Animal 15 (2021) 100085

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

Genomic prediction ability for feed efficiency traits using different models and pseudo-phenotypes under several validation strategies in Nelore cattle

L.C. Brunes [a,b,*], F. Baldi [c], F.B. Lopes [d], M.G. Narciso [b], R.B. Lobo [e], R. Espigolan [f], M.F.O. Costa [b], C.U. Magnabosco [g]

[a] Animal Science Department, Goiás Federal University, 74690-900 Goiânia, GO, Brazil
[b] Embrapa Rice and Beans, GO-462, km 12, 75375-000 Santo Antônio de Goiás, GO, Brazil
[c] Animal Science Department, São Paulo State University – Júlio de Mesquita Filho (UNESP), Prof. Paulo Donato Castelane, 14884-900 Jaboticabal, SP, Brazil
[d] Cobb-Vantress, Inc., 72761 Siloam Springs, AR, USA
[e] National Association of Breeders and Researchers, 14020-230 Ribeirão Preto, Brazil
[f] Department of Veterinary Medicine, Faculty of Animal Science and Food Engineering, University of Sao Paulo, 13635-900 Pirassununga, SP, Brazil
[g] Embrapa Cerrados, BR-020, 18 Sobradinho, 70770-901 Brasilia, DF, Brazil

ARTICLE INFO

Article history:
Received 27 March 2020
Received in revised form 9 September 2020
Accepted 15 September 2020
Available online 24 December 2020

Keywords:
Genomic breeding value
Residual body weight gain
Residual feed intake
Single nucleotide polymorphisms
Zebu

ABSTRACT

There is a growing interest to improve feed efficiency (**FE)** traits in cattle. The genomic selection was proposed to improve these traits since they are difficult and expensive to measure. Up to date, there are scarce studies about the implementation of genomic selection for FE traits in indicine cattle under different scenarios of pseudo-phenotypes, models, and validation strategies on a commercial large scale. Thus, the aim was to evaluate the feasibility of genomic selection implementation for FE traits in Nelore cattle applying different models and pseudo-phenotypes under validation strategies. Phenotypic and genotypic information from 4 329 and 3 467 animals were used, respectively, which were tested for residual feed intake, DM intake, feed efficiency, feed conversion ratio, residual BW gain, and residual intake and BW gain. Six prediction methods were used: single-step genomic best linear unbiased prediction, Bayes A, Bayes B, Bayes Cπ, Bayesian least absolute shrinkage and selection operator (**BLASSO**), and Bayes R. Phenotypes adjusted for fixed effects (Y*), estimated breeding value (**EBV**), and EBV deregressed (**DEBV**) were used as pseudo-phenotypes. The validation approaches used were: (1) random: the data was randomly divided into ten subsets and the validation was done in each subset at a time; (2) age: the partition into training and testing sets was based on year of birth and testing animals were born after 2016; and (3) EBV accuracy: the data was split into two groups, being animals with accuracy above 0.45 the training set; and below 0.45 the validation set. In the analyses that used the Y* as pseudo-phenotype, prediction ability (**PA**) was obtained by dividing the correlation between pseudo-phenotype and genomic EBV (**GEBV**) by the square root of the heritability of the trait. When EBV and DEBV were used as the pseudo-phenotype, the simple correlation of this quantity with the GEBV was considered as PA. The prediction methods show similar results for PA and bias. The random cross-validation presented higher PA (0.17) than EBV accuracy (0.14) and age (0.13). The PA was higher for Y* than for EBV and DEBV (30.0 and 34.3%, respectively). Random validation presented the highest PA, being indicated for use in populations composed mainly of young animals and traits with few generations of data recording. For high heritability traits, the validation can be done by age, enabling the prediction of the next-generation genetic merit. These results would support breeders to identify genomic approaches that are more viable for genomic prediction for FE-related traits.

Implications

Genetic selection for feed efficiency allows for reduced production costs and the negative effects on the environment. The prediction ability depends on pseudo-phenotypes and validation approaches used,

* Corresponding author at: Animal Science Department, Goiás Federal University, 74690-900 Goiânia, GO, Brazil.
E-mail address: ludmillabrunes@hotmail.com (L.C. Brunes).

https://doi.org/10.1016/j.animal.2020.100085

mainly, for traits that are poorly evaluated, as feed efficiency. The genetic parameters and genomic models evaluated provide support for the adoption of these traits on a large evaluation scale. The random cross-validation and phenotype adjusted for fixed effects displayed superior prediction ability. Despite the prediction models show similar prediction ability, the single-step genomic best linear unbiased prediction simplify genomic evaluations, since this model use phenotypes instead of pseudo-phenotypes and accounting with the whole population to estimate genomic value.

Introduction

Animal feedstuff plays a significant role in the production costs of beef cattle systems (Savietto et al., 2014). Thus, there is a big pressure to improve feed efficiency (**FE)-**related traits, due to the influence on production costs and reduce the environmental impacts of livestock. Besides that, competition from beef cattle production with other agribusiness activities has forced breeders and industries to seek new strategies to maximize the efficiency and profitability of meat systems (Olivieri et al., 2016).

In recent years, many efforts were done to improve the ratio of average daily gain (**ADG**) and DM intake (**DMI**). Novel traits like residual feed intake (**RFI**), residual BW gain (**RG**) (Koch et al., 1963), and residual intake and BW gain (**RIG**) (Berry and Crowley, 2012) can be used to identify animals with lower feed intake or higher weight gain (Berry and Crowley, 2012). These traits show the advantage of being phenotypically not associated with growth and body size (Koch et al., 1963; Berry and Crowley, 2012).

The genomic selection was proposed to improve the FE-related traits, since these traits are difficult and expensive to measure, limiting their full-scale use in beef cattle breeding programs (Pryce et al., 2012; Silva et al., 2016). Several methods are available for genomic prediction, which differ statistically and also for use under commercial conditions (Aguilar et al., 2010; Habier et al., 2011; Erbe et al., 2012; Daetwyler et al., 2013; De Los Campos et al., 2013; Chiaia et al., 2018). For Nelore cattle, only traits as ADG, DMI, feed conversion ratio (**FCR**), and RFI, and models as genomic best linear unbiased prediction, single-step genomic best linear unbiased prediction, and Bayes Cπ were evaluated in an experimental herd (Silva et al., 2016).

To the best of our knowledge, there are no reports of pseudo-phenotypes evaluation in the genomic prediction of FE traits in Nelore cattle (Pryce et al., 2012; Bolormaa et al., 2013; Silva et al., 2016). As pseudo-phenotype or response variable, the true breeding value would be the ideal parameter. However, this value is unknown (De Los Campos et al., 2013) and other values were evaluated for this purpose (Fernandes Júnior et al., 2016; Chiaia et al., 2018), to identify within each scenario, which is ideal to be used. Proper estimation of marker effects requires an adequate pseudo-phenotype to summarize the genetic information on training animals, as well as statistical methods that efficiently associate pseudo-phenotypes to marker information (Ostersen et al., 2011). The pseudo-phenotype prediction ability (**PA**) depends on the structure of the data and pedigree, the estimated breeding value (**EBV**) accuracy, the availability of phenotypes, genotypes, and pedigree, that is, the quantity and quality of the information available (Garrick et al., 2009; Daetwyler et al., 2013; Boddhireddy et al., 2014; Neves et al., 2014).

Under commercial situations, it is important to enable the prediction of the next-generation genetic merit, using genomic information. Thus, older animals or with more reliable EBV can be used as a reference population to define the prediction equations that will be validated in younger animals or with less accurate EBV, that is indicated when a structured data set with older animals, phenotypic, pedigree, and genomic information is available (VanRaden et al., 2009; Habier et al., 2011; Silva et al., 2016). However, in beef cattle, records from FE-related traits are normally available in young unproven animals, since these traits have recently been included as a selection criterion, and proven sires only have genomic information.

Thus, the data set for FE traits have commonly few generations, few animals with progeny record, and older animals have low accuracy genetic evaluations. Under these conditions, cross-validation may be feasible (Pérez-Cabal et al., 2012; Pryce et al., 2012; Silva et al., 2016), but this approach brings as a disadvantage the random formation of training and validation populations, which may result in the use of few accuracy animals for training the prediction equation and in the bias of the estimates (Runcie and Cheng, 2019).

It is important to identify the most appropriate validation approach and pseudo-phenotype for application in different data set structure and population evaluated for FE-related traits, allowing to increase the PA of genomic EBV (**GEBVs**).

Most of the previous genomic studies for FE performed with zebu breeds were carried out with experimental or small populations with little environmental variation and promising traits as RG and RIG were not evaluated for this breed (Silva et al., 2016). Thus, there is a need to carry out studies with a larger sample size to evaluate the implementation feasibility of genomic selection for FE-related traits under different commercial scenarios and data, since FE-related traits are a group of denominated novel traits with scarce phenotyping in bull test stations and complex data structure due to unbalance phenotyping and genotyping.

The aim was to evaluate the feasibility of genomic selection implementation for FE-related traits in Nelore cattle applying different models and pseudo-phenotypes under several validation strategies using records from commercial herds. Preliminary results of this study were published as an abstract (Magnabosco et al., 2020), and the same data set was shared with Brunes et al. (2020).

Material and methods

Data from the FE tests carried out between 2011 and 2018, phenotypic and genotypic information of 4 329 and 3 594 animals, respectively, were considered, and more details about the data set used were presented by Brunes et al. (2020). Animals belonged to 39 farms located in the Midwest, Southeast, Northeast, and North Brazilian regions. The Nelore Brazil Breeding Program, coordinated by the National Association of Breeders and Researchers, performs the genetic evaluation of these herds. The relationship matrix was calculated based on pedigree information from 58 374 animals, provided by the National Association of Breeders and Researchers. The data and pedigree have a consistent connection through common sires in the FE tests and due to achievement of progeny testing for more than 30 years among herds participating in the breeding program (Lobo et al., 2019).

A total of 125 FE tests were performed and animals were evaluated with an average age of 13.5 $\pm$ 3.92 months at the beginning of the tests under similar management and environmental conditions in the feedlot. Tests were conducted using the same protocol (Mendes et al., 2020). The FE tests were performed in five different places, three ranches (HoRa Hofig Ramos, Rancho da Matinha, and AgroNova) and two research centers (Embrapa Cerrados in Goias and Federal University of Uberlandia). Diets offered over the years differed in composition and ingredients but were formulated based on silage and commercial concentrate, with an average of 64% total digestible nutrients, 13% CP, 76% DM, and formulated for gains of 1.2 kg/day (Mendes et al., 2020). During the tests, the animals' weight was obtained by periodic weighing, and at the beginning and end of the evaluation period. Roughage, concentrate and wastes samples were collected every week to evaluate chemical composition.

