Journal of Animal Science, 2022, **100**, 1–13
https://doi.org/10.1093/jas/skac009
Advance access publication 14 January 2022
Animal Genetics and Genomics

Development of genomic predictions for Angus cattle in Brazil incorporating genotypes from related American sires

Gabriel Soares Campos,[†,1] Fernando Flores Cardoso,[‡] Claudia Cristina Gulias Gomes,[‡] Robert Domingues,[‡] Luciana Correia de Almeida Regitano,[‖] Marcia Cristina de Sena Oliveira,[‖] Henrique Nunes de Oliveira,[$] Roberto Carvalheiro,[$] Lucia Galvão Albuquerque,[$] Stephen Miller,[¶] Ignacy Misztal,[†] and Daniela Lourenco[†]

[†]Department of Animal and Dairy Science, University of Georgia, Athens, GA 30602, USA
[‡]Embrapa Pecuária Sul, Bagé, RS 96401-970, Brazil
[‖]Embrapa Pecuária Sudeste, São Carlos, SP 13560-970, Brazil
[$]Departamento de Zootecnia, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, SP 14884-900, Brazil
[¶]Angus Genetics Inc., St. Joseph, MO 64506, USA

[1]Corresponding author: gabrielsoarescampos@hotmail.com

Abstract

Genomic prediction has become the new standard for genetic improvement programs, and currently, there is a desire to implement this technology for the evaluation of Angus cattle in Brazil. Thus, the main objective of this study was to assess the feasibility of evaluating young Brazilian Angus (BA) bulls and heifers for 12 routinely recorded traits using single-step genomic BLUP (ssGBLUP) with and without genotypes from American Angus (AA) sires. The second objective was to obtain estimates of effective population size (N_e) and linkage disequilibrium (LD) in the Brazilian Angus population. The dataset contained phenotypic information for up to 277,661 animals belonging to the Promebo breeding program, pedigree for 362,900, of which 1,386 were genotyped for 50k, 77k, and 150k single nucleotide polymorphism (SNP) panels. After imputation and quality control, 61,666 SNPs were available for the analyses. In addition, genotypes from 332 American Angus (AA) sires widely used in Brazil were retrieved from the AA Association database to be used for genomic predictions. Bivariate animal models were used to estimate variance components, traditional EBV, and genomic EBV (GEBV). Validation was carried out with the linear regression method (LR) using young-genotyped animals born between 2013 and 2015 without phenotypes in the reduced dataset and with records in the complete dataset. Validation animals were further split into progeny of BA and AA sires to evaluate if their progenies would benefit by including genotypes from AA sires. The N_e was 254 based on pedigree and 197 based on LD, and the average LD (±SD) and distance between adjacent single nucleotide polymorphisms (SNPs) across all chromosomes were 0.27 (±0.27) and 40743.68 bp, respectively. Prediction accuracies with ssGBLUP outperformed BLUP for all traits, improving accuracies by, on average, 16% for BA young bulls and heifers. The GEBV prediction accuracies ranged from 0.37 (total maternal for weaning weight and tick count) to 0.54 (yearling precocity) across all traits, and dispersion (LR coefficients) fluctuated between 0.92 and 1.06. Inclusion of genotyped sires from the AA improved GEBV accuracies by 2%, on average, compared to using only the BA reference population. Our study indicated that genomic information could help us to improve GEBV accuracies and hence genetic progress in the Brazilian Angus population. The inclusion of genotypes from American Angus sires heavily used in Brazil just marginally increased the GEBV accuracies for selection candidates.

Lay Summary

There was a desire to implement genomic selection for Angus cattle in Brazil since the technology has been proved to increase genetic gain in animal breeding programs. Single-step genomic best linear unbiased prediction (ssGBLUP), which simultaneously combines pedigree and genomic information, was used to estimate individuals' genomic breeding values (GEBV) or genetic merit. Genomic selection can accelerate genetic progress by increasing accuracy, especially in young animals without progeny. The accuracy of GEBV can also be improved by combing data from other countries to increase the reference population (i.e., genotyped and phenotyped animals) in small, genotyped populations. Thus, the main objective of this study was to evaluate the accuracy of GEBV for young Brazilian Angus (BA) bulls and heifers with ssGBLUP, including or not the genotypes from American Angus sires. The accuracies with ssGBLUP were higher than those from traditional BLUP (EBV calculated from pedigree), improving accuracies by, on average, 16% for young bulls and heifers. Including genotypes from American Angus sires heavily used in Brazil just marginally increased the GEBV accuracies for selection candidates.

Key words: beef cattle, exchange of genotypes, genomic selection, multi-country evaluation, single-step GBLUP

Abbreviations: AA, American Angus; BA, Brazilian Angus; BLUP, best linear unbiased prediction; BW, birth weight; CG, contemporary group; Corr, correlation between true and imputed genotypes; CR, call rate; EBV, estimated breeding value; G, genomic relationship matrix; GEBV, genomic estimated breeding value; LD, linkage disequilibrium; LR, linear regression; MAF, minor allele frequency; N_e, effective population size; PCG, preconditioned conjugate gradient algorithm; Perc, percentage of correctly imputed genotypes; PWG, post-weaning gain; QTL, quantitative trait loci; QC, quality control; RP, reference population; ssBR, single-step Bayesian regression; ssGBLUP, single-step genomic best linear unbiased prediction; ssGBLUP_AA, single-step genomic best linear unbiased prediction including genotypes from American Angus Association; SNP, single nucleotide polymorphism; TC, tick count; TCP, perineum tick count; TM, total maternal; WC, weaning conformation; HCW, weaning hair coat; WP, weaning precocity; WM, weaning muscling; WWG, weaning weight gain; YC, yearling conformation; YHC, yearling hair coat; YP, yearling precocity; YM, yearling muscling

Received July 6, 2021 Accepted January 12 2022.

© The Author(s) 2022. Published by Oxford University Press on behalf of the American Society of Animal Science.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Introduction

The traditional genetic improvement program for Angus cattle in Brazil based on phenotypes and pedigree was established in 1974 to improve growth, carcass quality, and more recently disease resistance (tick resistance) and adaptability (hair coat). However, genomic selection has become the new standard for genetic improvement programs because of its ability to produce larger genetic gains than traditional programs based on pedigree and phenotypes (Schaeffer, 2006; García-Ruiz et al., 2016). There is currently a desire to implement this technology for the evaluation of Angus cattle in Brazil. The success of genomic selection depends on the size of the reference population, effective population size (N_e), linkage disequilibrium (LD) between markers and quantitative trait loci (QTL), and statistical methods used to estimate genomic breeding values (GEBV) (Hayes et al., 2009). In particular, the number of genotyped Angus animals in Brazil will likely need to be augmented, as sizeable reference populations are generally required to achieve the larger genetic gains expected from genomic selection (Goddard and Hayes, 2009; VanRaden et al., 2011).

One of the most successful methods for genomic evaluation that has been implemented worldwide is the single-step genomic best linear unbiased prediction (ssGBLUP; Aguilar et al., 2010; Christensen and Lund, 2010). This method combines pedigree, phenotypes, and genotypes into one single analysis to compute GEBV for all animals in the pedigree. The ssGBLUP has computational advantages over multistep methods that lead to more accurate and less biased GEBV (Lourenco et al., 2013). Variations of the single-step method exist, and one example is the single-step Bayesian Regression (ssBR; Fernando et al., 2014). Single-step is now the method of choice for beef cattle genomic evaluations in large and small populations.

