

# Genome-Enabled Prediction of Breeding Values for Feedlot Average Daily Weight Gain in Nelore Cattle

Adriana L. Somavilla,\* Luciana C. A. Regitano,<sup>†</sup> Guilherme J. M. Rosa,\*<sup>‡</sup> Fabiana B. Mokry,<sup>§</sup>

Mauricio A. Mudadu,<sup>†</sup> Polyana C. Tizoto,<sup>§</sup> Priscila S. N. Oliveira,<sup>§</sup> Marcela M. Souza,<sup>§</sup>

Luiz L. Coutinho,\*\* and Danísio P. Munari\*<sup>1</sup>

\*Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal 14884-900, Brazil, <sup>†</sup>Embrapa Pecuária Sudeste, São Carlos 13560-970, Brazil, <sup>‡</sup>Department of Animal Science, University of Wisconsin–Madison, Wisconsin 53706, <sup>§</sup>Departamento de Genética e Evolução, Universidade Federal de São Carlos, 13565-905, Brazil, and <sup>\*\*</sup>Departamento de Zootecnia, Universidade de São Paulo, Piracicaba 13418-900, Brazil

ORCID IDs: 0000-0002-0718-4856 (A.L.S.); 0000-0001-9242-8351 (L.C.A.R.); 0000-0001-9172-6461 (G.J.M.R.); 0000-0002-2767-2507 (M.A.M.); 0000-0001-6915-038X (D.P.M.)

**ABSTRACT** Nelore is the most economically important cattle breed in Brazil, and the use of genetically improved animals has contributed to increased beef production efficiency. The Brazilian beef feedlot industry has grown considerably in the last decade, so the selection of animals with higher growth rates on feedlot has become quite important. Genomic selection (GS) could be used to reduce generation intervals and improve the rate of genetic gains. The aim of this study was to evaluate the prediction of genomic-estimated breeding values (GEBV) for average daily weight gain (ADG) in 718 feedlot-finished Nelore steers. Analyses of three Bayesian model specifications [Bayesian GBLUP (BGBLUP), BayesA, and BayesC $\pi$ ] were performed with four genotype panels [Illumina BovineHD BeadChip, TagSNPs, and GeneSeek High- and Low-density indicus (HDi and LDi, respectively)]. Estimates of Pearson correlations, regression coefficients, and mean squared errors were used to assess accuracy and bias of predictions. Overall, the BayesC $\pi$  model resulted in less biased predictions. Accuracies ranged from 0.18 to 0.27, which are reasonable values given the heritability estimates (from 0.40 to 0.44) and sample size (568 animals in the training population). Furthermore, results from *Bos taurus indicus* panels were as informative as those from Illumina BovineHD, indicating that they could be used to implement GS at lower costs.

## KEYWORDS

genomic  
selection  
*Bos taurus indicus*  
growth  
feedlot  
performance  
GenPred  
Shared data  
resources

Brazil has the world's second largest cattle herd with over 200 million heads (Instituto Brasileiro de Geografia e Estatística 2013), with the Nelore (*Bos taurus indicus*) being the most widespread and economically important breed. As the total pasture area in Brazil has decreased

over the decades, productivity gains have become an important factor for beef production (Martha *et al.* 2012). The Nelore breed has been selected for growth rate traits on pasture based on traditional pedigree and phenotype analysis; however, the Brazilian beef feedlot industry has grown about 50% in the last decade (Millen *et al.* 2011), and novel breeding objectives and criteria are required.

In this context, the application of technologies to improve animal performance and thus to supply genetically improved animals for both pasture and feedlot systems are a critical factor to overcome the challenge of increasing Brazilian beef production efficiency. Nowadays, exploring the availability of technology to genotype thousands of single nucleotide polymorphisms (SNPs) distributed across the genome allows the application of GS. Phenotypic and SNP data information are then combined to predict GEBV earlier in the life of the animals (Meuwissen *et al.* 2001). It has been argued that GS could lead to a decrease in generation interval, an improvement of the rate of genetic

Copyright © 2017 Somavilla *et al.*

doi: <https://doi.org/10.1534/g3.117.041442>

Manuscript received February 4, 2017; accepted for publication April 5, 2017; published Early Online April 7, 2017.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material is available online at [www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.041442/-/DC1](http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.041442/-/DC1).