The DNA samples were obtained from hair follicles taken from animals' tails and placed in card with adhesive film. The Nelore cattle were genotyped for single nucleotide polymorphism markers using low-density panel (CLARIFIDE® Nelore 3.1), containing 29 004

molecular markers. DNA extraction and sample genotyping were performed by Zoetis® (Kalamazoo, MI), through its protocol.

Performance traits

The DMI was measured by collective stalls equipped with automated systems (GrowSafeSystem® and Intergado®) for a minimum of 70 days preceded by adaptation. The DMI, measured in kg/day, was obtained by calculating the average of all valid daily intake values during the test period. This parameter was calculated as the amount of individually consumed feed automatically recorded by the electronic systems. The FE, measured in kg ADG/kg DMI, was obtained as the ratio between ADG and DMI. The FCR, measured in kg DMI/kg ADG, was obtained by the inverse ratio (DMI/ADG). The ADG (kg/day) was estimated by the linear regression coefficient of the weights as a function of the days in test, using the *lm* function of R Program - R Core Team, 2018 and the following equation:

$$y_{ij} = \alpha + \beta * DIT_j + \varepsilon_{ij}$$

where y_{ij} is the *j*th observation of weight of *i*th animal; α is the intercept of the regression equation which represents the initial weight; β is the linear regression coefficient which represents the ADG; DIT_j is the day in the performance test of *j*th observation; and ε_{ij} is the *j*th residual associated to the *i*th record. It was assumed that the residuals were independent and not correlated.

The metabolic BW was estimated from BW and ADG, by:

$$MW_j^{0.75} = \left[\alpha + \beta * (DIT_j/2)\right]^{0.75}$$

where $MW_j^{0.75}$ is the metabolic BW of *j*th animal; α is the intercept of the regression equation which represents the initial weight; and β is the linear regression coefficient which represents the ADG, as described and obtained above in estimating ADG.

Residual feed intake (kg of DM/day) was estimated, within each contemporary group, by the residual of the DMI regression as a function of ADG and metabolic BW, using the R Program (2018) and the equation below (Koch et al., 1963):

$$y_i = \beta_o + \beta_1 ADG + \beta_2\, MW^{0.75} + \varepsilon_i\ (RFI)$$

where *y* is individual DMI of *i*th animal; β_o is the intercept; β_1 and β_2 are the linear regression coefficient of *ADG* and $MW^{0.75}$, respectively; and ε_i is the residual associated with the *i*th observation, that is, RFI. Regression analysis was performed and no effect of backfat thickness on RFI was observed, thus the RFI was not adjusted for fat thickness.

The RG (Koch et al., 1963; Berry and Crowley, 2012) (kg of ADG/day) was obtained as the difference between the observed ADG and the estimated based on DMI and metabolic BW. The estimated ADG (ADG_{ei}) was obtained using the *lm* function on the R Program - R Core Team, 2018, within contemporary group and by:

$$ADG_{e_i} = \beta_o + \beta_1 DMI + \beta_2\, MW^{0.75} + \varepsilon_i\ (RG)$$

where β_0 is the intercept, β_1 and β_2 are the regression coefficients of *DMI* and $MW^{0.75}$, respectively; and ε_i is the residual associated with the *i*th observation, that is, RG.

Residual intake and BW gain was calculated as RG-RFI, after standardizing both traits to a variance of 1, allowing their combination into a' single value (Berry and Crowley, 2012). Both traits, RFI and RG, are linear functions of their component traits: DMI, ADG, and metabolic BW.

Statistical and quality control analyses

The contemporary group was composed by farm, management group, FE test, sex, year, and birth season (dry season from April to September and the rainy season from October to March). The fixed effects included in the contemporary group were those whose *P*-value was lower than 0.001 obtained from ANOVA results. Records within ± 3.5 SD from the contemporary group mean were considered in the analysis. Additionally, all contemporary group should have at least four animals to proceed with the analysis. Descriptive statistics obtained for the traits evaluated are summarized in Table 1.

Table 1
Number of observations (N), phenotypic mean, SD, number of contemporary group (N° CG), additive genetic variance (σ^2a), residual variance (σ^2e), and heritability ± SE ($h^2 \pm$ SE) for FE-related traits in Nelore cattle[1].

Trait	N	Mean	Median	SD	N° CG	σ^2a	σ^2e	$h^2 \pm$ SE
RFI	4 080	0.00	0.03	0.70	125	0.09	0.042	0.17 ± 0.04
DMI	4 097	7.97	7.86	1.75	126	0.21	0.68	0.23 ± 0.04
FE	2 242	0.09	0.09	0.03	93	0.000030	0.000404	0.07 ± 0.03
FCR	2 235	12.18	11.16	4.43	125	0.80	8.14	0.09 ± 0.03
RG	2 056	0.00	0.02	0.20	93	0.03	0.16	0.17 ± 0.05
RIG	2 033	0.02	−0.01	0.74	93	0.11	0.43	0.20 ± 0.05

[1] This table was previously presented in Brunes et al. (2020). RFI: residual feed intake; DMI: DM intake; FE: feed efficiency; FCR: feed conversion ratio; RG: residual BW gain; RIG: residual intake and BW gain.

In the quality control for genomic data, single nucleotide polymorphisms with minor allele frequency, call rate, and *P*-value for Hardy–Weinberg equilibrium test less than 0.02; 0.95 and 0.15, respectively, were excluded. Only single nucleotide polymorphisms in autosome chromosomes and with known position according to UMD 3.1 bovine genome were considered. Samples with call rates below 0.95 were not considered in the analyses. The quality control criteria were defined based on the data set, averages obtained after analysis, and in previous studies (Olivieri et al., 2016; Silva et al., 2016). This process was performed with R Program - R Core Team, 2018 resulting in a data set with 19 602 single nucleotide polymorphisms and 3 467 animals.

To evaluate the existence population substructure a principal component analysis was performed using information from single nucleotide polymorphisms and genomic relationship matrix of individuals who were approved in the quality control criteria (VanRaden, 2008) (Fig. 1), and more details about the data set was described by Brunes et al. (2020). The proportion of variance explained by the two first principal components was 33.95%. The PC1 e PC2 did not group the animals into clear-cut clusters, implying that genetic admixture probably existed for the evaluated population. The animals' dispersion in the principal component analysis plot indicated the absence of subgroups among the evaluated animals, since there is no formation of major components. Principal component analyses were also performed between

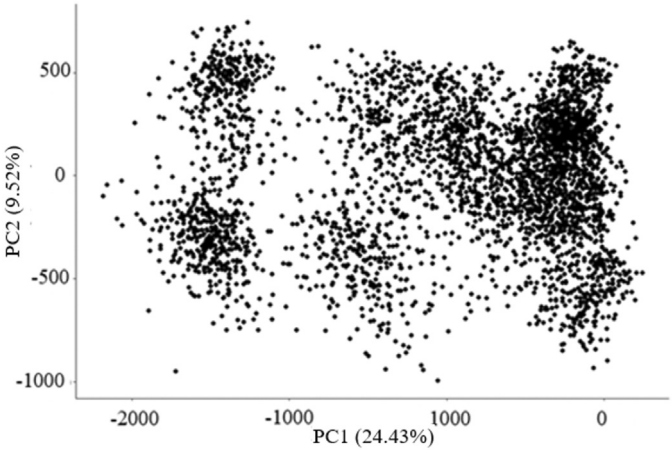

Fig. 1. Principal components (PC) of the genomic relationship among Nelore cattle evaluated for related feed efficiency traits.

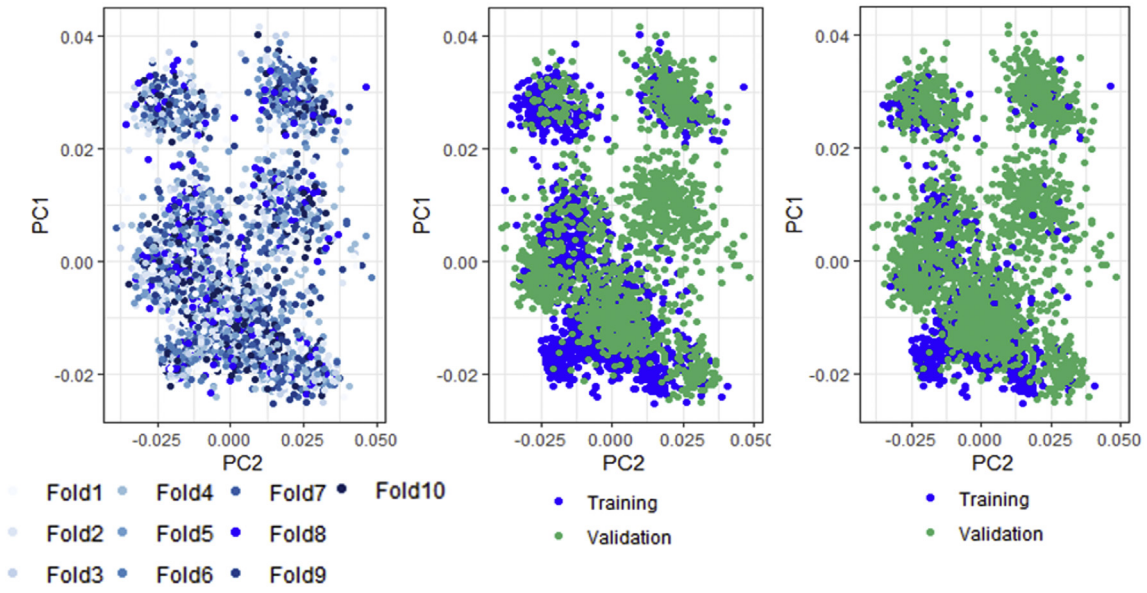

Fig. 2. Distribution of training and validation populations of Nelore cattle for clustering by random (A), age (B), and genetic breeding value (C) accuracy by principal components (PC).

the subsets according to the validation method (Fig. 2). The graphic analyses of the validation and training groups, regardless of the method, demonstrates the absence of population substructure formation and reinforces the interrelationship between them.

The genetic distance between individuals was calculated based on their genotypes using the method Jukes–Cantor (Jukes and Cantor, 1969) and R Program - R Core Team, 2018. Genetic distances values were, on average, 0.08 (0.0 to 0.1816), indicating that the data are not dispersed or sub-grouped in whole population (Table 2), as reported by Brunes et al. (2020). The genetic distances between training and validation populations displayed the same pattern, with low values and no dispersion, pointed out that considering the animals' genetic structure constitute a unique population.