The accuracy of GEBV in small-genotyped populations may be limited regardless of method. One way to increase GEBV accuracies is to combine populations from several countries to conduct a multi-breed or single-breed multi-country evaluation (Berry et al., 2016). Lund et al. (2011) found an increase of 10% in reliability (accuracy squared) when combining small Holstein populations from France, Nordic countries, Germany, and The Netherlands. Andonov et al. (2017) simulated a large dairy cattle population with more than 20 generations and a small dairy cattle population (3 generations), and the authors reported an increase in accuracy in the small populations when all information (pedigree, phenotypes, and markers) from a larger population was added to the evaluation. However, these authors warned about the difficulties regarding extensive data sharing among countries. In populations where a small proportion of animals is genotyped and semen from foreign sires is heavily used, as is the case of Angus in Brazil, incorporation of genotypes from foreign relatives may help us to boost GEBV accuracies for Brazilian Angus selection candidates without records or progeny. This is especially true if foreign sires have progeny with phenotypes, and possibly genotypes, in the local population. Thus, the main objective of this study was to assess the feasibility of evaluating young Brazilian Angus bulls and heifers for 12 routinely recorded traits using ssGBLUP with and without genotypes from American Angus sires. The second objective was to obtain estimates of effective population size and linkage disequilibrium in the Brazilian Angus population.

Material and Methods

Animal Care and Use Committee approval was not needed because information was obtained from pre-existing databases.

Phenotype, pedigree, and genotype data

The phenotypes, pedigree, and genotypes for the Brazilian Angus cattle used in this study were provided by the Promebo breeding program (Promebo®, 2020). Animals were born from 1974 to 2018 and had phenotypes collected for birth weight (BW), weaning weight gain (WWG), weaning conformation (WC), weaning precocity (WP), weaning muscling (WM), weaning hair coat (WHC), post-weaning gain (PWG), yearling conformation (YC), yearling precocity (YP), yearling muscling (YM), yearling hair coat (YHC), and tick count (TC). Table 1 shows the number of animals with records for each trait. The conformation, precocity, and muscling data at

Table 1. Data structure and descriptive statistics for the Brazilian Angus population

Trait[1]	Number of observations	Mean ± SD	Min	Max	Number of contemporary groups	Number of sires	Number of dams	Number of animals with phenotypes and genotypes	Number of animals in validation set
BW	140,043	33.70 ± 4.93	15	60	3,218	4,306	65,844	943	334
WWG	277,661	138.40 ± 39.44	20.57	410	11,732	6,775	123,260	1018	373
WC	249,480	3.17 ± 1.08	1	5	12,306	6,346	110,832	1020	376
WP	228,596	3.23 ± 1.06	1	5	10,938	5,910	101,730	1021	377
WM	228,575	3.17 ± 1.06	1	5	10,937	5,912	101,734	1021	377
WHC	79,003	2.05 ± 0.72	1	3	4,904	2,801	39,571	795	193
PWG	180,090	148.24 ± 113.00	0.74	510	12,448	5,749	90,797	942	327
YC	174,015	3.23 ± 1.07	1	5	15,021	5,635	86,000	953	332
YP	159,408	3.31 ± 1.03	1	5	13,599	5,261	79,242	953	332
YM	159,342	3.22 ± 1.03	1	5	13,596	5,261	79,183	953	332
YHC	64,572	1,89 ± 0.71	1	3	5,823	2,792	35,644	698	103
TC	2,263	41.57 ± 36.94	0	150	30	150	1,052	921	344

[1]BW, birth weight; WWG, weaning weight gain; WC, weaning conformation; WP, weaning precocity; WM, weaning muscling; WHC, weaning hair coat; PWG, postweaning gain; YC, yearling conformation; YP, yearling precocity; YM, yearling muscling; YHC, yearling hair coat; TC, tick count.

weaning and yearling were recorded by trained technicians (Cardoso et al., 2001, 2004). The conformation score is related to the body volume of the carcass, basically considering the body length and the depth of ribs. Finishing precocity score is a measure of the animal capacity to store fat reserves, and it is used to evaluate the animal's ability to achieve a minimum market required degree of finishing for slaughter. Muscling score reflects apparent muscle mass, and individuals with more convex muscles receive higher grades. Each animal received visual scores from 1 to 5, where 5 refers to the maximum expression and 1 to the lowest expression of a trait relative to its contemporaries. Additional information can be found in Cardoso et al. (2001) and Cardoso et al. (2004). The visual score for weaning and yearling hair coats ranged from 1 to 3, where 1 refers to short, 2 to medium, and 3 to long hair coat (Reimann et al., 2018). Tick counts were performed manually by counting adult female ticks with at least 4.5 mm on one side of each animal (Wharton et al., 1970). One to three subsequent tick counts were obtained between 2012 and 2017 from five different herds. Mean age during the evaluation period was 546 ± 60 d, and the mean TC was 51 ± 53 ticks. Tick counts records were log-transformed to normalize the distribution, and a constant of 1.001 was added to the counts prior to this transformation because log10(1.0)=0.0 and null values are treated as missing by the used software (Cardoso et al., 2015).

Three different SNP panels were used to genotype 1386 animals, of which 1247 animals with GeneSeek Genomic Profiler 150k, 92 animals with GeneSeek Genomic Profiler 50k, and 17 animals with GeneSeek Genomic Profiler 77k (Neogen Agrigenomics, Lincoln, NE). Quality control of genotypes (QC) was implemented using the R/SNPStats package (Clayton, 2014). Samples with genotyping call rates (CR) < 0.90, heterozygosities three SD above or below the observed mean, mismatched sex, and duplicate records were removed. Only SNPs mapped to autosomes with CR > 0.98, minor allele frequencies (MAF) > 0.03, and with a probability of deviation from Hardy–Weinberg Equilibrium > 10^{-7} were considered in the analyses. Lastly, when SNPs were observed in the same position or genotypes were highly correlated (r > 0.98), only the SNP with the highest MAF was retained.

American Angus genotypes

Genotypes from 332 American Angus (AA) sires widely used in Brazil were retrieved from the AA Association database. These sires were genotyped with the Illumina Bovine 50k v2 SNP panel (50k_AA), which contains 54,609 SNP. The chosen American Angus sires had to have at least one progeny in the Promebo dataset. The AA provided sire genotypes, but no phenotypes were shared for any trait. Nonetheless, these sires had 19,944 progeny phenotypes (i.e., weaning weight) in the Brazilian dataset. A QC was separately conducted using the same procedure described above. After QC, 330 American Angus sires and 33,692 SNPs remained for subsequent analyses.

Genotype imputation

A joint imputation of Brazilian and American Angus genotypes was performed by combining SNPs in common among 150k, 77k, 50k, and 50k_AA. After editing and merging SNP from the four SNP panels, 61,666 SNPs and 1,292 samples from BA and 330 samples from AA animals remained for imputation. Missing genotypes were imputed using the FImpute software v2.2 (Sargolzaei et al., 2014).

To evaluate the accuracy of genotype imputation, animals were divided into reference and validation sets using the 150k SNP panel. The reference set included BA animals born before 2013 (n = 835) and the validation set comprised all BA cattle born between 2013 and 2014 (n = 281). The validation set included the younger animals, because they are the ones that would be genotyped in practice. Animals from the validation set had their 150k genotypes masked, except for SNP genotypes present in the commercially available lower density SNP panels under evaluation (50k or 77k), thus mimicking a situation in which these animals were genotyped with lower density SNP panel (Carvalheiro et al., 2014; Piccoli et al., 2014a). We pretended the animals in the validation set were genotyped with lower density SNP panel (50k or 77k), calculating statistics only for the masked genotypes.

The FImpute software uses a deterministic algorithm and family and population information. Family information is considered only when pedigree information is available. FImpute assumes that all animals are related by using the overlapping sliding windows method, and the shared haplotypes among individuals may differ in length and frequency based on relationship (Sargolzaei et al., 2014).