<sup>1</sup>Corresponding author: Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Via de Acesso Prof. Paulo Donato Castellane, Jaboticabal 14884-900, Brazil. E-mail: [danisio@fcav.unesp.br](mailto:danisio@fcav.unesp.br)

gain (Schaeffer 2006), and also assist the better control of inbreeding rates (Daetwyler *et al.* 2007).

Based on the importance of the Nelore cattle in Brazil and the increasing use of feedlot systems, it is necessary to identify appropriate methodologies that allow GS of animals with higher growth rates on feedlots. The aim of the current study was to compare different regression models and SNP panels in terms of accuracy, bias, and precision of GEBV for ADG in feedlot-finished Nelore steers.

## MATERIALS AND METHODS

### Samples

During the mating seasons of 2006/07 through 2008/09, 804 steers, offspring of 34 Nelore bulls from 17 lineages, chosen to represent the genealogies of the Nelore breed in Brazil, were generated through fixed-time artificial insemination in five farms. They were raised to 21 months of age and then moved to either the Embrapa Southeast Livestock (São Carlos, SP, Brazil) or the Embrapa National Beef Cattle Center (Campo Grande, MS, Brazil) during three seasons in feedlot experiment periods (2009, 2010, and 2011). Animals were fed with a total mixed ration diet with 13% crude protein and 71% total digestible nutrients (dry matter basis, corn or sorghum, soybean meal, soybean hull, cotton seed, limestone, mineral mixture, urea, and monensin). The diet was provided twice a day in which the feed offered (total mixture composed by concentrate:silage, 40:60 ratio) was adjusted daily *ad libitum*. The animals were weighed every 14 d without fasting, for an average period of 91 d. Steer rearing and sample collection protocols were approved by the Animal Care and Use Committee from the Embrapa Southeast Livestock.

### Phenotype and genotype datasets

The initial dataset consisted of 7236 weight records from the 804 steers, but only those from the 15th up to 77th d in feedlot were considered to estimate ADG, to disregard the first weight, and also because after this period >30% of the animals had already been slaughtered. A linear regression analysis of live weight over time was performed using the remained 3523 records from 803 steers, using the *lm* function of the R software (R Development Core Team 2014). The slope was used as the ADG during the feedlot period for the purpose of considering only the linear weight gain and avoiding comparison with different feedlot period lengths.

Steers were assigned to 39 contemporary groups (CG) containing from 5 to 42 animals, which combined information on mating season (three levels), experimental feedlot (two levels), and slaughter group (32 levels of animals slaughtered in the same week). After that, the phenotype and genotype datasets were merged to ensure that they had the same individuals. The summary of age at feedlot entry, starting weight, ADG, and days in feedlot on the remaining animals are presented in Table 1.

In total, 780 steers and 34 bulls were genotyped with the Illumina BovineHD BeadChip (Illumina, San Diego, CA). The initial dataset contained 742,906 markers, in which unplaced, mitochondrial, and sex-linked SNPs were first discarded, as well as duplicated markers (*e.g.*, two different names and positions for the same SNP). SNPs were also filtered based on two other panels: GeneSeek Genomic Profiler (GGP) HDi 80K and GGP LDi 20K (Gene Seek Inc., Lincoln, NE). The panels were built specifically for *B. taurus indicus* breeds. Originally, the GGP HDi 80k/LDi 20k contained 74,085/19,721 markers, of which 69,942/18,464 were available in the primary dataset.

**Table 1 Summary of age and weight at feedlot entry, ADG, and days in feedlot for the 718 Nelore steers**

	Age (d)	Weight (kg)	ADG (kg/d)	Days in Feedlot
Minimum	542	226	0.193	48
Mean ( $\pm$ SD)	649 (45)	361 (51)	1.235 (0.407)	92 (20)
Maximum	745	510	2.457	119

ADG, average daily weight gain.

Paternity correction and quality control (QC) were performed to improve results. To deal with SNPs presenting significant deviation from the Hardy-Weinberg Proportions (HWP) deviation, we checked plots of HWP vs. percentage of heterozygous, and 17 SNPs with >80% of heterozygous were excluded from the three datasets because they could reflect an error during the genotyping procedure (Ziegler 2009). QC was performed using the R package SNPtats (Clayton 2015). SNPs were kept for further analysis only if they had call rate >98% and minor allele frequency (MAF) >1%. The MAF filter excluded 20.0, 1.9, and 7.3% of the total SNPs from the 770k, HDi, and LDi panels, respectively.