Estimation of variance components

Variance components and genetic parameters were estimated considering a genomic pedigree-based animal model through the single-step genomic best linear unbiased prediction (Aguilar et al., 2010). Single-trait analyses were performed using the restricted maximum likelihood method with REMLF90 and AIREMLF90 (Misztal, 2017). Initially, the REMLF90 (EM-algorithm) was used, and then the obtained estimates were used as initial values for AIREMLF90 (AI-algorithm), a program used to obtain the SE of the variance components and heritability (Meyer and Houle, 2013). Direct additive genetic and residual effects were included as random effects; the contemporary group and animals' age as a covariate were included as a fixed effect. The model used for all traits can be represented as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z_a} + \boldsymbol{e}$$

where y is the vector of the observed trait (RFI, DMI, FE, FCR, RG, and RIG); X is the incidence matrix of fixed effects; β is the vector of fixed effects; Z are the incidence matrix of direct genetic effects; a is the vector of additive genetic random effects; and e is the vector of residual random effects. It was assumed that $E[\mathbf{y}] = \mathbf{X\beta}$; with the direct additive genetic, and residual effects assumed normally distributed with mean zero and $\text{Var}(\boldsymbol{a}) = \mathbf{H} \otimes \mathbf{S}_a$, and $\text{Var}(\mathbf{e}) = \mathbf{I} \otimes \mathbf{S}_e$; in which $\mathbf{S}_a$ and $\mathbf{S}_e$ is the additive genetic and residual covariance matrix, respectively, and $\mathbf{I}$ is an identity matrix of an appropriate order.

The single-step genomic best linear unbiased prediction is a modification of the traditional best linear unbiased prediction, where the inverse of the numerator relationship matrix (A^{-1}) was replaced by H^{-1} (Aguilar et al., 2010), which combines pedigree and genomic information. The inverse of H matrix was obtained as (Aguilar et al., 2010):

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

where H is the relationship coefficients matrix between the animal; A_{22} is an additive relationship matrix for the genotyped animals; and G is the genomic relationship matrix. The G matrix was created as proposed by VanRaden (2008) as follow:

$$G = ZDZ'q$$

where Z is a matrix of gene content adjusted for allele frequencies; D is a weight matrix for single nucleotide polymorphism (initially $D = \text{I}$); and q is a normalizing factor. These factors were obtained ensuring that the average diagonal in G is close to that of A_{22}.

Genomic analyses

Three pseudo-phenotypes were used in the genomic analyses: EBV, deregressed EBV (DEBV) and phenotype adjusted for fixed effects (Y*). The EBVs obtained in the genomic pedigree-based animal model

Table 2
Descriptive analysis of genetic distance in Nelore cattle[1].

Statistics	Mean	Median	Minimum	Maximum	1° quartile	3° quartile
Whole population	0.0865	0.0835	0.0000	0.1816	0.0526	0.1030
Between folds of random validation	0.0912	0.0971	0.0000	0.1710	0.0421	0.0991
Training and validation population for age approach	0.1161	0.1158	0.0002	0.1816	0.0635	0.1174
Training and validation population for estimated breeding value accuracy	0.1095	0.1087	0.0001	0.1796	0.0603	0.1125

[1] This table was previously presented in Brunes et al. (2020).

analyses were deregressed using the method proposed by Garrick et al. (2009), with the support of the DPR package (Lopes, 2017) of R Program - R Core Team, 2018. The Y* was obtained based on the phenotype observed that was adjusted using an animal model considering the fixed effects of the contemporary group and age of evaluation of FE:

$$y_{ij}^{*} = \mu + CG_j + \beta age + e_{ij}$$

where y_{ij}^{*} is the jth observation of phenotype adjusted for fixed effects of ith animal; μ is the average of the data; CG_j is the fixed effect of the jth contemporary group; β is the regression coefficients of age; and e_{ij} is the random effect of the jth residue of the ith observation with $\mu = 0$ and variance $= \sigma_e^2$.

The EBVs and DEBVs accuracy were calculated using the formula below:

$$Acc = \sqrt{\left(\frac{PEV}{(1+F_i)*\sigma_a^2}\right)}$$

where PEV is the variance of prediction error; F_i is the inbreeding coefficient of ith animal; and σ_a^2 is the additive genetic variance of the trait. Table 3 shows the descriptive statistics and accuracy for pseudo-phenotype used in genomic analyses.

The general model used for genomic predictions was:

$$\boldsymbol{y} = \boldsymbol{\mu} + \sum_{\boldsymbol{j=1}}^{\boldsymbol{k}} \boldsymbol{Z_{ij} a_j} + \boldsymbol{e}$$

where y is the pseudo-phenotype vector (Y*, EBV ou DEBV) of ith animal; μ is an overall constant; k is the genetic effect of the jth single nucleotide polymorphism; Z_{ij} is a genotype indicator variable for individual i at locus j; a_j for $j = 1, 2\ldots$; and e is residual effect vector is the residual associated with the observation on individual i. single nucleotide polymorphisms were coded as 0, 1, and 2 for AA, AB e BB, respectively. The residual distribution vector (e) was:

$$e\tilde{N}(0, I\sigma_e^2)$$

where σ_e^2 is residual variance.

For Bayesian analyses, Gibbs chains of 200 000 iterations were generated, with a burn-in period of 150 000 cycles and a sampling interval of 10. The burn-in period and chain convergence were verified by visual inspection of a posteriori means of additive and residual genetic variances, using BOA package of R Program - R Core Team, 2018, with trace plots for visual inspection and Geweke method (Geweke, 1992).

Single-step genomic best linear unbiased prediction

The animal model was applied to the single-step genomic best linear unbiased prediction method (Aguilar et al., 2010). The model used was the same one used to estimate variance components, as described above, by combining A and G from the H matrix.

Bayes A

For Bayes A (Meuwissen et al., 2001), it was assumed that the prior conditional distribution of a marker effect (a_j) is Gaussian, with mean equal to zero and independent variance σ_{aj}^2. The variance associated with the effect of each marker has an inverse χ^2 a prior distribution:

$$p\left(\sigma_{aj}^2\right) = \chi^{-2}\left(\sigma_{aj}^2 | \nu, S^2\right)$$

where ν is degrees of freedom and S^2 is scale parameters. After these specifications, the marginal prior distribution of each marker effect is:

$$p\left(a_j | \nu, S^2\right) = \int N\left(a_j | 0, \sigma_{a_j}^2\right) \chi^{-2}\left(\sigma_{a_j}^2 | \nu, S^2\right) \partial\sigma_{a_j}^2$$

which presents distribution *t-Student* $t(0, \nu_a, S_a^2)$ (Rosa et al., 2003):

$$p\left(a_j | \nu, S^2\right) = t\left(0, \nu, S^2\right)$$

Bayes B

In Bayes B (Meuwissen et al., 2001), it was assumed that most markers have null effect and only a few contribute to the genetic variance (σ_{aj}^2), conditional on the specific variance of each marker, non-zero and Gaussian distribution $N(a_j | 0, \sigma_{aj}^2)$. Therefore, the distribution of marker effects can be described with the mixed model:

Table 3
Number of observations (N), mean, SD, minimum, maximum, and average accuracy ± SD (accuracy ± SD) of adjusted phenotype (Y*), estimated breeding value (EBV), and deregressed estimated breeding value (DEBV) for FE-related traits in Nelore cattle.

Trait	Pseudo-phenotype	N	Mean	SD	Minimum	Maximum	Accuracy ± SD
RFI	Y*[1]	4 103	0.0008	0.69	−3.7449	2.7051	–
	EBV	3 467	−0.0533	0.11	−0.5373	0.3323	0.85 ± 0.08
	DEBV[2]	3 467	−0.0807	0.17	−0.9679	0.6123	0.66 ± 0.03
DMI	Y*	4 106	7.9718	0.91	3.5410	11.2700	–
	EBV	3 467	−0.0828	0.15	−0.9183	0.6649	0.48 ± 0.04
	DEBV	3 467	−0.1583	0.29	−1.8029	1.2328	0.52 ± 0.01
FE	Y*	2 246	0.0916	0.02	0.0223	0.2584	–
	EBV	3 467	0.0007	0.00	−0.0030	0.0044	0.42 ± 0.02
	DEBV	3 467	0.0013	0.00	−0.0059	0.0087	0.51 ± 0.01
FCR	Y*	2 239	12.1776	2.92	0.5220	26.7131	–
	EBV	3 467	−0.1216	0.24	−1.1414	0.8668	0.46 ± 0.05
	DEBV	3 467	−0.4047	0.42	−2.0524	1.5146	0.57 ± 0.01
GR	Y*	2 059	0.0045	0.13	−0.6409	0.5181	–
	EBV	3 467	0.0037	0.02	−0.1130	0.0910	0.57 ± 0.04
	DEBV	3 467	−0.0131	0.06	−0.2662	0.3618	0.50 ± 0.03
RIG	Y*	2 062	0.0189	0.71	−2.6711	3.5023	–
	EBV	3 467	0.1051	0.15	−0.5773	0.9320	0.73 ± 0.14
	DEBV	3 467	0.0963	0.16	−0.6977	1.3512	0.72 ± 0.03

RFI: residual feed intake; DMI: DM intake; FE: feed efficiency; FCR: feed conversion ratio; RG: residual BW gain; RIG: residual intake and BW gain.
[1] The Y* was obtained based on the phenotype observed that was adjusted using an animal model considering the fixed effects of the contemporary group and age of evaluation of feed efficiency.
[2] The DEBV was obtained by the EBV deregression using the method proposed by Garrick et al. (2009).

$$p\left(a_j|\sigma^2_{a_j},\pi\right)=\begin{cases}0 & \text{with probability } \pi\\ N\left(0,\sigma^2_{aj}\right) & \text{with probability } (1-\pi)\end{cases}$$

where π is the proportion of makers with null genetic effects. Similar to Bayes A, this model shows inverse χ^2 prior distribution, thus, marginally after integrating the variances of the markers:

$$p(a_j|\pi)=\begin{cases}0 & \text{with probability } \pi\\ t\left(0,v,S^2\right) & \text{with probability}(1-\pi)\end{cases}$$

Thereby, Bayes B can be reduced to Bayes A assuming that $\pi=0$. S^2, given by:

$$S^2=\frac{\sigma^2_a(\nu-2)}{\nu}$$

where:

$$\sigma^2_a=\frac{\sigma^2_S}{(1-\pi)\sum_{j=1}^{K}2p_j\left(1-p_j\right)}$$

σ^2_a is the additive genetic variance; p_j is the frequency of jth single nucleotide polymorphism; σ^2_s is the additive genetic variance explained by single nucleotide polymorphisms; and π is the probability that jth single nucleotide polymorphism has null effect (Habier et al., 2011).