We performed imputations with and without pedigree information to evaluate the performance of FImpute when using family information. Imputation accuracy was assessed using the percentage of correctly imputed genotypes (Perc) and the mean correlation between true and imputed genotypes (Corr; Hickey et al., 2012).

Statistical analysis

Contemporary groups (CG) were formed by animals from the same farm, sex, year and season of birth, management group, and date of phenotypic evaluation. For continuous traits (BW, WWG, PWG, and TC), CG with less than three animals and data exceeding 3.5 SD above or below the CG mean were excluded. For visual score traits (WC, WP, WM, WHC, YC, YP, YM, and YHC), CG with less than three animals and without variability were eliminated (Table 1).

Variance components based on phenotypes and pedigree were estimated using Bayesian inference via Gibbs sampling. The GIBBS2F90 software (Misztal et al., 2014) was applied with a linear animal model for all traits, except for HCW and HCY that were analyzed using a threshold model with the THRGIBBS1F90 program (Misztal et al., 2014). The GIBBS2F90 analyses utilized a single chain of 200,000 cycles, a burn-in of 50,000 cycles, and a thinning interval of 10 cycles. The THRGIBBS1F90 analyses used a single chain of 400,000 cycles, a burn-in of 100,000 cycles, and a thinning interval of 10 cycles. Posterior estimates were obtained using the POSTGIBBSF90 program (Misztal et al., 2014). Convergence was tested using the criterion proposed by Geweke (1992), with the "boa" package in R (Smith, 2007).

The variance component analyses utilized the same bivariate animal models currently used in the Promebo genetic evaluations. The bivariate models for growth traits always used WWG as one of the traits because WWG was recorded in the largest number of animals. The models for conformation, muscling, precocity, and hair coat considered measurements at weaning and yearling as different traits. The model for TC included perineum tick counts (TCP) as a correlated

trait, although only results from whole side body tick count are described in this study.

The models included the systematic effects of CG and age of dam by sex of calf interaction and linear and quadratic covariates for animal age. Additionally, direct and maternal additive genetic, maternal permanent environmental, and residual were included as random effects for traits measured at weaning, whereas only direct additive genetic, maternal permanent environmental, and residual random effects were considered for yearling traits (model 1). The model for the trait with repeated measures (TC) included direct additive genetic, permanent environmental, and residual random effects (model 2). These bivariate animal models can be represented in matrix notation as follows:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & 0\\ 0 & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{a1} & 0\\ 0 & \mathbf{Z}_{a2}\end{bmatrix}\begin{bmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{m1} & 0\\ 0 & \mathbf{Z}_{m2}\end{bmatrix}\begin{bmatrix}\mathbf{m}_1\\ \mathbf{m}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{mpe1} & 0\\ 0 & \mathbf{Z}_{mpe2}\end{bmatrix}\begin{bmatrix}\mathbf{mpe}_1\\ \mathbf{mpe}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix}, \quad (1)$$

$$\begin{bmatrix}\mathbf{tc}_1\\ \mathbf{tc}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & 0\\ 0 & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{a1} & 0\\ 0 & \mathbf{Z}_{a2}\end{bmatrix}\begin{bmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{pe1} & 0\\ 0 & \mathbf{Z}_{pe2}\end{bmatrix}\begin{bmatrix}\mathbf{pe}_1\\ \mathbf{pe}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix}, \quad (2)$$

where

a) For model (1): $\mathbf{y}_i$ is the vector of observations for the ith trait measured at weaning or at yearling; $\boldsymbol{\beta}_i$ is the vector of systematic effects for the ith trait; $\mathbf{a}_i$ is the vector of direct additive genetic effects for the ith trait; $\mathbf{m}_i$ is the vector of maternal additive genetic effects for the ith trait; $\mathbf{mpe}_i$ is the vector of maternal permanent environmental effects for the ith trait; and $\mathbf{e}_i$ is the vector of random residuals for the ith trait. The $\mathbf{X}_i$, $\mathbf{Za}_i$, $\mathbf{Zm}_i$, and $\mathbf{Zmpe}_i$ are incidence matrices relating observations in vector $\mathbf{y}_i$ to effects in vectors $\boldsymbol{\beta}_i$, $\mathbf{a}_i$, $\mathbf{m}_i$, and $\mathbf{mpe}_i$, respectively.
b) For model (2): $\mathbf{tc}_i$ is the vector of tick counts on the whole side body (i = 1) and between the legs (i = 2); $\mathbf{pe}_i$ is the vector of random permanent environmental effects for the ith tick count; and $\mathbf{Zpe}_i$ is the incidence matrix relating tick counts in vector $\mathbf{tc}_i$ to animal permanent environmental effects in vector $\mathbf{pe}_i$. Vectors $\boldsymbol{\beta}_i$ and $\mathbf{a}_i$ as well as matrices $\mathbf{X}_i$ and $\mathbf{Z}$ai are as defined for model (1).

The following assumptions associated with the sampling distribution of the data were considered for the models 1 and 2:

$$\mathbf{y}|\beta\ ,\ \mathbf{a},\ \ \mathbf{m},\ \ \mathbf{mpe}\ \ \mathbf{R} \sim \mathrm{N}\,(\mathbf{X}\beta\ + \mathbf{Za} + \mathbf{Zm} + \mathbf{Zmpe},\ \mathbf{R})\,,$$
$$\mathbf{y}|\beta\ ,\ \mathbf{a},\ \ \mathbf{pe},\ \ \mathbf{R} \sim \mathrm{N}\,(\mathbf{X}\beta\ + \mathbf{Za} + \mathbf{Zpe},\ \mathbf{R})\,,$$

where $\boldsymbol{\beta}$, **a**, **m**, **mpe**, and **pe** are the positional parameters of observations conditional distribution; $\mathbf{R} = \mathbf{R}_0 \otimes \mathbf{I}$; $\mathbf{R}_0 = \begin{bmatrix}\sigma^2_{e1} & \sigma_{e1,e2}\\ \sigma_{e1,e2} & \sigma^2_{e2}\end{bmatrix}$, where $\mathbf{R}_0$ is the (co)variance matrix of residuals for the ith trait or for the ith tick count trait, **I** is the identity matrix, and ⊗ denotes the direct product between the matrices.

For the systematic effects, it was assumed the following prior distribution: $\beta \sim N(0, V_\beta)$, in which V_β is the non-informative diagonal variance matrix, assuming $V_\beta \rightarrow 10^{12}$. For the genetic effects, we had the following prior distributions of position parameters for the models (1) and (2):

$$\begin{bmatrix}\mathbf{a}_1\\ \mathbf{a}_2\\ \mathbf{m}_1\\ \mathbf{m}_2\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{a1} & \sigma_{a1,a2} & 0 & 0\\ & \sigma^2_{a2} & 0 & 0\\ & & \sigma^2_{m1} & 0\\ \text{Symm.} & & & 0\end{bmatrix} \otimes \boldsymbol{\Lambda}\right) \text{ and}$$
$$\begin{bmatrix}\mathbf{tc}_1\\ \mathbf{tc}_2\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{atc1} & \sigma_{tc1,tc2}\\ \sigma_{tc2,tc1} & \sigma^2_{tc2}\end{bmatrix} \otimes \boldsymbol{\Lambda}\right),$$