After QC, the Beagle v.3.3.2 (Browning and Browning 2009) software was used for phase inference and imputation of missing genotypes for each SNP panel. Finally, to constitute a fourth SNP panel scenario, Tagger (Bakker *et al.* 2005), which is based on linkage disequilibrium (LD) between markers ( $r^2$ ), was used. This tool estimates the  $r^2$  between all SNP pairs and then selects a minimal set (TagSNPs) of markers with a  $r^2 \geq 0.3$  with at least one another marker on the same chromosome. We have chosen this threshold because it is the overall average  $r^2$  at the distance of 10–25 kb, obtained in a previous analysis of the same animals (Mudadu *et al.* 2016). The final number of SNP was 15,863, 63,945, 82,933, and 534,787 for the LDi, HDi, TagSNP, and 770k panels, respectively.

### Fixed effects modeling and adjusted phenotypes

The adjusted phenotype ( $\hat{y}$ ) was represented as  $\hat{y} = y - 1\hat{\mu} - W\hat{\alpha}$ , in which  $y$  is the vector of observations,  $\hat{\mu}$  is the overall mean,  $W$  is an incidence matrix for fixed effects (CG and animal age at feedlot entry), and  $\hat{\alpha}$  is the vector of fixed effects estimates. A residual analysis was performed at this point and animals with the normalized residuals with absolute values >3.5 were removed, thus 718 steers remained in the dataset.

### Models for genomic-enabled prediction

Three specifications were considered for building genome-enabled prediction models: BayesA, BayesC $\pi$ , and BGLUP. The R package BGLR (de los Campos and Rodriguez 2016) was used to fit the models, a flat (noninformative) prior was assigned to the intercept. For the BayesA method, a normal distribution was assigned to the marker effects,  $\beta_j \sim N(0, \sigma_{\beta_j}^2)$ , where  $j = (1, \dots, p)$ ,  $p$  is the number of SNPs, and  $\sigma_{\beta_j}^2$  is the individual variance for the SNP effect. In a second level of hierarchy, each  $\sigma_{\beta_j}^2$  was assigned independent and identically distributed (iid) Scaled-inverse  $\chi^2$  density, with degrees of freedom ( $df_\beta$ ) set to 5 and scale parameter ( $S_\beta$ ) treated as unknown, following a  $\gamma$  distribution with shape ( $s$ ) and rate ( $r$ ) parameters. The parameter  $s$  was set to  $s = 1.1$  and  $r$  was solved so that 80% of proportion of the variance of the response was attributed the linear predictor. On this model, the prior marginal distribution of marker effects is a scaled-t density, with parameters  $df_\beta$  and  $S_\beta$  (Rosa *et al.* 2003).

For the BayesC $\pi$  model, the prior for each marker effect was an iid mixture of point of mass ( $1 - \pi$ ) at zero (spike) and a slab that

■ **Table 2** Parameters of Gibbs sampler for each model

MCMC Samples	Model		
	BayesA	BayesC $\pi$	BGBLUP
Total	400,000	600,000	160,000
Burn-in	150,000	200,000	60,000
Thinning	10	20	10
Posterior <sup>a</sup>	25,000	20,000	10,000

MCMC, Markov chain Monte Carlo; BGBLUP, Bayesian genomic best linear unbiased prediction.

<sup>a</sup>Final number of samples used to calculate features of posterior distributions.

follows a Gaussian distribution,  $\beta_j \sim N(0, \sigma_\beta^2)\pi$ , where  $\sigma_\beta^2$  is the common variance for the SNP effects. The additional parameter  $\pi$  represents the prior proportion of nonzero effects and was treated as an unknown, with a  $\beta$  prior distribution  $\pi \sim \beta(p_0, \pi_0)$ , with  $p_0 > 0$  and  $\pi_0 \in [0, 1]$ . The parameters were set to  $p_0 = 2$  and  $\pi_0 = 0.5$ , which gave a uniform prior in the interval  $[0, 1]$ . Thus, differently from BayesA, BayesC $\pi$  sets some SNP effects to zero, within a variable selection framework.