Bayes Cπ

In Bayes Cπ method, proposed by Habier et al. (2011), the markers effects have a common variance, following a scaled inverse χ^2 prior, with V_a degrees of freedom and S^2_a a scale parameter. As a result, the effect of an single nucleotide polymorphism fitted with probability $(1-\pi)$ and presents a mixture of multivariate student's t-distribution $t(0, V_a, IS^2_a)$, where π is the probability of a marker presenting null effect and has a uniform prior distribution (0,1).

In this method, π is estimated from the analyzed data. Subsequently, the inclusion or exclusion of each marker in the model was performed by an indicator variable δ_j, which is equal to 1 if the marker j is fitted in the model, otherwise, it is zero. The common effect variance is sampled from full-conditional posterior, which presents scaled inverse χ^2 distribution with degrees of freedom $\widetilde{V_a}=V_a+m^t$ and scale:

$$\widetilde{S^2_a}=\left(v_aS^2_a+\sum_{j=1}^{k}a^2_j\right)/\widetilde{v_a}$$

where m^t is the number of markers fitted with non-zero effects in iteration t.

Bayesian least absolute shrinkage and selection operator

In the BLASSO (Park and Casella, 2008), the conditional prior distribution of each marker effect

$$p\left(a_j|\tau^2_j,\sigma^2_e\right)$$

is Gaussian with a zero-mean and marker-specific prior variance, independent from each other. Thus,

$$p\left(a_j|\tau^2_j,\sigma^2_e\right)=\prod_{j=1}^{p}N\left(a_j|0,\tau^2_j\sigma^2_e\right)$$

This prior induces marker-specific shrinkage, whose extent depends on τ_j^{-2}. The variances parameters (τ_j^{-2}) are assigned exponential IIID prior:

$$p(\tau^2_1,\tau^2_2,\ldots,\tau^2_K|\lambda)=\prod_{j=1}^{K}Exp\left(\tau^2_j|\lambda\right)$$

and a Gamma process is assumed as the prior distribution of the square of regularization parameter λ (Park and Casella, 2008),

$$p\left(\lambda^2\right)=Gamma(r,\theta)$$

Under these settings, the marginal prior of each marker is a double exponential, as follows:

$$p(a_j|\lambda)=\int N\left(a_j|0,\sigma^2_e\tau^2_j\right)Exp\left(\tau^2_j|\lambda^2\right)\partial\tau^2_j$$

Finally, the residual variance (σ^2_e) is assigned and a priori density of X^2 scaled inverse prior density, with degrees of freedom df_e, and scale parameter S_e.

Bayes R

In Bayes R (Erbe et al., 2012), which is an extension of Bayes Cπ with the assumption that the single nucleotide polymorphism effects (g) are derived from a mixture of normal distributions N(0, σ^2_k), according to the ratio vector

$$\boldsymbol{PR}=\{Prk|k=1,2,3,4\}$$

Variances of each component were fixed

$$\sigma^2_K=\left\{0,0.0001{*}\sigma^2_g,0.001{*}\sigma^2_g,0.01{*}\sigma^2_g\right\},$$

where σ^2_g is the total genetic variance. The prior distribution of the probability vector pr = ($pr1, pr2, pr3, pr4$) is a Dirichlet distribution (a multivariate generalization of a beta distribution), with $\alpha = 1$ (where α is a vector 4×1 of pseudo-counts, all with value 1, providing prior not informative). The Dirichlet distribution is a conjugate of the prior multinomial distribution, so that the posterior distribution of **pr** é ~ Dir (α + β), where β is a vector containing the number of single nucleotide polymorphisms in each distribution, estimated from the data.

Fixed effects included in the model were the same as single-step genomic best linear unbiased prediction describe above, however, a residual polygenic effect was added in Bayes R, with a covariance structure proportional to the relationship matrix.

Estimation of genomic breeding values

Based on the solutions of single nucleotide polymorphism effects estimated by the models studied, the direct genetic value was calculated using the following formula:

$$DGV_i=\sum_{j=1}^{p}w_{ij}\widehat{g}_j$$

where DGV_i is the direct genetic value; p is the number of single nucleotide polymorphisms; w_{ij} is the genotype of animal i for single nucleotide polymorphism j (coded as 0, 1 or 2); and $\widehat{g}_j$ is the estimated single nucleotide polymorphism substitution effect for single nucleotide polymorphism j that was estimated from the training population. The genomic values prediction was performed using the program BLUPF90 (Misztal, 2017) to single-step genomic best linear unbiased prediction; Bayesian generalized linear regression (Campos and Rodriguez, 2015) from R Program - R Core Team, 2018 to Bayes A, Bayes B, Bayes Cπ, and BLASSO; and Bayes R with scripts supported by R Program - R Core Team, 2018, to the method of the same name.

Table 4
Prediction ability of breeding genomic values for residual feed intake (RFI), DM intake (DMI), feed efficiency (FE), feed conversion ratio (FCR), residual BW gain (RG), and residual intake and BW gain (RIG) in Nelore cattle using different pseudo-phenotypes, models, and validation approaches.

Trait	Pseudo-phenotype	Bayes A	Bayes B	Bayes Cπ	BLASSO	Bayes R	ssGBLUP
Random							
RFI	Y						0.221 ± 0.02
	Y*	0.264 ± 0.01	0.251 ± 0.02	0.272 ± 0.01	0.259 ± 0.02	0.262 ± 0.02	
	EBV	0.143 ± 0.01	0.121 ± 0.01	0.182 ± 0.02	0.153 ± 0.01	0.152 ± 0.01	
	DEBV	0.112 ± 0.01	0.148 ± 0.01	0.134 ± 0.02	0.136 ± 0.01	0.121 ± 0.01	
DMI	Y						0.272 ± 0.01
	Y*	0.339 ± 0.01	0.300 ± 0.01	0.337 ± 0.01	0.345 ± 0.01	0.345 ± 0.01	
	EBV	0.221 ± 0.02	0.161 ± 0.01	0.239 ± 0.02	0.23 ± 0.02	0.239 ± 0.02	
	DEBV	0.121 ± 0.01	0.198 ± 0.01	0.235 ± 0.02	0.211 ± 0.01	0.253 ± 0.02	
FE	Y						0.143 ± 0.01
	Y*	0.170 ± 0.01	0.144 ± 0.01	0.126 ± 0.01	0.194 ± 0.02	0.157 ± 0.01	
	EBV	0.109 ± 0.01	0.107 ± 0.01	0.162 ± 0.03	0.124 ± 0.01	0.103 ± 0.01	
	DEBV	0.095 ± 0.01	0.106 ± 0.01	0.123 ± 0.01	0.103 ± 0.00	0.088 ± 0.01	
FCR	Y						0.120 ± 0.01
	Y*	0.130 ± 0.01	0.150 ± 0.01	0.113 ± 0.01	0.147 ± 0.01	0.148 ± 0.01	
	EBV	0.168 ± 0.02	0.091 ± 0.01	0.151 ± 0.02	0.167 ± 0.02	0.134 ± 0.01	
	DEBV	0.089 ± 0.01	0.089 ± 0.01	0.089 ± 0.01	0.065 ± 0.00	0.118 ± 0.01	
RG	Y						0.130 ± 0.02
	Y*	0.223 ± 0.01	0.221 ± 0.01	0.239 ± 0.01	0.214 ± 0.02	0.200 ± 0.02	
	EBV	0.114 ± 0.01	0.105 ± 0.01	0.139 ± 0.01	0.127 ± 0.01	0.124 ± 0.01	
	DEBV	0.109 ± 0.01	0.108 ± 0.01	0.117 ± 0.01	0.107 ± 0.01	0.123 ± 0.02	
RIG	Y						0.214 ± 0.02
	Y*	0.245 ± 0.02	0.270 ± 0.02	0.244 ± 0.01	0.251 ± 0.02	0.200 ± 0.02	
	EBV	0.121 ± 0.01	0.098 ± 0.01	0.121 ± 0.01	0.111 ± 0.01	0.141 ± 0.01	
	DEBV	0.154 ± 0.01	0.126 ± 0.01	0.128 ± 0.01	0.138 ± 0.01	0.16 ± 0.01	
Age							
RFI	Y						0.16
	Y*	0.19	0.18	0.18	0.19	0.22	
	EBV	0.10	0.11	0.10	0.11	0.14	
	DEBV	0.11	0.11	0.11	0.12	0.12	
DMI	Y						0.24
	Y*	0.25	0.24	0.25	0.24	0.22	
	EBV	0.15	0.11	0.21	0.16	0.14	
	DEBV	0.17	0.16	0.16	0.15	0.19	
FE	Y						0.08
	Y*	0.08	0.06	0.10	0.06	0.07	
	EBV	0.04	0.03	0.03	0.04	0.04	
	DEBV	0.06	0.04	0.06	0.04	0.09	
FCR	Y						0.10
	Y*	0.13	0.15	0.12	0.16	0.13	
	EBV	0.06	0.06	0.05	0.06	0.08	
	DEBV	0.05	0.06	0.05	0.06	0.08	
RG	Y						0.15
	Y*	0.19	0.18	0.18	0.16	0.08	
	EBV	0.10	0.10	0.11	0.11	0.06	
	DEBV	0.11	0.12	0.10	0.13	0.10	
RIG	Y						0.18
	Y*	0.26	0.25	0.27	0.23	0.19	
	EBV	0.13	0.13	0.13	0.13	0.09	
	DEBV	0.14	0.14	0.14	0.14	0.11	
EBV accuracy							
RFI	Y						0.17
	Y*	0.22	0.24	0.26	0.25	0.24	
	EBV	0.18	0.18	0.18	0.19	0.15	
	DEBV	0.18	0.16	0.15	0.14	0.15	
DMI	Y						0.23
	Y*	0.25	0.23	0.23	0.21	0.24	
	EBV	0.23	0.23	0.23	0.23	0.25	
	DEBV	0.14	0.13	0.19	0.18	0.25	
FE	Y						0.07
	Y*	0.12	0.15	0.11	0.13	0.07	
	EBV	0.12	0.10	0.08	0.10	0.05	
	DEBV	0.06	0.07	0.07	0.07	0.05	
FCR	Y						0.08
	Y*	0.13	0.16	0.13	0.08	0.11	
	EBV	0.05	0.10	0.09	0.09	0.09	
	DEBV	0.07	0.13	0.07	0.08	0.09	
RG	Y						0.14
	Y*	0.17	0.16	0.12	0.15	0.12	
	EBV	0.18	0.12	0.13	0.13	0.14	
	DEBV	0.12	0.13	0.13	0.13	0.14	

(continued on next page)

Table 4 (*continued*)

Trait	Pseudo-phenotype	Bayes A	Bayes B	Bayes Cπ	BLASSO	Bayes R	ssGBLUP
RIG	Y						0.14
	Y*	0.13	0.11	0.13	0.13	0.18	
	EBV	0.11	0.14	0.12	0.15	0.12	
	DEBV	0.11	0.11	0.10	0.11	0.12	

ssGBLUP: single-step genomic best linear unbiased prediction; BLASSO: Bayesian least absolute shrinkage and selection operator; Y: phenotype; Y*: adjusted phenotype; EBV: estimated breeding value; DEBV: deregressed estimated breeding value.

