Thompson, (1971)) utilizing airemlf90 software (<http://nce.ads.uga.edu/>). Furthermore, preGSf90 software from this same source was used to calculate the genomic relationship coefficients.

Results and discussion

GBLUP appeared to be a poorer fitting model compared to GLRNM based on a lower Akaike's criterion (5589.1214 vs. 5407.2545). Hence, it seems necessary to consider G*E for genetic evaluations of tick resistance in Hereford and Braford beef cattle.

GBLUP and GLRNM intercept variances for additive genetic and permanent environment effects were rather similar, as anticipated, whereas significant slope variance components were determined using GLRNM, thereby again indicating G*E to be driven by EB as defined by TC (Table 1). However, the GLRNM model did not present considerable re-rankings of genetic merit across environments (Figure 1) suggesting perhaps that use of GLRNM would not accelerate genetic progress in specific environments relative to the use of GBLUP, given that the same breeding stock would be selected for all environments by both models. Nevertheless, Figure 1 also demonstrates that genetic merit also depends upon EB and that differences in GEBV between animals decrease with a low EB. It further indicates the difficulty to identify superior breedstock in better (i.e., low EB) environments. Furthermore, although G*E did not indicate significant re-rankings in genetic merit, the same was not true for producing ability. The negative covariance between intercept and slope for permanent environment effect lead to substantial re-ranking between animals (Figure 2).

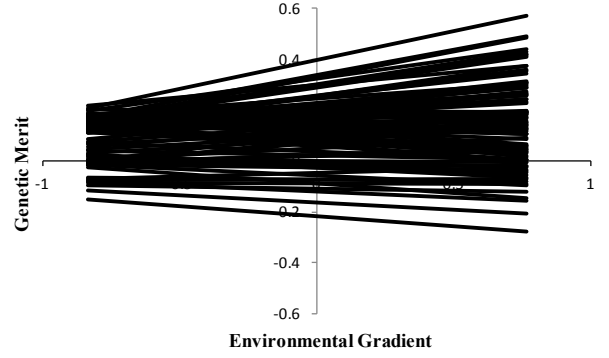


Figure 1: Genetic merit reaction norms of all animals in data set.

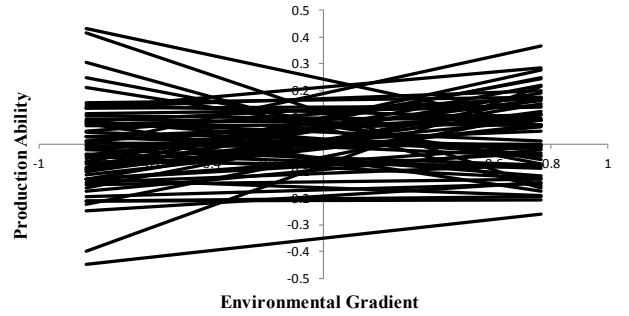


Figure 2 Production ability reaction norms of all animals in data set.

The use of homocedastic error models as those fitted in this study, could lead to biased inferences on variance components and hence genetic parameters. Thus, fitting more complex models such as those considering heterogeneity of residual variance and hierarchical Bayesian models based on 1-step approach (Su et al., (2006)) which treat the covariate associated with the reaction norm as unknown and allows inferring for all unknowns parameters together, may result in better-fitting models and help better describe the presence of G*E for tick resistance.

Table 1. Variance component (VC) estimates and standard errors for the parameters of conventional genomic (GBLUP) and 2-step linear reaction norm model (GLRNM).

VC	GBLUP	GLRNM
σ_{ige}^2 ¹	0.014 (0.002)	0.014 (0.002)
σ_{sge}^2 ²	N/A	0.020 (0.010)
σ_{isge} ³	N/A	0.004 (0.003)
σ_{ipe}^2 ⁴	0.015 (0.002)	0.014 (0.002)
σ_{spe}^2 ⁵	N/A	0.077 (0.013)
σ_{ispe} ⁶	N/A	-0.013 (0.003)
σ_e^2 ⁷	0.072 (0.001)	0.064 (0.001)