The BGBLUP model was implemented as a Bayesian Reproducing Kernel Hilbert Spaces regression (de los Campos *et al.* 2009), using a single kernel, user-defined (co)variance matrix  $K$ . The vectors of additive random effects were assigned multivariate normal priors,  $u \sim N(0, K\sigma_u^2)$ , in which  $\sigma_u^2 \sim \chi^{-2}(S, df)$  and  $K$  was set as a marker-derived relationship matrix  $G$ , built as the first method proposed by VanRaden (2008). Briefly, let  $M_{n \times m}$  be a genotype matrix with  $n$  (number of samples) rows and  $m$  (number of SNPs) columns,  $Z_{n \times m}$  be the centered M matrix, and  $G = ZZ' / [2 \sum p_j(1 - p_j)]$ , where the denominator is the total variance across loci. The degrees of freedom (df) was set to 5 and the scale parameter ( $S$ ) was solved so that 80% of proportion of the variance of the response was attributed the linear predictor.

The number of iterations, burn-in, and thinning interval parameters were graphically evaluated and were different for each model (Table 2), and the length of the chain used to compute posterior statistics was 25,000, 20,000, and 10,000 for BayesA, BayesC $\pi$ , and BGBLUP, respectively. For BayesA and BayesC $\pi$ , the marker-based genetic variance ( $\sigma_g^2$ ) was computed as the sum of the variance explained by each SNP marker ( $\sigma_{\beta_j}^2$ ), while for BGBLUP the genetic variance was equal to  $\sigma_u^2$ . For the three models, the narrow sense heritability was estimated as:  $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2)$ , where  $\sigma_e^2$  is the residual variance.

### Validation

The dataset was divided into training (animals from seasons 1 and 2) and testing (animals from season 3) subgroups, which contained 568 and 150 animals, respectively. For the BayesA and BayesC $\pi$  models, the GEBV on the testing set was defined as  $GEBV_{i(tst)} = \sum_{j=1}^p g_{ij} \hat{\beta}_{tm}$ , where  $g_{ij}$  is the genotype of the  $j$ th SNP on the  $i$ th animal and  $\hat{\beta}_{tm}$  is the vector of the SNP marker effect estimated on the training set. For BGBLUP, phenotypes of the testing subgroup were set as missing and samples of  $u$  were obtained in each iteration from the posterior distribution  $[u, \sigma_u^2, \sigma_e^2 | \hat{y}]$ .

The correlation between GEBV and the adjusted phenotype of animals on testing subgroup,  $r(GEBV_{i(tst)}, \hat{y}_{i(tst)})$ , was used as an estimation of prediction accuracy. The slopes of regressing adjusted phenotypes on GEBV for animals in the testing subgroup ( $b_{\hat{y}_{tst}, GEBV_{tst}}$ ) were evaluated as a measure of bias, which can be used to verify whether genomic predictions are inflated or deflated. The last comparison criterion was the mean square error,

■ **Table 3** Pearson correlation coefficients used as proxy estimates of prediction accuracies of GEBV for ADG of the 150 animals in the testing subgroup

Model	SNP Panel <sup>a</sup>			
	770k	TagSNP	HDi	LDi
BGBLUP	0.26	0.24	0.25	0.26
BayesA	0.26	0.25	0.26	0.27
BayesC $\pi$	0.26	0.25	0.25	0.26

SNP, single nucleotide polymorphism; HDi, high-density indicus; LDi, low-density indicus; BGBLUP, Bayesian genomic best linear unbiased prediction.

<sup>a</sup>Actual number of SNPs included in the analysis: 770k, 534,787; TagSNP, 82,933; HDi, 63,945; and LDi, 15,863.

$MSE = \sum_{i=1}^{n_{tst}} (GEBV_i - \hat{y}_i)^2 / (n_{tst})$ , where  $n_{tst}$  is the size of testing dataset, that was used as a measure of precision and bias of the point estimator.

### Data availability

The phenotypic and genotypic data are available at the figshare repository and their description and accession numbers are listed in Supplemental Material, File S1. File S2 contains a custom R script used in the analysis.