Validation and model comparison

Three validation strategies were used: age, EBV accuracy, and random cross-validation. For age, the animals were divided into two groups, being the animals born between 2010 and 2016, the training set; and animals born in 2017 the validation set. For EBV accuracy, the animals were divided into two groups, being animals with accuracy above 0.45 the training set; and below 0.45 the validation set.

Considering the random cross-validation strategy, the analyses were conducted repeatedly, considering the k-fold technique. The whole data set is randomly partitioned into 10 subsets (folds) each of approximately equal sizes. Then, nine subsets were used for training the genomic prediction model and the remaining subset was used for validation (Resende et al., 2012). At each repetition, cross-validation was performed in the group not used in the training set. After 10 repetitions, animals from the 10 sets were part of the validation group and had their genomic values predicted. All validation analyses, was implemented on R Program - R Core Team, 2018.

The GEBVs were calculated by an index combining parent average and direct genetic value (VanRaden et al., 2009):

$$GEBV_i = b_{DGV}DGV + b_{PA}PAV$$

where *PAV* is the parent average.

In the analyses that the used phenotype adjusted or Y* as pseudo-phenotype, PA was obtained by dividing the correlation between pseudo-phenotype and GEBV by the square root of the heritability of the trait (Pryce et al., 2012). This quantity is an approximation of the correlation between GEBV and true breeding value, which corresponds to true accuracy (Meuwissen et al., 2013). When EBV and DEBV were used as the pseudo-phenotype, the simple correlation of this quantity with the GEBV was considered as PA as an empirical measure of accuracy. In scenarios that used EBV or DEBV, the correlation between them and GEBV may be seen as an upper limit of the correlation between true breeding value and GEBV, which corresponds to a nonbiased accuracy estimate (Boddhireddy et al., 2014).

The regression between the pseudo-phenotype and the GEBVs was used to express the GEBV bias toward them, that is, the magnitude of inflation/deflation in relation to the GEBV. Hypothesis tests were performed to check the adequacy of the regression equations, through ANOVA (Montgomery and Runger, 2013). These hypothesis tests were performed considering a significance level of 0.05. The mean square error, which measures the individual differences between the values predicted by the model, was also calculated.

Results

Variance components and heritability obtained for evaluated traits are summarized in Table 1.

Table 4 shown the PA values obtained for FE-related traits considering validation approaches and pseudo-phenotypes. For traits with low heritability estimates, like FE and FCR, the average PA was lower, 0.07 ± 0.03 and 0.09 ± 0.03, respectively. Therefore, compared to FCR and FE, the higher selection response may be expected when using DMI, RFI, RG, and RIG as selection criterion. As expected, traits showing higher heritability estimates also displayed higher PA, such as DMI which show heritability estimates of 0.23 ± 0.04 and PA of 0.20 ± 0.05. The prediction models display similar PA, whose average values were 0.14 ± 0.06, 0.14 ± 0.06, 0.14 ± 0.06, 0.13 ± 0.06, 0.13 ± 0.06, and 0.15 ± 0.06 for Bayes A, Bayes B, Bayes Cπ, BLASSO, Bayes R, and single-step genomic best linear unbiased prediction, respectively (Table 4).

For cross-validation approaches, the highest PA was achieved by random, and the lowest PA was obtained using the age. The exception to this pattern was observed for DMI, which presented the highest heritability, and similar results for all validation approaches.

The phenotype adjusted for fixed effects as pseudo-phenotype displayed higher PA compared to EBV and DEBV. The average PA using EBV and DEBV as pseudo-phenotype were lower than those obtained with Y*, 30.0 and 34.3%, respectively. The EBV and DEBV as pseudo-phenotype displayed similar results for PA, 0.12 ± 0.05 and 0.11 ± 0.04, respectively.

In general, the regression coefficients were similar for all the models, validation approaches, and pseudo-phenotype; and it was close to 1 (1.06 ± 0.06) (Table 5), pointed out unbiased predictions.

The prediction models, validation approaches and pseudo-phenotype show similar mean squared error (0.05 ± 0.05), except for the adjusted phenotype (0.58 ± 1.93) (Supplementary Table S1).

Discussion

Heritability estimates for RFI, DMI, RG, and RIG were moderate and similar to those reported for beef cattle, ranging from 0.13 to 0.28 for RFI; from 0.25 to 0.36 for DMI; from 0.11 to 0.36 for RG, and from 0.12 to 0.34 for RIG (Berry and Crowley, 2012; de Oliveira et al., 2014; Grion et al., 2014; Ceacero et al., 2016; Olivieri et al., 2016; Silva et al., 2016; de Moraes et al., 2017). For FE and FCR, heritability estimates were lower than those presented in the literature for Nelore cattle, whose estimates range from 0.13 to 0.17 (FE) and 0.11 to 0.19 (FCR) (Grion et al., 2014; Ceacero et al., 2016; Olivieri et al., 2016; Silva et al., 2016; Polizel et al., 2018). Compared to other FE-related traits, the lower heritability obtained for FCR and FE is due to these traits are obtained as a ratio (Ceacero et al., 2016).

Higher PA for higher heritability traits was also observed by Silva et al. (2016) who evaluated a Nelore cattle experimental population and reported that ADG and DMI has shown the highest heritability (0.39 to 0.43, respectively) and PA (0.45 to 0.47 and 0.45 and 0.49, respectively). Similarly, Bolormaa et al. (2013), evaluating *Bos Taurus* (1 743; 223; 717; and 613 Angus, Murray Grey, Shorthorn, and Hereford cattle, respectively), *Bos indicus* (3 384 Brahman cattle), and composite beef cattle (550, 598, and 1 826 Belmont Red, Santa Gertrudis, and Tropical Composites cattle) for RFI, DMI, ADG, and metabolic BW, observed higher PA (0.13 to 0.36) for higher heritability traits (0.36 to 0.56).

Similar results for PA between the genomic methods are in agreement with the literature (De Los Campos et al., 2013; Lourenco et al., 2015; Strandén et al., 2017). These results can be attributed to the moderate to high accuracy of the EBV and DEBV as pseudo-phenotype (Table 3). The single-step genomic best linear unbiased prediction does not require pre-processing of phenotypes compared to multi-step methods (Aguilar et al., 2010; Christensen et al., 2012), accounting for the entire population structure to estimate GEBVs (Lourenco et al.,

Table 5
Phenotype and pseudo-phenotype regression coefficient on genomic breeding value with their respective *P*-values for residual feed intake (RFI), DM intake (DMI), feed efficiency (FE), feed conversion ratio (FCR), residual BW gain (RG), and residual intake and BW gain (RIG) in Nelore cattle using different models and validation approaches.