where σ^2_{ai} is the additive genetic variance for the ith trait; $\sigma_{ai,j}$ is the additive genetic covariance between the ith and jth traits; σ^2_{mi} is the maternal genetic variance for the ith trait, and $\boldsymbol{\Lambda}$ is the pedigree relationship matrix (**A**) in BLUP and the realized relationship matrix in ssGBLUP (**H**). The maternal effect was considered only for the traits measured at weaning, and the covariance between direct and maternal effect was considered null. For the model (2), σ^2_{tci} is the additive genetic variance for the ith tick count trait; $\sigma_{tci,j}$ is the additive genetic covariance between the ith and jth tick count traits. The other parameters were defined above:

$$\begin{bmatrix}\mathbf{mpe}_1\\ \mathbf{mpe}_2\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{mpe1} & 0\\ 0 & \sigma^2_{mpe2}\end{bmatrix} \otimes \mathbf{I}\right) \text{ and}$$
$$\begin{bmatrix}\mathbf{pe}_1\\ \mathbf{pe}_2\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2_{pe1} & \sigma_{pe1,pe2}\\ \sigma_{pe1,pe2} & \sigma^2_{pe2}\end{bmatrix} \otimes \mathbf{I}\right),$$

where σ^2_{mpei} is the maternal permanent environmental variances for the ith trait and **I** represent an identity matrix. For the model (2), σ^2_{pei} is the permanent environmental variances for the ith tick count trait; $\sigma_{pe1,pe2}$ is the covariance between permanent environmental effects.

For the (co)variances, the prior distributions were inverted Wishart that have been described previously by Gianola and Fernando (1986).

The full Bayesian threshold model for WHC and YHC contained systematic, additive, maternal genetic, and maternal permanent environmental effects. A linear model using Markov Chain Monte Carlo (MCMC) with Gibbs sampling was used for analysis of the underlying liability and it was defined as the following data distribution:

$$\mathbf{L}|\beta\ ,\ \mathbf{a},\ \mathbf{m}, \mathbf{mpe}, \mathbf{R} \sim N(\mathbf{X}\beta\ + \mathbf{Za} + \mathbf{Zm} + \mathbf{Zmpe}, \mathbf{R}),$$

where **L** is a vector of unobserved liabilities of all animals. For the systematic and genetic effects, the prior distributions were previously defined in a Bayesian framework. The model assumed existence of an underlying unobservable normal variable (Li), to analyze WHC and YHC with three scores:

$$f\,(y \mid L) = \Pi\,{}^{n}_{i=1} f\,(y \mid L_i) = \ \Pi\,{}^{n}_{i=1} I\,(L_i < t_1)\,I\,(y_i = 1) + I\,(t_1 < L_i < t_2)\,I\,(y_i = 2) + I\,(L_i > t_2)\,I\,(y_i = 3)\,,$$

where y is the observed score for WHC and YHC; t_1 and t_2 are the thresholds that categorize the three levels of response for WHC and YHC. The I is an indicator function that assumes value 1 if evaluated expression is true and 0 otherwise.

Subsequently, genomic estimated breeding values (GEBV) for all traits were computed using ssGBLUP with the same models used to estimate variance components. The inverse of the realized relationship matrix ($\mathbf{H}^{-1}$) was obtained as in Aguilar et al., (2010):

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

where $\mathbf{A}^{-1}$ is the inverse of the pedigree relationship matrix, $\mathbf{A}_{22}^{-1}$ is the inverse of the pedigree relationship matrix for genotyped animals, and $\mathbf{G}^{-1}$is the inverse of the genomic relationship matrix (G) constructed using the first method of (VanRaden, 2008), with current allele frequencies. To ensure that the inverse of G exists, it was constructed as 0.95**G** + 0.05 $\mathbf{A}_{22}$. Parameters like τ and ω that help make $\mathbf{G}^{-1}$ and $\mathbf{A}_{22}^{-1}$ compatible, avoiding inflation of GEBV, were not changed from the default values of 1. This is because including inbreeding in the calculation of all the relationship matrices aids in matrix compatibility (Tsuruta et al., 2019). The pedigree included animals up to 10 generations back from animals with phenotypes and/or genotypes. Inbreeding was considered in the construction of all relationship matrices.

Two different **G** matrices were used to investigate the impact of adding genotypes from American Angus sires to the ssGBLUP analyses: one with only genotypes for Brazilian Angus animals, and another one with genotypes for Brazilian and American Angus animals. The blending, as mentioned earlier (0.95**G** + 0.05 $\mathbf{A}_{22}$), was also used when Brazilian and American Angus genotypes were combined.

The BLUP90IOD software (Misztal et al., 2002, 2014), which is based on iteration on data and the preconditioned conjugate gradient algorithm (PCG; Strandén & Lidauer, 1999; Tsuruta et al., 2001), was used to obtain solutions for the ssGBLUP mixed model equations for all traits, except for WHC and YHC. Solutions for these two categorical traits were obtained with program THRGIBBS1F90 (Misztal et al., 2014). Official genetic evaluations for Brazilian Angus still use traditional pedigree BLUP evaluation procedures. Thus, pedigree BLUP analyses were also conducted to assess gains in prediction accuracy when using genomic information in addition to pedigree and phenotypes.

Validation procedure

The validation set was created using a range of years of birth (2013 to 2017), instead of a single year, because of limited number of genotyped animals with phenotypes (ranging from 698 to 1021) and because more influential animals usually have genotyping priority. This ensured that approximately 15% to 35% of genotyped animals with phenotypes were in the validation set (Table 1). The validation animals had no phenotypes in the partial (p) dataset for both traits in the bivariate model but at least one record in the whole (w) dataset. We used accuracy and dispersion from the linear regression (LR) method (Legarra and Reverter, 2018) to investigate the usefulness of utilizing genotypic information from both Brazilian and American Angus sires for genomic predictions. The prediction accuracy ($\widehat{acc}$), which uses the covariance between (G)EBV in the whole and partial datasets, was calculated as follows:

$$\widehat{acc} = \sqrt{\frac{cov[(G)\,EBV_w, (G)\,EBV_p]}{(1-\overline{F}_p)\,\sigma_a^2}}, \quad (3)$$

where $\overline{F}_p$ is the average inbreeding coefficient of validation animals and σ_a^2 is the additive genetic variance in the whole data set. The LR coefficient of (G)EBV$_w$ on (G)EBV$_p$ was used to assess the degree of inflation/deflation of pedigree and genomic predictions.

Validations were conducted for analyses with only genotypes from Brazilian Angus animals and analyses with genotypes from both Brazilian and American Angus animals. Additionally, to evaluate if the progeny of American Angus sires benefitted from the inclusion of genotypes from their sires, validation animals were split into progeny of Brazilian and American Angus sires before computing the aforementioned statistics.

Linkage disequilibrium and effective population size

Pairwise linkage disequilibrium (r^2) was estimated using allele and haplotype frequencies (Hill and Robertson, 1968) with R scripts provided by Badke et al. (2012). The LD values between all pairs of SNPs from all chromosomes were grouped according to pairwise physical distances into intervals of 100 kb starting from 0 up to 10 Mb. Overall mean values of r^2 for SNP pairs in each interval were obtained by calculating means across all chromosomes.