## RESULTS AND DISCUSSION

### Accuracy of GEBV

Pearson correlation coefficients between adjusted phenotypes and GEBV were used as a proxy of genome-enabled prediction accuracies (Table 3). All estimates were quite similar, ranging from 0.24 to 0.27. Bolormaa *et al.* (2013) reported even lower accuracies (from 0.13 to 0.24) of GEBV for ADG in feedlot using GBLUP estimates in *B. taurus taurus* and *B. taurus indicus* animals. When analyzing ADG of almost 4000 Nelore young bulls in pasture using traditional BLUP, Fragomeni *et al.* (2013) reported an EBV accuracy of 0.56, which suggests that we could achieve higher accuracies than we found in the present study.

It is known that the success of GS depends on the accuracy of GEBV, which in turn is a function of heritability, size of training population, and effective population size ( $N_e$ ) (Goddard and Hayes 2009). Based on the simulation presented by van der Werf (2013), who considered a population with  $N_e = 250$  (estimated  $N_e$  of Nelore cattle = 214 (Mudadu *et al.* 2016)) and a trait with  $h^2 = 0.5$ , a training population of 500 animals would reach an accuracy of 0.2, similar to our results. Moreover, the authors showed that a training population of >2000 individuals would be required to achieve an accuracy of 0.4. Another key factor is the level of relationship among animals in the training and testing sets. The present study evaluated half-sib families and, according to Hayes *et al.* (2009), this structure only allows estimation of the effects of paternal alleles with high accuracy, decreasing the reliability of the GEBVs.

Taking into account the above-mentioned factors, we point out that the crucial points would be to increase the number of reference animals and to include animals with different levels of relationship, thus the SNP marker effects could be better estimated. Since ADG in feedlot-finished steers could be viewed as a new selection criterion for Nelore cattle, it is important to estimate the GEBVs with high accuracy in order to allow the selection of young animals and genetic gains at reduced genotyping costs.

### Bias and precision measures of GEBVs

Regression coefficients of adjusted phenotypes on GEBV (Table 4) were used to measure the extent of prediction bias, since values

■ **Table 4 Regression coefficients (b) of GEBV on adjusted phenotype and MSE of predictions for the 150 animals in testing subgroup**

Model	SNP Panel <sup>a</sup>							
	770k		TagSNP		HDi		LDi	
	b	MSE	b	MSE	b	MSE	b	MSE
BGBLUP	1.15	1.58	0.46	1.59	1.10	1.58	1.11	1.59
BayesA	1.29	1.09	0.69	1.24	1.68	1.32	1.99	1.37
BayesC $\pi$	0.98	1.12	0.45	1.12	0.94	0.94	0.93	0.94

SNP, single nucleotide polymorphism; HDi, high-density indicus; LDi, low-density indicus; b, regression coefficient; MSE, mean squared errors; BGBLUP, Bayesian genomic best linear unbiased prediction.

<sup>a</sup>Actual number of SNPs included in the analysis: 770k, 534,787; TagSNP, 82,933; HDi, 63,945; and LDi, 15,863.

greater or lower than one are related to deflated or inflated GEBV, respectively. For the 770k panel, only the results from BayesC $\pi$  models were not considered to be biased. Also, it is clear that estimates from BayesA models (except for TagSNP) were deflated, which means that the GEBVs were not in the same scale as the adjusted phenotypes. The opposite was observed for all models applied to the TagSNP dataset, thus it seems that selecting markers based only on their pairwise  $r^2$  resulted in overestimated predictors.

Differences among prediction accuracies were negligible, thus information on slopes and mean squared errors (MSE) (Table 4) were combined and the models resulting in less biased GEBV were 770k-BayesC $\pi$ , HDi-BayesC $\pi$ , and LDi-BayesC $\pi$ . The current average cost of genotyping can easily reach \$150.00, \$100.00, and \$50.00 per animal for 770k, HDi, and LDi, respectively. Therefore, if it was possible to predict accurate GEBV using less dense panels of SNPs at lower cost, the implementation and application of GS would be better accepted by the beef cattle industry.

### Estimates of variance components

The divergences in the variance components (Table 5) were expected, since the markers included in each model capture different proportions

of the genetic variance. For example, the marker-based genetic variance estimated using BGBLUP was the lowest (about 0.02) in this study. For BayesA and BayesC $\pi$ , the genetic variance is a function of SNP effects and their uncertainty variances and allelic frequencies (Gianola *et al.* 2009). Results from BayesA models were not consistent among SNP panels and we hypothesized that, by fitting a larger number of markers, the captured marker-based genetic variance is greater (Table 5).