Trait	Pseudo-phenotype	Bayes A	Bayes B	Bayes Cπ	BLASSO	Bayes R	ssGBLUP
Random							
RFI	Y						1.081 (0.001)
	Y*	1.026 (0.001)	1.033 (0.001)	1.043 (0.001)	1.024 (0.001)	1.026 (0.001)	
	EBV	1.057 (0.001)	1.044 (0.001)	1.045 (0.001)	1.053 (0.001)	1.057 (0.001)	
	DEBV	1.049 (0.001)	1.071 (0.001)	1.058 (0.001)	1.075 (0.001)	1.049 (0.001)	
DMI	Y						1.082 (0.001)
	Y*	1.046 (0.001)	1.000 (0.003)	0.997 (0.002)	1.041 (0.001)	1.046 (0.001)	
	EBV	1.062 (0.001)	1.015 (0.003)	1.019 (0.002)	1.031 (0.001)	1.062 (0.001)	
	DEBV	1.062 (0.001)	1.015 (0.002)	1.047 (0.001)	1.051 (0.001)	1.050 (0.001)	
FE	Y						1.072 (0.001)
	Y*	1.021 (0.001)	1.058 (0.001)	1.029 (0.001)	1.095 (0.001)	1.021 (0.002)	
	EBV	1.057 (0.001)	1.081 (0.001)	1.069 (0.001)	1.048 (0.001)	1.057 (0.001)	
	DEBV	1.057 (0.001)	1.081 (0.001)	1.056 (0.001)	1.045 (0.001)	1.033 (0.001)	
FCR	Y						1.064 (0.001)
	Y*	1.043 (0.001)	1.094 (0.001)	1.059 (0.001)	1.071 (0.001)	1.043 (0.001)	
	EBV	1.078 (0.001)	1.058 (0.001)	1.028 (0.003)	1.027 (0.001)	1.078 (0.001)	
	DEBV	1.078 (0.001)	1.058 (0.001)	1.014 (0.002)	1.044 (0.001)	1.041 (0.001)	
RG	Y						1.062 (0.001)
	Y*	1.091 (0.001)	1.064 (0.001)	1.053 (0.001)	1.071 (0.001)	1.091 (0.001)	
	EBV	1.062 (0.001)	1.044 (0.001)	1.075 (0.001)	1.072 (0.001)	1.069 (0.001)	
	DEBV	1.069 (0.001)	1.044 (0.001)	1.043 (0.001)	1.071 (0.001)	1.074 (0.001)	
RIG	Y						1.046 (0.001)
	Y*	1.103 (0.003)	1.056 (0.001)	1.088 (0.001)	1.074 (0.001)	1.103 (0.001)	
	EBV	1.059 (0.001)	1.036 (0.001)	1.027 (0.003)	1.055 (0.001)	1.059 (0.002)	
	DEBV	1.059 (0.001)	1.036 (0.002)	1.067 (0.001)	1.040 (0.001)	1.055 (0.001)	
Age							
RFI	Y						1.073 (0.001)
	Y*	0.980 (0.003)	1.090 (0.001)	1.020 (0.001)	0.940 (0.002)	1.120 (0.001)	
	EBV	1.140 (0.001)	1.120 (0.001)	1.070 (0.001)	1.110 (0.001)	1.080 (0.001)	
	DEBV	1.160 (0.001)	1.110 (0.001)	1.120 (0.001)	1.080 (0.001)	1.120 (0.001)	
DMI	Y						1.023 (0.001)
	Y*	0.960 (0.002)	0.960 (0.003)	0.970 (0.002)	0.960 (0.002)	1.030 (0.001)	
	EBV	1.160 (0.001)	1.180 (0.001)	1.120 (0.001)	1.170 (0.001)	1.090 (0.001)	
	DEBV	1.150 (0.001)	1.120 (0.001)	1.060 (0.001)	1.140 (0.001)	1.120 (0.001)	
FE	Y						1.133 (0.001)
	Y*	0.950 (0.002)	1.060 (0.001)	1.010 (0.001)	0.920 (0.002)	1.150 (0.001)	
	EBV	1.070 (0.001)	0.950 (0.001)	1.190 (0.001)	1.120 (0.001)	0.960 (0.001)	
	DEBV	1.080 (0.001)	0.990 (0.002)	1.060 (0.001)	0.960 (0.001)	1.090 (0.001)	
FCR	Y						1.050 (0.001)
	Y*	1.030 (0.001)	0.980 (0.002)	1.070 (0.001)	0.990 (0.002)	1.040 (0.001)	
	EBV	1.180 (0.001)	1.190 (0.001)	0.980 (0.002)	1.190 (0.001)	1.090 (0.001)	
	DEBV	1.040 (0.001)	1.070 (0.001)	1.030 (0.002)	1.190 (0.001)	1.020 (0.002)	
RG	Y						1.117 (0.001)
	Y*	1.040 (0.001)	0.930 (0.002)	1.040 (0.002)	0.950 (0.002)	1.060 (0.001)	
	EBV	1.080 (0.001)	1.050 (0.001)	1.130 (0.001)	1.090 (0.001)	1.060 (0.001)	
	DEBV	1.090 (0.001)	1.020 (0.002)	1.050 (0.001)	0.960 (0.002)	1.020 (0.002)	
RIG	Y						1.057 (0.001)
	Y*	0.940 (0.002)	1.060 (0.001)	0.980 (0.002)	1.030 (0.001)	1.090 (0.001)	
	EBV	1.170 (0.001)	1.120 (0.001)	1.100 (0.001)	1.070 (0.001)	0.960 (0.002)	
	DEBV	1.090 (0.001)	0.910 (0.003)	0.940 (0.002)	0.930 (0.003)	1.040 (0.001)	
EBV accuracy							
RFI	Y						1.073 (0.001)
	Y*	1.090 (0.001)	0.910 (0.003)	0.940 (0.002)	0.930 (0.002)	1.040 (0.001)	
	EBV	1.150 (0.001)	1.120 (0.001)	1.030 (0.002)	1.020 (0.002)	0.940 (0.002)	
	DEBV	1.080 (0.001)	1.120 (0.001)	1.150 (0.001)	1.130 (0.001)	1.090 (0.001)	
DMI	Y						1.097 (0.001)
	Y*	1.070 (0.001)	1.040 (0.001)	1.060 (0.001)	1.040 (0.001)	1.060 (0.001)	
	EBV	0.920 (0.002)	1.090 (0.001)	1.060 (0.001)	1.090 (0.001)	0.960 (0.002)	
	DEBV	0.950 (0.002)	0.960 (0.002)	1.050 (0.001)	1.190 (0.001)	1.090 (0.001)	
FE	Y						1.113 (0.001)
	Y*	1.040 (0.001)	1.010 (0.001)	1.030 (0.002)	0.950 (0.002)	1.130 (0.001)	
	EBV	0.910 (0.002)	1.010 (0.001)	1.030 (0.001)	0.980 (0.002)	1.100 (0.001)	
	DEBV	1.120 (0.001)	1.160 (0.001)	1.120 (0.001)	1.160 (0.001)	1.120 (0.001)	
FCR	Y						1.090 (0.001)
	Y*	1.050 (0.001)	1.010 (0.002)	1.030 (0.001)	1.030 (0.001)	1.010 (0.002)	
	EBV	1.060 (0.001)	1.090 (0.001)	1.040 (0.001)	1.010 (0.002)	1.050 (0.001)	
	DEBV	1.150 (0.001)	1.130 (0.001)	0.910 (0.002)	1.090 (0.001)	1.060 (0.001)	
RG	Y						1.127 (0.01)
	Y*	0.910 (0.002)	0.950 (0.002)	1.090 (0.001)	1.080 (0.001)	1.040 (0.002)	
	EBV	1.090 (0.001)	1.050 (0.001)	1.060 (0.001)	1.040 (0.002)	1.080 (0.001)	
	DEBV	1.140 (0.001)	1.130 (0.001)	0.910 (0.002)	1.150 (0.001)	1.170 (0.001)	

(continued on next page)

Table 5 (*continued*)

Trait	Pseudo-phenotype	Bayes A	Bayes B	Bayes Cπ	BLASSO	Bayes R	ssGBLUP
RIG	Y						1.70 (0.001)
	Y*	1.010 (0.003)	1.040 (0.001)	1.020 (0.002)	1.130 (0.001)	1.080 (0.001)	
	EBV	0.920 (0.002)	1.090 (0.001)	1.080 (0.001)	1.030 (0.002)	1.100 (0.001)	
	DEBV	1.140 (0.001)	0.930 (0.002)	1.160 (0.001)	1.170 (0.001)	1.030 (0.002)	

ssGBLUP: single-step genomic best linear unbiased prediction; BLASSO: Bayesian least absolute shrinkage and selection operator; Y: phenotype; Y*: adjusted phenotype; EBV: estimated breeding value; DEBV: deregressed estimated breeding value. [a,b]Coefficients from regression analyses that were significant in ANOVA at a significance level of 0.05 (Christensen, 2015).

2014). However, when high accuracy pseudo-phenotypes are used, the markers effects may be estimated properly in Bayesian methods. The similarity between the prediction methods can also be attributed to a large number of records and because most animals with phenotypes had genotypic information, in these conditions frequentist and Bayesian methods reach the same PA (Meuwissen, 2009).

Another reason for the similarity between the prediction methods is the complex and polygenic nature of the FE traits, since different models tend to show similar predictive ability when the traits are affected by many small-effect loci (De Los Campos et al., 2013). In this case, prediction models show similar PA, since the genetic architecture appeared to approach the infinitesimal model (Lee et al., 2017).

Still, other factors besides the PA must be taken into consideration in choosing the genomic prediction method used as data set. The single-step genomic best linear unbiased prediction may be indicated for the population that has genotyped and non-genotyped animals and the individuals are related and connected through the pedigree (Silva et al., 2016). Since, this method associates phenotypic information from non-genotyped animals that are related to genotyped animals via a combined relationship matrix and using all available information (Lourenco et al., 2014; Silva et al., 2016). In single-step genomic best linear unbiased prediction, the information from animals' relatives is taken into account priority rather than the individual information (Silva et al., 2016). Another advantage of single-step genomic best linear unbiased prediction was the lower computational demand and time, also simplifying the evaluations and allowing the selection of animals only with genotype information (Aguilar et al., 2010). The key aspect of genomic selection implementation is obtaining GEBVs for the selection of non-phenotyped animals.

On the other hand, frequentist methods require more records to reach a high PA than does a Bayesian method, and when the number of records is small and the marker density is high, Bayesian methods may show superior PA (Meuwissen, 2009). For data set with missing pedigree, genomic prediction using single-step genomic best linear unbiased prediction can be compromised, increasing bias and leading to convergence problems, and incompatibility between kinship and genomic matrices (Misztal et al., 2013; Tonussi et al., 2017). In addition, for data set with missing pedigree, pseudo-phenotypes can be of low accuracy, compromising genomic selection. For these data sets the Bayesian methods and Y* as pseudo-phenotype can be recommended.

The low PA for validation using age or EBV accuracy may be due to few generations and few animals with progeny records evaluated for the FE traits. In general, older animals have more information and EBV with higher accuracy to be used as a training population, which does not apply here, as these traits were recently included as a selection criterion, so the training and validation population display the same information available. As a result, there is no gain in PA using older animals or based on accuracy. Moreover, validation with age and EBV accuracy grouping gives no consideration of the relationship between the evaluation subsets. Prediction ability depends not just on a large population, but also on how closely related the individuals are. In addition, animals grouping by age or EBV accuracy can lead to higher variation in the number of animals in training and validation sets when compared to the random method, which may also have influenced PA (VanRaden et al., 2009).

In the randomness approaches, predictions may be optimistically biased because the individuals were more related than validation by age or EBV accuracy. This leads to higher PA, even for low or moderate heritability traits (Pérez-Cabal et al., 2012). Results reported by Silva et al. (2016), who evaluated RFI, FCR, ADG, and DMI, are in agreement with this study, where higher PA values with random cross-validation were obtained compared to grouping by animal's age. Evaluating RFI and weaning weight in Holstein heifers Pryce et al. (2012) also observed that the GEBVs PA is influenced by the relationship between two individuals, and the higher levels of genomic relationships among animals, more accurate are the predicted genomic values. Thus, in a related population, with young animals and without progenies cross-validation can be used to better test results and prediction and lead to drawn robust conclusions. On another hand, for high heritability traits, the genomic prediction using validation based on age or EBV accuracy may be performed without losing PA. The feasibility of using older animals to estimate prediction equations for traits with higher heritability and to validate in young animals would enable predict next- generations' performance and anticipate evaluation and decision-making.

The ratios between genetic signal and noise differ for Y*, EBV, and DEBV (Daetwyler et al., 2013). Thus, different results for genomic predictions using each one of the pseudo-phenotype were expected and an advantage for either one of these depends on the data set used and the specific application (Boddhireddy et al., 2014). Estimated breeding value accuracies are influenced by the trait heritability and this fact supports the superiority of the PA obtained with the adjusted phenotype over EBVs and DEBVs, mainly for FCR and FE. Previous studies also reported higher PA using Y* as pseudo-phenotype to estimate markers effects (Fernandes Júnior et al., 2016). When low heritability or low EBV accuracy trait and population with missing pedigree is evaluated, the use of Y* as pseudo-phenotype may be indicated. The adoption of Y* can be a strategy for commercial populations that present phenotype and genotype, but without pedigree records.