The effective population size (N_e) at a given time point was estimated based on the relationship between r^2 and N_e, rearranging the equation from Sved (1971) as follows:

$$E\left(r^2\right) = 1/\left(4cN_e + 1\right), \quad (4)$$

where c is the genetic distance between two SNPs expressed in Morgans. The N_e was estimated considering each SNP pair located within 100 Mb window of the same chromosome with physical distances converted to genetic distances between each SNP pair located in the same chromosome, assuming 1cM = 1 Mb (Qanbari et al., 2009). The N_e for past generation t (N_{et}) was estimated using the relationship between t and c (t = 1/2c; Hayes et al., 2003) and solving for N_{et} in equation (4), as follows:

$$N_{et} = \left(1r^2\right)/\left(4cr^2\right), \quad \text{for } 0.0 < r^2 < 1.0.$$

Because generations were assumed to be discrete and distances between SNP are continuous, the calculation of N_{et} for a given generation t = 1/2c in the past was assessed by selecting SNP pairs with a map distance within corresponding ranges of c values. When applying t = 1/2c, the resulting t value was rounded to the target generation. For example, r^2 of all SNP pairs with distances between 0.333 Mb (t = 1.5) and 1 Mb (t = 0.5) were selected and averaged across all chromosomes to calculate N_e at t = 1. To ensure that sufficient numbers of SNP pairs were available to obtain reliable estimates of N_{et} for each t, wider intervals around t were used to define the corresponding ranges of c, due to the inverse relationship between t and c (Sved, 1971). Values of N_{et} were obtained with increments of one generation for t between 1 and 10, of five generations for t between 15 and 100, and of 50 generations for t between 150 and 1000 (Corbin et al., 2010).

The N_e was also estimated using the rate of inbreeding (ΔF) per generation using the formula N_e = 1/ 2 ΔF (Falconer and Mackay, 1996), where ΔF per generation was estimated based on the average inbreeding between

2010 and 2017. This period corresponds to approximately 1 generation interval for Angus cattle in Brazil (Piccoli et al., 2014b). The rate of inbreeding was calculated as $\Delta F = (\overline{F}_{2017} - \overline{F}_{2010})/(1 - \overline{F}_{2010})$, where $\overline{F}_{2010}$ and $\overline{F}_{2017}$ are the mean coefficients of inbreeding for animals born in 2010 and 2017, and estimated with the algorithm of Meuwissen and Luo (1992) using the RENUMF90 software (Misztal et al., 2014).

Principal component analysis

A principal component analysis (PCA) of the genomic relationship matrix was performed to investigate the connectedness between AA and BA in the reference and validation population sets. The PCA was obtained using the preGSf90 software (Misztal et al., 2014).

Results and Discussion

Estimates of variance components and heritabilities for all evaluated traits are presented in Table 2. Heritabilities ranged from 0.11 to 0.20 for growth traits, 0.13 to 0.16 for visual scores, and 0.18 to 0.32 for adaptability traits. Genetic correlations varied from 0.23 to 0.31 for growth traits, 0.84 to 0.87 for visual scores, and 0.34 to 0.96 for adaptability traits (Table 3). These values generally agree with the ones reported in the literature for Angus cattle (Cardoso et al., 2001, 2004) and for different beef cattle breeds in Brazil (Campos et al., 2018; Reimann et al., 2018; Teixeira et al., 2018).

Accuracy of imputation

The average accuracy of imputation from the 50k and 77k SNP panels to the 150k SNP panel ranged from 0.979 to 0.986 for Corr and from 97.54% to 98.29% for Perc (Table 4). Imputation accuracy was slightly higher (0.006) from 77k to 150k than from 50k to 150k. As expected, imputation accuracy increased when the number of SNP to be imputed decreased. Piccoli et al. (2014a) observed average concordance rates of 0.943 and 0.921 among all scenarios when imputing from low density to 50k and 777k with FImpute in Hereford and Braford cattle. Ventura et al. (2014) imputed genotypes from purebred Angus and Charolais and crossbred Angus-Charolais populations with FImpute; overall, average imputation accuracies ranged from 94.20% to 97.93% for purebred animals and from 54.15% to 97.53% for crossbred animals.

Comparison of Corr and Perc values with and without pedigree in Table 4 indicated that there was no significant

Table 2. Variance components and heritabilities for the Brazilian Angus population

Trait[1]	σ^2_a	σ^2_m	σ^2_{mpe}	σ^2_{pe}	σ^2_e	$h^2_a \pm SD$	$h^2_m \pm SD$
BW	3.15 ± 0.12	0.49 ± 0.06	0.38 ± 0.06	–	11.43 ± 0.09	0.20 ± 0.005	0.03 ± 0.004
WWG	92.14 ± 3.26	31.39 ± 2.03	46.07 ± 2.00	–	399.30 ± 2.38	0.16 ± 0.005	0.06 ± 0.003
WC	0.126 ± 0.004	–	0.0345 ± 0.034	–	0.663 ± 0.004	0.15 ± 0.006	–
WP	0.147 ± 0.005	–	0.050 ± 0.012	–	0.729 ± 0.004	0.16 ± 0.005	–
WM	0.131 ± 0.005	–	0.081 ± 0.027	–	0.740 ± 0.004	0.14 ± 0.006	–
WHC	0.089 ± 0.004	–	0.005 ± 0.001	–	0.219 ± 0.003	0.28 ± 0.012	–
PWG	92.65 ± 4.66	–	–	–	734.22 ± 4.49	0.11 ± 0.006	–
YC	0.097 ± 0.005	–	–	–	0.667 ± 0.004	0.13 ± 0.006	–
YP	0.109 ± 0.004	–	–	–	0.720 ± 0.005	0.13 ± 0.006	–
YM	0.112 ± 0.005	–	–	–	0.711 ± 0.005	0.14 ± 0.006	–
YHC	0.121 ± 0.005	–	–	–	0.253 ± 0.004	0.32 ± 0.012	–
TC	0.0112 ± 0.002	–	–	0.003 ± 0.001	0.052 ± 0.002	0.18 ± 0.012	–

[1]BW, birth weight; WWG, weaning weight gain; WC, weaning conformation; WP, weaning precocity; WM, weaning muscling; WHC, weaning hair coat; PWG, postweaning gain; YC, yearling conformation; YP, yearling precocity; YM, yearling muscling; YHC, yearling hair coat; TC, tick count.

Table 3. Covariances and correlations for the Brazilian Angus population

Traits	WWG-BW	WWG-PWG	WC-YC	WP-YP	WM-YM	WHC-YHC	TC-TCP
Covariances							
$\sigma_{a1,a2}$	5.32	21.47	0.098	0.111	0.101	0.099	0.005
$\sigma_{a1,m1}$	0	0	0	0	0	0	–
$\sigma_{pe1,pe2}$	–	–	–	–	–	–	0.006
$\sigma_{e1,e2}$	0.07	−109.77	0.194	0.167	0.185	0.148	0
Correlations							
$r_{a1,a2}$	0.31	0.23	0.89	0.87	0.84	0.96	0.34
$r_{pe1,pe2}$	–	–	–	–	–	–	0.56
$r_{e1,e2}$	0.01	−0.20	0.29	0.23	0.26	0.63	0

BW, birth weight; WWG, weaning weight gain; WC, weaning conformation; WP, weaning precocity; WM, weaning muscling; WHC, weaning hair coat; PWG, postweaning gain; YC, yearling conformation; YP, yearling precocity; YM, yearling muscling; YHC, yearling hair coat; TC, tick count; TCP, perineum tick count.

Table 4. Number of total SNPs shared between the 150k SNP panel and the 77k and 50k SNP panels before and after quality control, mean and SD of imputation accuracy for different SNP panels with and without pedigree using FImpute for the combined genotype dataset from Brazilian and American Angus animals

SNP panel	Number of SNPs in common with 150k before quality control	Number of SNPs in common with 150k after quality control	With pedigree		Without pedigree	
			Correlation between imputed and observed genotypes	% Correctly imputed genotypes	Correlation between imputed and observed genotypes	% Correctly imputed genotypes
GeneSeek Genomic Profiler 150k	138,888	86,279	–	–	–	–
GeneSeek Genomic Profiler 77K	73,144	41,250	0.986 ± 0.014	98.29 ± 1.63	0.985 ± 0.013	98.24 ± 1.61
GeneSeek Genomic Profiler 50K	37,694	23,824	0.980 ± 0.018	97.59 ± 2.14	0.979 ± 0.017	97.54 ± 2.11

increase in imputation accuracy when population information was used for imputation. The proportion of unknown sires was 16% for genotyped animals and 22% for all animals; thus, these outcomes could be due to poor pedigree quality in the Brazilian Angus population.