BayesC $\pi$  models resulted in less biased GEBVs, and its coefficients of heritability ranged from 0.41 to 0.44 (Table 5). This was similar to the coefficient reported by Olivieri *et al.* (2016) for ADG in Nelore cattle in a postweaning feedlot performance test ( $h^2 = 0.43$ ). Although heritability is a population parameter, it is known that magnitudes of heritability estimates of similar traits are often similar across populations.

### Conclusions

For the purpose of comparing GEBV estimates using different SNP panels and Bayesian models, we considered some of the most common criteria used to evaluate the quality of the genome-enabled predictions. Overall, all SNP panels and models provided similar accuracies; however, *B. taurus indicus* SNP chips (HDi and LDi) and methods that zero a proportion of the SNP effects, such as BayesC $\pi$ , seem to result in less biased predictions. Furthermore, results from less dense marker panels based on *B. taurus indicus* were as good as the high-density panel, and at lower genotyping costs.

### ACKNOWLEDGMENTS

We thank Dante Pazzanese Lanna, Michele Nascimento, Amália Chaves, Andrea Souza, Rymer Tullio, Sérgio Medeiros, and Antônio Rosa for data collection, help, and technical assistance. We acknowledge the collaborative efforts among the Brazilian Agricultural Research Corporation (EMBRAPA), São Paulo State University, the São Paulo State Research Foundation (FAPESP), and the University of Wisconsin–Madison. This study was funded by EMBRAPA and FAPESP scholarships no. 2012/23702-8 and 2013/21644-3, and grant no. 2012/23638-8.

■ **Table 5 Estimates of residual ( $\sigma_e^2$ ) and genetic ( $\sigma_g^2$ ) variance components, heritability ( $h^2$ ), and proportion of nonzero effects ( $\pi$ ) for all models**

SNP Panel <sup>a</sup>	Parameter	BGBLUP <sup>b</sup>	BayesA <sup>b,c</sup>	BayesC $\pi$ <sup>b,c</sup>
770k	$\sigma_e^2$	0.05 (0.04–0.06)	0.06 (0.05–0.07)	0.05 (0.04–0.06)
	$\sigma_g^2$	0.02 (0.01–0.04)	0.06	0.03
	$h^2$	0.31 (0.19–0.45)	0.53 (0.49–0.58)	0.41 (0.36–0.47)
	$\pi$	–	–	0.98 (0.96–1.00)
TagSNP	$\sigma_e^2$	0.05 (0.04–0.06)	0.06 (0.05–0.07)	0.05 (0.04–0.06)
	$\sigma_g^2$	0.02 (0.01–0.04)	0.04	0.03
	$h^2$	0.32 (0.19–0.46)	0.40 (0.36–0.45)	0.42 (0.37–0.48)
	$\pi$	–	–	0.98 (0.96–1.00)
HDi	$\sigma_e^2$	0.05 (0.04–0.06)	0.06 (0.05–0.07)	0.05 (0.04–0.06)
	$\sigma_g^2$	0.02 (0.01–0.04)	0.03	0.03
	$h^2$	0.32 (0.19–0.46)	0.31 (0.28–0.35)	0.42 (0.37–0.48)
	$\pi$	–	–	0.98 (0.96–1.00)
LDi	$\sigma_e^2$	0.05 (0.04–0.06)	0.06 (0.05–0.07)	0.05 (0.03–0.06)
	$\sigma_g^2$	0.02 (0.01–0.04)	0.02	0.04
	$h^2$	0.32 (0.19–0.45)	0.28 (0.25–0.32)	0.44 (0.36–0.47)
	$\pi$	–	–	0.98 (0.96–1.00)

SNP, single nucleotide polymorphism; HDi, high-density indicus; LDi, low-density indicus; BGBLUP, Bayesian genomic best linear unbiased prediction; HPD, highest posterior density intervals.

<sup>a</sup>Actual number of SNPs included in the analysis: 770k, 534,787; TagSNP, 82,933; HDi, 63,945; and LDi, 15,863.

<sup>b</sup>Numbers in brackets refers to the HPD at 95% (lower bound–upper bound).

<sup>c</sup>HPD for  $\sigma_g^2$  for models BayesA and BayesC $\pi$  could not be estimated.