The similarity in PA using EBV and DEBV as pseudo-phenotype may be attributed to the full knowledge of the kinship matrix (Chiaia et al., 2018), increasing the reliability of the deregressed EBV. Therefore, for data set without missing pedigree, use of DEBV as pseudo-phenotype brings the same PA than EBV, with the advantage to avoid double-counting of information and double shrinkage of the direct genomic values using EBVs as pseudo-phenotypes (Garrick et al., 2009). However, the use of multiple sires and the unknown parental identification are common in extensive beef cattle production systems, which results in missing pedigree, in this case, the use of Y* as pseudo-phenotype for Bayesian models or phenotype for single-step genomic best linear unbiased prediction can be recommended.

It is important to emphasize that higher PA values for FE-related traits were obtained by Silva et al. (2016) evaluating Nelore cattle. This difference may be due to the closely related experimental population composed of a unique herd with smaller environmental variance. This study evaluated animals from different herds and was subjected to varying environmental conditions, which may lead to the lowest additive genetic variance. Nevertheless, the use of genomic information to predict breeding values proved to be a feasible alternative, aiming for the accurate identification of more efficient animals for feed utilization. Considering the genomic PA for FE is around 0.30 in beef and dairy cattle (Mujibi et al., 2011; Pryce et al., 2012; Bolormaa et al., 2013), and results obtained in this study are within the range observed in preview reports.

The small difference between models for prediction bias can be attributed to the larger sample size (Silva et al., 2016), since in this case, pseudo-phenotypes are more accurate (Table 3), reducing the differences between models. In addition, most animals with phenotype have genotypic information in this study, as mentioned before, reducing the predictive advantage of single-step genomic best linear unbiased prediction. The higher number of phenotyped and genotyped animals may result in a decrease in prediction bias (VanRaden et al., 2009), even for low heritability traits and regardless of the prediction model used. Bolormaa et al. (2013) reported that traits with a large number of evaluated and genotyped animals provided less biased GEBVs prediction.

Another factor that influenced the prediction bias is the kinship matrix between the animals of the validation and training population. Although most of the evaluated animals are not a selected population for FE, and come from several commercial herds with higher genetic diversity, the level of genomic relationship between individuals is high, since there are several sires that connect the FE tests. Genomic predictions tend to be more precise when animals in the validation population are related to animals in the training population (Saatchi et al., 2011). This pattern agrees with the results presented by Pryce et al. (2012), evaluating FE Holstein heifers.

The similarity between the models based on bias may be attributed also to the polygenic nature of the evaluated traits. The FE traits show complex nature and controlled by several quantitative trait loci with small effect (Rolf et al., 2012; Santana et al., 2014; Olivieri et al., 2016). In accordance, Daetwyler et al. (2013) observed that the models show the same prediction capacity when the traits are affected by many small-effect loci.

The highest mean square error observed for Y* may be due to this parameter follows the magnitude of the pseudo-phenotype. Fernandes Júnior et al. (2016) evaluated carcass traits in Nelore cattle and also reported variation in mean square error as a function of the magnitude of the trait. These results do not necessarily suggest the rejection of models with high mean square error using Y* when compared to other pseudo-phenotypes included in this study, since the adjusted phenotype as a pseudo-phenotype showed higher PA and similar bias.

The use of genomic information to predict breeding values proved to be a feasible alternative, aiming for the accurate identification of animals with better feed utilization. Residual BW gain and RIG show similar PA to RFI, being an effective selection criterion for FE. The prediction methods show similar results for PA and bias. Random cross-validation schemes presented the highest PA than EBV accuracy and age, being indicated for use in populations composed mainly of young animals, and for traits with few generations of data recording. For high heritability traits, the validation can be done by age reaching the same PA as random validation, enabling the training in older animals, validation in young animals, and the prediction of the next-generation genetic merit, using genomic information. The adjusted phenotype pseudo-phenotype was more appropriate to estimate marker effects when few generations of phenotypes and genotypes of young animals are available and pedigree information is not available. The results from this study would support breeders to identify prediction methods, pseudo-phenotypes, and validation approaches that are more viable for genomic prediction for FE-related traits, considering the data structure and information available for traits with different genetic background.

Supplementary materials

Supplementary data to this article can be found online at https://doi.org/10.1016/j.animal.2020.100085.

Ethics approval

The research project was approved by the Committee on Ethics in the Use of Animals (CEUA/PRPI) of the Federal University of Goiás (UFG), according to protocol N° 088/18 issued by this institution.

Data and model availability statement

The data that support the findings of this study are available on request from the corresponding author upon reasonable request. The data are not publicly available due to privacy or legal restrictions, because it belongs to a commercial breeding program.

Author ORCIDs

L. C. Brunes https://orcid.org/0000-0001-9012-520X.
F. Baldi https://orcid.org/0000-0003-4094-2011.
F. B. Lopes https://orcid.org/0000-0003-1292-3331.
R. B. Lobo https://orcid.org/0000-0001-6016-5817.
R. Espigolan https://orcid.org/0000-0003-2586-1643.
C. U. Magnabosco https://orcid.org/0000-0002-7274-0134.

Author contributions

L. C. Brunes: Conceptualization, methodology, formal analysis, writing – original draft. F. Baldi: Conceptualization, methodology, formal analysis, writing – original draft. F. B. Lopes: Methodology, formal analysis, writing – review & editing. M. G. Narciso: Formal analysis, data curation. B. Lobo: Data curation, writing – review & editing. R. Espigolan: writing – review & editing. M. F. O. e Costa: Writing – review & editing. C. U. Magnabosco: Conceptualization, methodology, formal analysis, writing – review & editing.

Declaration of interest

The authors declare that they do not have any conflict of interest regarding the topics approached within this study.

Acknowledgements

The authors thank National Association of Breeders and Researchers for providing data set and genealogical data. This study is part of the doctoral thesis of the first author.

Financial support statement

This work was supported by Coordination for the Improvement of Higher Education Personnel (CAPES) for providing doctoral financial support for the first, grant #88887.162699/2018-00.

References

Aguilar, I., Misztal, I., Johnson, D., Legarra, A., Tsuruta, S., Lawlor, T., 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. Journal of Dairy Science 93, 743–752.
Berry, D.P., Crowley, J.J., 2012. Residual intake and body weight gain: a new measure of efficiency in growing cattle. Journal of Animal Science 90, 109–115.
Boddhireddy, P., Kelly, M.J., Northcutt, S., Prayaga, K.C., Rumph, J., DeNise, S., 2014. Genomic predictions in Angus cattle: comparisons of sample size, response variables, and clustering methods for cross-validation. Journal of Animal Science 92, 485–497.
Bolormaa, S., Pryce, J.E., Kemper, K., Savin, K., Hayes, B.J., Barendse, W., Zhang, Y., Reich, C. M., Mason, B.A., Bunch, R.J., Harrison, B.E., Reverter, A., Herd, R.M., Tier, B., Graser, H. U., Goddard, M.E., 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. Journal of Animal Science 91, 3088–3104.
Brunes, L.C., Baldi, F., Lopes, F.B., Lobo, R.B., Espigolan, R., Costa, M.F.O., Stafuzza, N.B., Magnabosco, C.U., 2020. Weighted single-step genome-wide association study and pathway analyses for feed efficiency traits in Nellore cattle. Journal of Animal Bredding and Genetics 00, 1–22.
Campos, G.D.L., Rodriguez, P., 2015. BLR: Bayesian linear regression. R Package Retrieved on 01 April 2019, from. https://cran.r-project.org/web/packages/BLR/index.html.
Ceacero, T.M., Mercadante, M.E.Z., Cyrillo, J.N.D.S.G., Canesin, R.C., Bonilha, S.F.M., De Albuquerque, L.G., 2016. Phenotypic and genetic correlations of feed efficiency traits with growth and carcass traits in Nelore cattle selected for postweaning weight. PLoS ONE 11, 1–11.
Chiaia, H.L.J., Peripolli, E., Silva, R.M.O., Feitosa, F.L.B., de Lemos, M.V.A., Berton, M.P., Olivieri, B.F., Espigolan, R., Tonussi, R.L., Gordo, D.G.M., de Albuquerque, L.G., de