Linkage disequilibrium

The average r^2 ± SD and distance between adjacent SNPs across all chromosomes in Brazilian Angus cattle were 0.27 ± 0.27 and 40743.68 pb. As expected, there was a rapid decrease in LD as physical distances between markers increased (Figure 1). Lu et al. (2012) estimated an r^2 value of 0.29 ± 0.30 in an Angus population from Canada with the distance between 0 and 30 kb. De Roo et al. (2008) and Villa-Angulo et al. (2009) obtained higher r^2 values for Angus cattle than those estimated in this study at smaller distances (0–5 kb) of approximately 0.6. These differences in r^2 values could be attributed to differences in sample size and number of SNPs in these studies.

Average r^2 declined rapidly with increasing distance between markers; it decreased to 50% of its initial value by ~ 50 kb (Table 5). The average LD was 0.49 ± 0.35 at distances of 1 kb, and 0.25 ± 0.26 at distances of 50 kb. Usually *Bos indicus* cattle exhibit smaller LD and rapid decay at small distances compared to *Bos taurus* cattle (Gibbs et al., 2009; Biegelmeyer et al., 2016).

Lastly, r^2 values were greater than 0.2 for 27.32% and greater than 0.3 for 19.94% of adjacent SNP markers in this study. The LD levels are an indicator of chance of success when implementing a genomic selection program in a population. According to Meuwissen et al. (2001), the average r^2 should be greater than 0.2 for genomic selection to be effective, which is the case in the Brazilian Angus population.

Effective population size

The N_e is a predictor of the effective number of independent chromosome segments that are represented in a population (Stam, 1980). Furthermore, the accuracy of genomic selection is highly associated with N_e (Daetwyler et al., 2010). The N_e based on LD was 197 in the Brazilian Angus population (Figure 2). Lu et al. (2012) computed an N_e equal to 207 in Canadian Angus cattle; however, only 597 animals were used

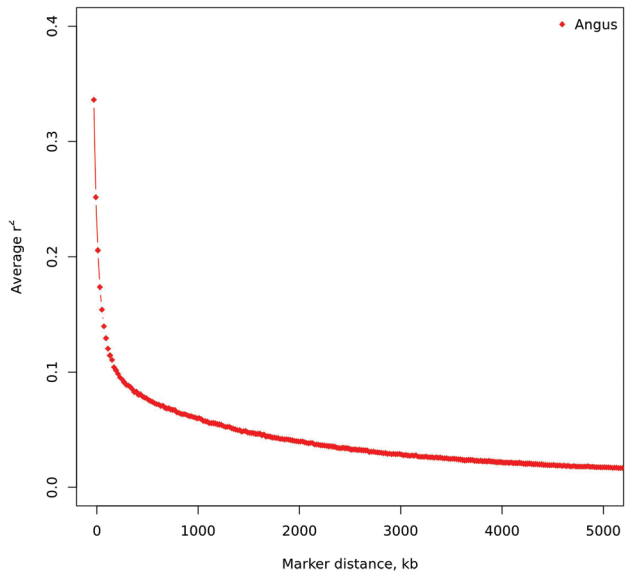

Figure 1. Decay of linkage disequilibrium (r^2) as a function of inter-marker distance in the Brazilian Angus population.

in their study. Pocrnic et al. (2016) estimated an N_e of 113 for American Angus based on the eigenvalue decomposition of **G** with over 80,000 genotyped animals. The N_e trend over past generations (Figure 2) suggests that the effective population size is decreasing, possibly due to the higher use increase of artificial insemination in Brazilian Angus in recent years (Gibbs et al., 2009). Furthermore, the use of a few purebred Angus sires in the most recent generations and high selection pressure for some traits may also be contributing to this reduction in population size. The historical N_e in the Brazilian Angus population decreased from 50 to 12 generations ago, was more stable between generations 10 and 4, and declined again two generations ago. In general, the observed trend for N_e reflects the historical process of domestication and breed formation.

The N_e based on pedigree (PED) in the Brazilian Angus population was 254. Piccoli et al. (2014b) estimated N_e based on pedigree of 185 for Hereford, 128 for Devon, and 303 for Shorthorn cattle. Differences between these N_e values and the N_e value in Brazilian Angus could be due to discrepancies

Table 5. Summary statistics for pairwise linkage disequilibrium between nonoverlapping adjacent markers in the Brazilian Angus population by intermarker distance

Intermarker distance	r^2 ± SD[1]	Meandistance (bp)	No. SNP[2]
0–1 kb	0.49 ± 0.35	674.92	270
1–5 kb	0.47 ± 0.34	3193.89	1996
5–10 kb	0.34 ± 0.30	7649.19	3594
10–20 kb	0.30 ± 0.28	15418.90	11021
20–40 kb	0.26 ± 0.26	29699.66	31310
40–60 kb	0.21 ± 0.23	50004.78	28167
60–80 kb	0.17 ± 0.20	70008.60	28753
80–100 kb	0.15 ± 0.18	90020.59	28211
0.1–0.5 Mb	0.10 ± 0.13	299479.13	61479
0.5–1 Mb	0.07 ± 0.09	749405.45	61136
1–2 Mb	0.05 ± 0.06	1497562.23	60682
2–5 Mb	0.03 ± 0.04	3479953.17	59818
5–10 Mb	0.01 ± 0.02	7468670.91	56794
Average/Total	0.27 ± 0.27	40743.68	61666

The last row contained the average r^2 ± SD and distance (pb) between nonoverlapping adjacent markers and the total number of SNP.
No. SNP[2]: Number of SNP in each inter-marker distance.
[1]Mean and standard deviation (r^2 ± SD) for pairwise linkage disequilibrium between nonoverlapping adjacent markers.

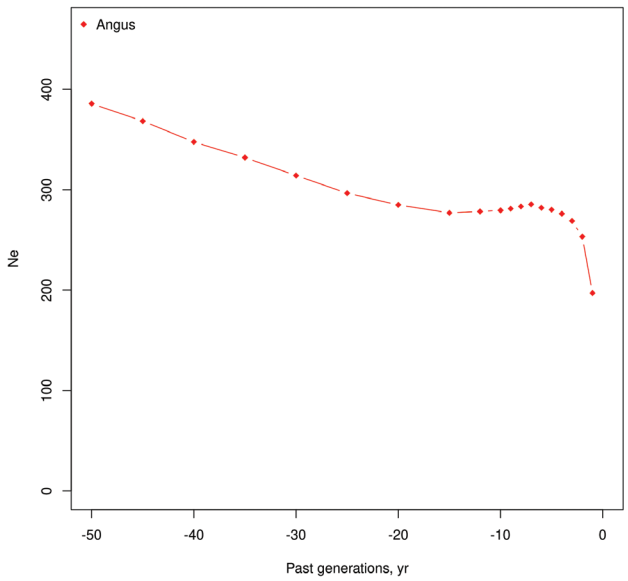

Figure 2. Estimated effective population size (N_e) as a function of past generations in the Brazilian Angus population.

in the numbers of animals used in these two studies. The N_e values based on LD and PED in this study were above the critical threshold of 50 (FAO, 1998), which indicates that a possible loss of genetic diversity is not of concern in the Brazilian Angus population at this time. The difference between the LD and PED N_e estimates is likely due to the quality and depth of pedigree as well as different methods and sources of information used in the calculations. The possible stratified sampling of animals for genotyping could also contribute to the lower estimated N_e using genotypes.