¹ σ_{ige}^2 : additive genetic intercept variance

² σ_{sge}^2 : additive genetic slope variance

³ σ_{isge} : additive genetic covariance between intercept and slope

⁴ σ_{ipe}^2 : permanent environment intercept variance

⁵ σ_{spe}^2 : permanent environment slope variance

⁶ σ_{ispe} : permanent environment covariance between intercept and slope

⁷ σ_e^2 : residual variance

Heritability estimates (h^2) as a function of CG levels slightly differed from GBLUP ($h^2= 0.14$) with $h^2=0.12$ for low environmental burden (TC= -0.408), $h^2=0.13$ for medium environmental burden (TC= 0.035), $h^2=0.15$ for high environmental burden (TC= 0.320). These results indicate that selection response might only change slightly according to the model fitted, GBLUP or GLRNM and also within GLRNM burdens.

The repeatabilities estimates were 0.30 under GBLUP and 0.44, 0.44 and 0.46 for GLRNM under low, medium and high environmental burdens which suggest that more than one TC on the body side is useful to model non-genetic and/or permanent environmental effects that might lead to the producing ability re-rankings across environments presented earlier (Figure 2), thereby having rather strong implications for management strategies.

Cross-validation correlations ($r_{y\hat{y}}$) within each of GBLUP and GLRNM were strong and the means and standard errors across replicates are presented in Table 2. Those correlations tended towards statistical significance between the two models under the K-means cross-validation partitioning strategy ($P=0.07$) whereas there was no evidence of a difference based on random partitioning ($P=0.18$). These results may reflect poor power with a low number (5) of cross-validation replicates. Silva et al. (2014) found higher genomic prediction accuracies for reaction norm models compared to a standard animal model in pigs. These authors also reported accuracy differences among EB with higher values in intermediate burdens.

Table 2. Cross validation correlations between observed (y) and estimated (\hat{y}) phenotypes, means and standard errors in a 5-fold cross validation using K-means and random partitioning (in bold) for conventional genomic model (GBLUP) and a 2-step reaction norm model (GLRNM).

M ¹ /F ²	F1	F2	F3	F4	F5	Mean
	0.75	0.64	0.68	0.65	0.66	0.68±0.02
GBLUP	0.70	0.71	0.71	0.71	0.73	0.71±0.01
	0.74	0.64	0.67	0.64	0.66	0.67±0.02
GLRNM	0.69	0.71	0.71	0.70	0.73	0.71±0.01

¹M: model

²F: fold

Genomic accuracy of predictions were generally higher using random than K-means partitioning for GBLUP and GLRNM (Table 1). This may be due to a greater occurrence of cross validation within CG using random partitioning with strong genetic relationships for animals between training and validation populations compared to a cross validation across CG. These values are in agreement with Saatchi et al. (2011) which reported substantial accuracy variation between groups in which lowest accuracies were found in groups which individuals were less related each other.

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. Our results suggest that the accuracy of prediction will decrease as the tick infestation level increases and as the relationship between animals in training data set with validation dataset decreases.

Acknowledgments

The authors thank Delta G Connection by providing the data used in this research; Embrapa Southern Region Animal Husbandry and Michigan State University for theoretical and technical support; CAPES, CNPq and FAPEMIG by granting the scholarship.

Literature cited

- Aguilar, I., Misztal, I., Johnson, D.L. et al. (2010). *J. Dairy Sci.*, 93:743-752.
- Budeli, M.A., Nephawe, K.A., Norris, D. et al. (2009). *S. Afr. J. Anim. Sci.*, 39:321-327.
- Kolmodin, R., Strandberg, E., Madsen, P. et al. (2002). *Acta Agri. Scand. A-An.*, 52:11-24.
- Mota, R.R., Cardoso F.F., Lopes P.S. et al. (2013). Proc 10th Brazilian Animal Breeding Symposium.
- Oliveira, M.M., Gomes, C.C.G., Roso, V.M. et al. (2012). Proc 9th Brazilian Animal Breeding Symposium.
- Patterson, H.D., Thompson, R. (1971). *Biometrics*. 58:545-554.
- Saatchi, M., McClure, M.C., McKay, S.D. et al. (2011). *Genet. Sel. Evol.* 43:40.
- Silva, F.F., Mulder, H.A., Knol, E.F. et al. (2014). *J. Anim. Sci.*, published online February 3.
- Su, G., Madsen, P., Lund, M.S. et al. (2006). *J. Anim. Sci.*, 84:1651-1657.