Oliveira, H.N., Ferrinho, A.M., Mueller, L.F., Kluska, S., Tonhati, H., Pereira, A.S.C., Aguilar, I., Baldi, F., 2018. Genomic prediction ability for beef fatty acid profile in Nelore cattle using different pseudo-phenotypes. Journal of Applied Genetics 59, 493–501.
Christensen, R, 2015. Analysis of Variance, Design, and Regression: Linear Modeling for Unbalanced Data. Chapman & Hall, London, UK.
Christensen, O.F., Madsen, P., Nielsen, B., Ostersen, T., Su, G., 2012. Single-step methods for genomic evoluation in pigs. Animal 6, 1565–1571.
Daetwyler, H.D., Calus, M.P.L., Pong-Wong, R., De Los Campos, G., Hickey, J.M., 2013. Genomic prediction in animals and plants: simulation of data, validation, reporting, and benchmarking. Genetics 193, 347–365.
De Los Campos, G., Hickey, J.M.M., Pong-Wong, R., Daetwyler, H.D.D., Calus, M.P.L., 2013. Whole genome regression and prediction methods applied to plant and animal breeding. Genetics 193 genetics.112.143313.
de Moraes, G.F., Abreu, L.R.A., Ferreira, I.C., Pereira, I.G., 2017. Genetic analysis of residual feed intake adjusted for fat and carcass and performance traits in a Nellore herd. Ciência Rural 47, 1–6.
de Oliveira, P.S.N., Cesar, A.S.M., do Nascimento, M.L., Chaves, A.S., Tizioto, P.C., Tullio, R.R., Lanna, D.P.D., Rosa, A.N., Sonstegard, T.S., Mourao, G.B., Reecy, J.M., Garrick, D.J., Mudadu, M.A., Coutinho, L.L., Regitano, L.C.A., 2014. Identification of genomic regions associated with feed efficiency in Nelore cattle. BMC Genetics 15, 100.
Erbe, M., Hayes, B.J., Matukumalli, L.K., Goswami, S., Bowman, P.J., Reich, C.M., Mason, B.A., Goddard, M.E., 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. Journal of Dairy Science 95, 4114–4129.
Fernandes Júnior, G.A., Rosa, G.J.M., Valente, B.D., Carvalheiro, R., Baldi, F., Garcia, D.A., Gordo, D.G.M., Espigolan, R., Takada, L., Tonussi, R.L., Andrade, W.B.F., Magalhaes, A. F.B., Chardulo, L.A.L., Tonhati, H., de Albuquerque, L.G., 2016. Genomic prediction of breeding values for carcass traits in Nellore cattle. Genetics, Selection, Evolution 48, 7.
Garrick, J.D., Taylor, J.F., Fernando, R.L., 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genetics, Selection, Evolution 41, 1–8.
Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernado, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian statistics 4. Clarendon Press, Oxford, UK, pp. 169–193.
Grion, A.L., Mercadante, M.E.Z., Cyrillo, J.N.S.G., Bonilha, S.F.M., Magnani, E., Branco, R.H., 2014. Selection for feed efficiency traits and correlated genetic responses in feed intake and weight gain of Nellore cattle. Journal of Animal Science 92, 955–965.
Habier, D., Fernando, R.L.R., Kizilkaya, K., Garrick, D.J.D., 2011. Extension of the bayesian alphabet for genomic selection. BMC Bioinformatics 12, 186.
Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules. In: Munro, H.N. (Ed.), Mammalian protein metabolism. Academic Press, New York, USA, pp. 21–132.
Koch, R., Swiger, L., Chambers, D., Gregory, K., 1963. Efficiency of feed use in beef cattle. Journal of Animal Science 22, 486–494.
Lee, J., Cheng, H., Garrick, D., Golden, B., Dekkers, J., Park, K., Lee, D., Fernando, R., 2017. Comparison of alternative approaches to single-trait genomic prediction using genotyped and non-genotyped Hanwoo beef cattle. Genetics, Selection, Evolution 49, 1–9.
Lobo, R.B., Bezerra, L.A.F., Faria, C.U., Vozzi, P.A., Magnabosco, C.U., Bergmann, J.A.G., Oliveira, H.N., Castro, L.M., Pereira, A.S.C., Aguilar, I., Baldi, F., 2019. Sumário de touros Nelore, Guzerá, Brahman, Tabapua. Associação Nacional de Criadores e Pesquisadores, Ribeirão Preto, SP, Brazil.
Lopes, F.B., 2017. DRP package. Retrieved on 01 August 2019, from https://github.com/camult/DRP.
Lourenco, D.A.L., Misztal, I., Tsuruta, S., Aguilar, I., Ezra, E., Ron, M., Shirak, A., Weller, J.L., 2014. Methods for genomic evaluation of a relatively small genotyped dairy population and effect of genotyped cow information in multiparity analyses. Journal of Dairy Science 97, 1742–1752.
Lourenco, D.A.L., Tsuruta, S., Fragomeni, B.O., Masuda, Y., Aguilar, I., Legarra, A., Bertrand, J. K., Amen, T.S., Wang, L., Moser, D.W., Misztal, I., 2015. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. Journal of Animal Science 93, 2653–2662.
Magnabosco, C.U., Lopes, F.B., Baldi, F., Magnabosco, V.S., Lobo, R.B., Pereira, L.S., Espigolan, R., Brunes, L.C., 2020. Genomic prediction accuracy for feed efficiency related traits using different pseudo-phenotypes, prediction and validation methods in Nellore cattle. Proceedings of the American Society of Animal Science, 19 July 2020, Wisconsin, USA, pp. 1–2.
Mendes, E.D.M., de Faria, C.U., Sainz, R.D., Silveira, A.C.L., Magnabosco, C.U., Eifert, E.C., Baldi, F., Ventura, H.T., de Castro, L.M., Ribeiro, L.B., Cavalcanti, L.F.L., Josakian, L.A., da Silva, L.O.C., Mercadante, M.E.Z., Ribas, M.N., Lobo, R.B., Arnandes, R.H.B., da Gomes, R.C., Farjalla, Y.B., 2020. Procedimentos para mensuração de consumo individual de alimento em bovinos de corte. Associação Nacional de Criadores e Pesquisadores, Ribeirão Preto, SP, Brazil.
Meuwissen, T., Hayes, B., Goddard, M., 2013. Accelerating improvement of livestock with genomic selection. Annual Review of Animal Biosciences 1, 221–237.
Meuwissen, T.H.E., 2009. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. Genetics, Selection, Evolution 41, 1–9.
Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.
Meyer, K., Houle, D., 2013. Sampling based approximation of confidence intervals for functions of genetic covariance matrices. Proceedings of the 20th Association for the advancement of animal breeding and genetics conference, 20 October 2013, Napier, New Zealand, pp. 523–526.
Misztal, I., 2017. BLUPF90 family of programs. Retrieved on 01 April 2019, from. http://nce.ads.uga.edu/html/projects/programs/.
Misztal, I., Vitezica, Z.G., Legarra, A., Aguilar, I., Swan, A.A., 2013. Unknown-parent groups in single-step genomic evaluation. Journal of Animal Breeding and Genetics 130, 252–258.
Montgomery, D.C., Runger, G.C., 2013. Applied statistics and probability for engineers. 6th edition. John Wiley & Sons, Hoboken, NJ, USA.
Mujibi, F.D.N., Nkrumah, J.D., Durunna, O.N., Grant, J.R., Mah, J., Wang, Z., Basarab, J., Plastow, G., Crews Jr., D.H., Moore, S.S., 2011. Associations of marker panel scores with feed intake and efficiency traits in beef cattle using preselected single nucleotide polymorphisms. Journal of Animal Science 89, 3362–3371.
Neves, H.R., Carvalheiro, R., O'Brien, A.M.P., Utsunomiya, Y.T., do Carmo, A.S., Schenkel, F. S., Solkner, J., McEwan, J.C., Tassel, C.P.V., Cole, J.B., da Silva, M.V.G.B., Queiroz, S.A., Sostegard, T.S., Garcia, F.J., 2014. Accuracy of genomic predictions in *Bos indicus* (Nellore) cattle. Genetics, Selection, Evolution 46, 17.
Olivieri, B.F., Mercadante, M.E.Z., Cyrillo, J.N.S.G., Branco, R.H., Bonilha, S.F., de Albuquerque, L.G., Silva, R.M., Baldi, F., 2016. Genomic regions associated with feed efficiency indicator traits in an experimental Nellore cattle population. PLoS ONE 11, 1–19.
Ostersen, T., Christensen, O.F., Henryon, M., Nielsen, B., Su, G., Madsen, P., 2011. Deregressed EBV as the response variable yield more reliable genomic predictions than traditional EBV in purebred pigs. Genetics, Selection, Evolution 43, 1–6.
Park, T., Casella, G., 2008. The Bayesian lasso. Journal of the American Statistical Association 103, 681–686.
Pérez-Cabal, M.A., Vazquez, A.I., Gianola, D., Rosa, G.J.M., Wiegel, K.A., 2012. Accuracy of genome-enabled prediction in a dairy cattle population using different cross-validation layouts. Frontiers in Genetics 3, 1–7.
Polizel, G.H.G., Grigoletto, L., Carvalho, M., Rossi Junior, P., Ferraz, J.B.S., Santana, M.H.A., 2018. Genetic correlations and heritability estimates for dry matter intake, weight gain and feed efficiency of Nellore cattle in feedlot. Livestock Science 214, 209–210.
Pryce, J.E., Arias, J., Bowman, P.J., Davis, S.R., Macdonald, K.A., Waghorn, G.C., Wales, W.J., Williams, Y.J., Spelman, R.J., Hayes, B.J., 2012. Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. Journal of Dairy Science 95, 2108–2119.
R Program - R Core Team, 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria Retrieved on 12 December 2018, from. https://www.r-project.org/.
Resende, M.D.V., Silva, F.F., Lopes, P.S.P., Azevedo, C.F., 2012. Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência Bayesiana (MCMC), Regressão Aleatória Multivariada (RRM) e Estatística Espacial. Universidade Federal de Viçosa/Departamento de Estatistica, Viçosa, MG, Brazil.
Rolf, M.M., Taylor, J.F., Schnabel, R.D., McKay, S.D., McClure, M.C., Northcutt, S.L., Kerley, M. S., Weaber, R.L., 2012. Genome-wide association analysis for feed efficiency in Angus cattle. Animal Genetics 43, 367–374.
Rosa, G.J.M., Padovani, C.R., Gianola, D., 2003. Robust linear mixed models with normal/independent distributions and Bayesian MCMC implementation. Biometrical Journal 45, 573–590.
Runcie, D., Cheng, H., 2019. Pitfalls and remedies for cross validation with multi-trait genomic prediction methods. G3: Genes, Genomes, Genetics 9, 3727–3741.
Saatchi, M., McClure, M.C., McKay, S.D., Rolf, M.M., Kim, J., Decker, J.E., Taxis, T.M., Chapple, R.H., Ramey, H.R., Northcutt, S.L., Bauck, S., Woodward, B., Dekkers, J.C.M., Fernando, R.L., Schnabel, R.D., Garrick, D.J., Taylor, F.J., 2011. Accuracies of genomic breeding values in American Angus beef cattle using k-means clustering for cross-validation. Genetics, Selection, Evolution 43, 40.
Santana, M.H.A., Utsunomiya, Y.T., Neves, H.H.R., Gomes, R.C., Garcia, J.F., Fukumasu, H., Silva, S.L., Olvieira Junior, G.A., Alexandre, P.A., Leme, P.R., Brassaloti, R.A., Coutinho, L.L., Lopes, T.G., Meirelles, F.V., Eler, J.P., Ferraz, J.B.S., 2014. Genome-wide association analysis of feed intake and residual feed intake in Nellore cattle. BMC Genetics 15, 1–8.
Savietto, S., Berry, D.P., Friggens, N.C., 2014. Towards an improved estimation of the biological components of residual feed intake in growing cattle. Journal of Animal Science 92, 467–476.
Silva, R.M., Fragomeni, B.O., Lourenco, D.A., Magalhães, A.F., Irano, N., Carvalheiro, R., Canesin, R.C., Mercadante, M.E., Boligon, A.A., Baldi, F.S., Misztal, I., Albuquerque, L. G., 2016. Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population. Journal of Animal Science 94, 3613–3623.
Strandén, I., Matilainen, K., Aamand, G.P., Mantysaari, E.A., 2017. Solving efficiently large single-step genomic best linear unbiased prediction models. Journal of Animal Breeding and Genetics 134, 264–274.
Tonussi, R.L., Silva, R.M.O., Magalhães, A.F.B., Espigolan, R., Peripolli, E., Oliveiri, B.F., Feitosa, F.L.B., Lemos, M.V.A., Berton, M.P., Chiaia, H.L.J., Pereira, A.S.C., Lobo, R.B., Bezerra, L.A.F., Magnabosco, C.U., Lourenço, D.A.L., Aguilar, I., Baldi, F., 2017. Application of single step genomic BLUP under different uncertain paternity scenarios using simulated data. PlosOne 12, 1–14.
VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. Journal of Dairy Science 91, 4414–4423.
VanRaden, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J. F., Schenkel, F.S., 2009. Invited review: reliability of genomic predictions for north American Holstein bulls. Journal of Dairy Science 92, 16–24.