Principal component analysis

The first and second principal components (PC) of **G** were used to evaluate the connectedness and the genetic distance between the reference and validation populations of AA and BA (Figure 3). Although the percentage of variance explained was low at 2.79 and 1.24 for PC1 and PC2, respectively, the AA sample used in this study is slightly less diverse than the Brazilian one. Cardoso et al. (2020) found similar results investigating Angus populations from Brazil and Canada. The authors indicated an overall genomic similarity between Angus subpopulations (Brazil x Canada), with noticeable signals of divergent selection in genomic regions associated with the adaptation in different environments.

Genomic predictions with Brazilian Angus genotypes

The LR method was used to validate ssGBLUP models without and with genomic information from AA sires heavily used in Brazil. The LR method is an attractive validation tool because it does not require pre-adjustment of phenotypes (as with predictive ability) and can be used for any type of traits and models, including categorical traits and models with maternal effects. In summary, the LR method can be applied to any model that follows the mixed model assumptions (Bermann et al., 2021).

The prediction accuracies of GEBV using ssGBLUP from the LR validation with young animals are presented in Table 6. The GEBV accuracies ranged from 0.37 to 0.45 across all traits and the dispersion values (LR coefficients of $GEBV_w$ on $GEBV_p$) varied from 0.92 to 1.06. Prediction accuracies from ssGBLUP outperformed those from BLUP for all traits, with an average superiority of 16%. This confirmed that utilization of genomic information would bring additional benefits to the genetic evaluation of Angus cattle in Brazil. The higher prediction accuracies with ssGBLUP occurred because genomic information provided more precise relationships among individuals and helped us to estimate Mendelian sampling. These improvements could result in greater genetic gains in the Brazilian Angus population after the implementation of genomic selection.

The average prediction accuracy among growth traits was 0.40, and WWG had the greatest accuracy (0.45). Lourenco et al. (2015) estimated predictive abilities for growth traits in American Angus and obtained higher gains for ssGBLUP vs. BLUP, likely because of a substantially larger number of genotyped animals with WW records (approximately 50k) than corresponding numbers in this study. The prediction accuracy was lower for total maternal (TM) than for WWG, which was expected because of the lower heritability of WWG maternal (0.06) than of WWG direct effects (0.16). However, accuracy gains by using genomic information were similar for TM and WWG. Lourenco et al. (2013) obtained similar results in a simulated beef cattle population and concluded that prediction accuracy gains for maternal effects could be as high as for direct effects.

The GEBV prediction accuracies for visual score traits ranged from 0.44 to 0.54. The visual score traits at yearling (YC, YP, and YM) had the highest realized accuracies among all visual scores (0.53 to 0.54). GEBV prediction accuracies for visual score traits in Hereford and Braford cattle computed with various single-step and multi-step methods ranged from 0.09 to 0.61 (Piccoli et al., 2020). The GEBV prediction

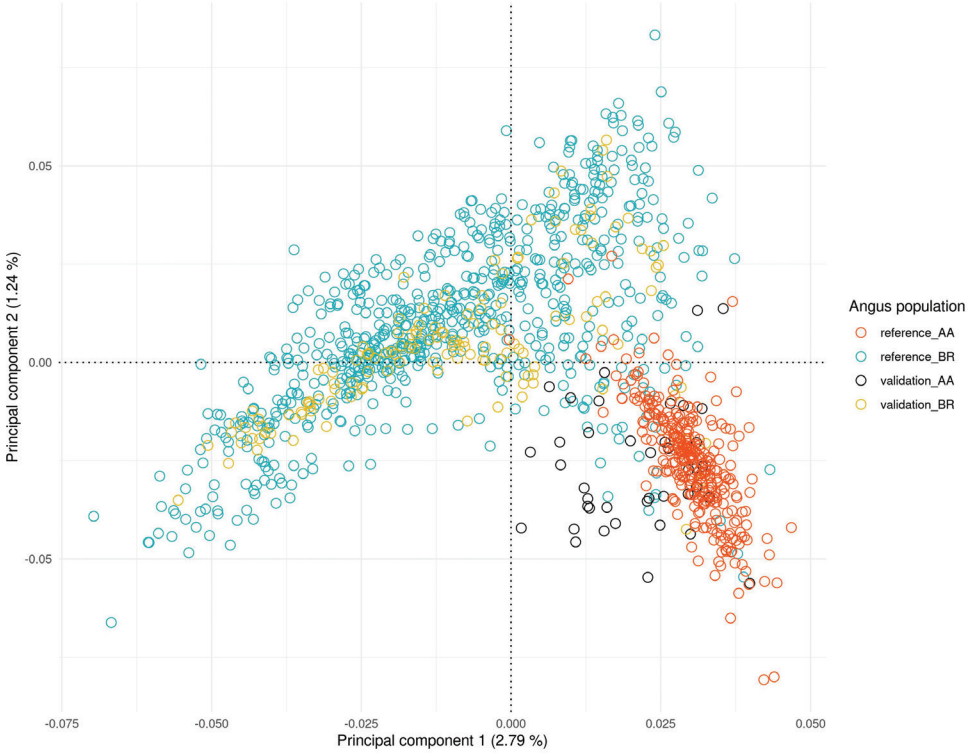

Figure 3. First and second principal components of the genomic relationship matrix for the reference and validation Brazilian and American Angus population.

accuracies for adaptability traits (WHC, YHC, and TR) fluctuated between 0.37 and 0.49. The YHC and TC had the greatest increases in accuracy when moving from BLUP to ssGBLUP among all evaluated traits (32% for YHC and 28% for TC). Cardoso et al. (2015) reported ssGBLUP prediction accuracies for TC of 0.48 in Hereford and 0.56 in Braford with *k*-means and random clustering. The LR coefficient of (G)EBV$_w$ on (G)EBV$_p$ was generally closer to 1 with ssGBLUP than with BLUP, indicating less inflated/deflated predictions. The dispersion across all traits ranged from 0.92 to 1.07 for ssGBLUP and from 0.91 to 1.13 for BLUP. Inflation/deflation may be a problem when predictions of genotyped and non-genotyped animals need to be compared for ranking and selection decisions. When the LR coefficient of GEBV$_w$ on GEBV$_p$ is less than 1, predictions are inflated and extreme GEBV may exist, favoring genotyped individuals (Cardoso et al., 2015).

Genomic predictions with Brazilian and American Angus genotypes

The reliability of genomic predictions increases as the size of the reference population (RP) increases (Goddard and Hayes, 2009). One way to increase RP is to combine information within a country with genotype data from foreign animals that are used locally (Andonov et al., 2017). When the number of progeny with records from foreign sires is small in the local population, EBV or GEBV from a foreign evaluation can be used as priors for BLUP and ssGBLUP (Legarra et al., 2007; Vandenplas et al., 2014). Additionally, all available information (phenotypes, pedigree, and genotypes) can be used if there are cooperative agreements among countries; however, this is seldom the case (Andonov et al., 2017).