## LITERATURE CITED

- Bakker, P. I. W., R. Yelensky, I. Pe'er, S. B. Gabriel, M. J. Daly *et al.*, 2005 Efficiency and power in genetic association studies. *Nat. Genet.* 37: 1217–1223.
- Bolormaa, S., J. E. Pryce, K. Kemper, K. Savin, B. J. Hayes *et al.*, 2013 Accuracy of prediction of genomic breeding values for residual feed intake and carcass and meat quality traits in *Bos taurus*, *Bos indicus*, and composite beef cattle. *J. Anim. Sci.* 91: 3088–3104.
- Browning, B. L., and S. R. Browning, 2009 A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* 84: 210–223.
- Clayton, D., 2015 snpStats: SnpMatrix and XSNpMatrix classes and methods. R package version 1.24.0. Available at: <http://bioconductor.org/packages/snpStats>. Accessed February 4, 2017.
- Daetwyler, H. D., B. Villanueva, P. Bijma, and J. A. Woolliams, 2007 Inbreeding in genome-wide selection: inbreeding in genome-wide selection. *J. Anim. Breed. Genet.* 124: 369–376.
- de los Campos, G., and P. P. Rodriguez, 2016 BGLR: Bayesian generalized linear regression. R package version 1.0.5. Available at: <https://CRAN.R-project.org/package=BGLR>. Accessed February 4, 2017.
- de los Campos, G., D. Gianola, and G. J. M. Rosa, 2009 Reproducing kernel Hilbert spaces regression: a general framework for genetic evaluation. *J. Anim. Sci.* 87: 1883–1887.
- Fragomeni, B. de O., D. C. B. Scalez, F. L. B. Toral, J. A. G. Bergmann, I. G. Pereira *et al.*, 2013 Genetic parameters and alternatives for evaluation and ranking of Nelore young bulls in pasture performance tests. *Rev. Bras. Zootec.* 42: 559–564.
- Gianola, D., G. de los Campos, W. G. Hill, E. Manfredi, and R. Fernando, 2009 Additive genetic variability and the Bayesian alphabet. *Genetics* 183: 347–363.
- Goddard, M. E., and B. J. Hayes, 2009 Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat. Rev. Genet.* 10: 381–391.
- Hayes, B. J., P. M. Visscher, and M. E. Goddard, 2009 Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* 91: 47.
- Instituto Brasileiro de Geografia e Estatística, 2013 *Anuário Estatístico do Brasil*. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.
- Martha, G. B., E. Alves, and E. Contini, 2012 Land-saving approaches and beef production growth in Brazil. *Agric. Syst.* 110: 173–177.
- 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.
- Millen, D. D., R. D. L. Pacheco, P. M. Meyer, P. H. M. Rodrigues, and M. De Beni Arrigoni, 2011 Current outlook and future perspectives of beef production in Brazil. *Anim. Front.* 1: 46–52.
- Mudadu, M. A., L. R. Porto-Neto, F. B. Mokry, P. C. Tizioto, P. S. N. Oliveira *et al.*, 2016 Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. *BMC Genomics* 17: 235 (erratum: *BMC Genomics* 17: 492).
- Olivieri, B. F., M. E. Mercadante, J. N. Cyrillo, R. H. Branco, S. F. Bonilha *et al.*, 2016 Genomic regions associated with feed efficiency indicator traits in an experimental Nelore cattle population. *PLoS One* 11: e0164390 (erratum: *PLoS One* 17: 492).
- R Development Core Team, 2014 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rosa, G. J. M., C. R. Padovani, and D. Gianola, 2003 Robust linear mixed models with normal/independent distributions and Bayesian MCMC implementation. *Biom. J.* 45: 573–590.
- Schaeffer, L. R., 2006 Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123: 218–223.
- van der Werf, J., 2013 Genomic selection in animal breeding programs, pp. 543–561 in *Genome-Wide Association Studies and Genomic Prediction*, edited by C. Gondro, C. J. van der Werf, and B. Hayes. Humana Press, Totowa, NJ.
- VanRaden, P. M., 2008 Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91: 4414–4423.
- Ziegler, A., 2009 Genome-wide association studies: quality control and population-based measures. *Genet. Epidemiol.* 33: S45–S50.

Communicating editor: D. J. de Koning