Genotypes from 330 AA sires heavily used in Brazil were available for this study. These AA sires represented an increase of 35% in the size of the RP because of the low number of genotyped Angus cattle in Brazil. Thus, an increase in ssGBLUP prediction accuracy could possibly be expected after adding the American Angus genotypes (ssGBLUP_AA) relative to using only Brazilian Angus genotypes (ssGBLUP). Table 6 shows LR prediction accuracies for BLUP, ssGBLUP, and ssGBLUP_AA for all validation animals and for the progeny of BA and AA sires. On average, the prediction accuracy for all young animals was 0.47 with ssGBLUP_AA and 0.46 with ssGBLUP. This marginal increase corresponds to only 2% of the prediction accuracy with ssGBLUP. The highest increase in prediction accuracy among all traits was for YHC (0.49 to 0.51), an increase of 4% after the addition of genotypes from AA sires. When validation animals were split into progeny of BA and AA sires, prediction accuracies for the progeny of AA sires were, on average, 0.11 lower than prediction accuracies for the progeny of Brazilian Angus sires. This possibly happened because progeny size and number of records for AA sires were smaller than for Brazilian Angus sires. American Angus sires had a total of 40 progeny in the validation set, whereas Brazilian Angus sires had 178; those progeny had phenotypes in the whole dataset. Another crucial point to the small increases in accuracies is that only a small number of genotypes from AA bulls were added; the number of records stayed the same. Overall, the GEBV of validation progeny from AA sires were more inflated/deflated than those from Brazilian Angus sires. However, the addition of genotypes from AA sires did not produce large dispersion changes in the three groups of validation animals. The average LR coefficient differences between ssGBLUP_AA and ssGBLUP were −0.005, 0.013, and

Table 6. Linear regression accuracy and dispersion of (G)EBV for all validation animals, validation animals that are progeny of Brazilian and American Angus sires

Trait[1]	Method[2]	Progeny of Brazilian Angus sires		Progeny of American Angus sires		All validation animals	
		Accuracy	Dispersion	Accuracy	Dispersion	Accuracy	Dispersion
BW	ssGBLUP	0.40	1.03	0.32	1.34	0.39	0.92
	ssGBLUP_AA	0.41	0.95	0.34	1.33	0.41	0.93
	BLUP	0.36	0.87	0.30	1.12	0.35	0.91
WWG	ssGBLUP	0.46	1.05	0.36	0.97	0.45	1.01
	ssGBLUP_AA	0.47	1.06	0.37	0.99	0.46	1.02
	BLUP	0.41	1.01	0.32	0.89	0.39	1.02
TM	ssGBLUP	0.35	1.15	0.26	0.88	0.37	1.02
	ssGBLUP_AA	0.35	1.09	0.27	0.89	0.37	1.02
	BLUP	0.31	1.19	0.21	0.84	0.30	1.13
PWG	ssGBLUP	0.37	0.92	0.38	0.98	0.42	0.97
	ssGBLUP_AA	0.38	0.94	0.40	1.00	0.42	0.97
	BLUP	0.33	0.91	0.32	0.95	0.36	0.96
WC	ssGBLUP	0.51	1.04	0.39	1.07	0.49	1.06
	ssGBLUP_AA	0.53	1.05	0.39	1.08	0.50	1.07
	BLUP	0.46	1.07	0.36	1.00	0.42	1.09
WP	ssGBLUP	0.49	1.03	0.31	0.84	0.44	1.02
	ssGBLUP_AA	0.51	1.03	0.35	0.88	0.46	1.02
	BLUP	0.44	1.05	0.29	0.78	0.39	1.05
WM	ssGBLUP	0.47	0.94	0.26	0.97	0.46	0.96
	ssGBLUP_AA	0.48	0.94	0.26	0.98	0.47	0.97
	BLUP	0.45	0.89	0.21	1.01	0.41	0.94
WHC	ssGBLUP	0.48	1.02	0.42	0.86	0.49	0.97
	ssGBLUP_AA	0.50	1.03	0.42	0.92	0.50	0.98
	BLUP	0.46	0.99	0.40	0.96	0.45	0.96
YC	ssGBLUP	0.53	1.05	0.36	1.12	0.53	1.07
	ssGBLUP_AA	0.54	1.06	0.36	1.30	0.53	1.08
	BLUP	0.50	1.05	0.33	1.10	0.49	1.10
YP	ssGBLUP	0.57	1.04	0.41	1.04	0.54	1.03
	ssGBLUP_AA	0.57	1.03	0.43	1.06	0.54	1.02
	BLUP	0.50	1.00	0.32	0.99	0.45	1.02
YM	ssGBLUP	0.52	0.95	0.33	0.97	0.53	0.96
	ssGBLUP_AA	0.53	0.95	0.33	1.00	0.53	0.96
	BLUP	0.50	0.86	0.29	1.04	0.47	0.92
YHC	ssGBLUP	0.45	1.18	0.38	0.66	0.48	1.01
	ssGBLUP_AA	0.48	1.26	0.40	0.68	0.51	0.95
	BLUP	0.40	1.19	0.31	1.08	0.37	0.95
TC	ssGBLUP	0.38	1.09	0.36	1.38	0.37	0.97
	ssGBLUP_AA	0.40	1.04	0.39	1.14	0.38	0.98
	BLUP	0.30	1.27	0.26	1.13	0.29	1.04

[1]BW, birth weight; WWG, weaning weight gain; TM, total maternal; WC, weaning conformation; WP, weaning precocity; WM, weaning muscling; WHC, weaning hair coat; PWG, postweaning gain; YC, yearling conformation; YP, yearling precocity; YM, yearling muscling; YHC, yearling hair coat; TC, tick count.
[2]ssGBLUP, single-step GBLUP without genotypes from the American Angus Association; ssGBLUP_AA, single-step GBLUP with genotypes from the American Angus Association; BLUP, pedigree BLUP.

0 for the validation progeny of BA sires, AA sires, and all validation animals.

The expected prediction accuracy gains resulting from the addition of AA sire genotypes to RP were not realized, and only a marginal increase in prediction accuracy was observed. This outcome may change when these AA sires gather more progeny with phenotypes and genotypes in Brazilian herds. Lund et al. (2011) obtained an average increase of 10% in reliability (prediction accuracy squared) when combining small Holstein populations from France, Nordic countries, Germany, and The Netherlands. The increase in reliability for some traits was up to 19% relative to using national reference data alone. Andonov et al. (2017) reported an increase in prediction accuracy in small-simulated populations

when information from a larger population was added to the evaluation. These authors observed a beneficial increase in prediction accuracies for GEBV relative to EBV in the small population when phenotypes, pedigree, and genotypes from the small and larger populations were combined. The level of relatedness between animals in RP and selection candidates is an important aspect that influences the accuracy of genomic predictions (Pszczola et al., 2012). Lourenco et al. (2015) pointed out that genotyping strategies in breeding populations should consider important animals with substantial information (high accuracy animals) and selection candidates related to them. Thus, the extent of the benefit of incorporating genotypes from foreign animals into local genomic evaluations is determined by the level of connection among them (Berry et al., 2016) and the shared amount of information.

An alternative to increase the RP, and consequently the accuracy of genomic predictions in small Angus populations, would be an across-country genomic evaluation where not only genotypes but also pedigree and phenotypes are shared. Several countries in South America use semen from American Angus sires; thus, a Pan-American genomic evaluation that includes data from the United States, Brazil, Argentina, Uruguay, and other countries would benefit a large group of breeders. However, this Pan-American genomic evaluation would require sharing available data across countries, which could be a sensitive issue. A Pan-American evaluation was established for Hereford cattle in 2016 (Berry et al., 2016) for countries sharing Hereford germplasm. The steady decrease of genotyping costs would help countries with smaller herds to genotype more animals, especially progeny of foreign sires, to increase connectedness among countries. Although the benefit of a combined evaluation seems to be greater for countries with small herds, countries with large datasets would also benefit by receiving information from sires that are more heavily used internationally than nationally. This information would help identify sires with high genetic potential in various countries and low levels of inbreeding (Andonov et al., 2017). Another interesting area to explore with genomic information would be the genotype by environment interaction (GxE), which would allow the selection of the best genotypes (i.e., young sires) to produce under specific conditions and more robust animals for the Brazilian and American environments. However, that was not possible in this study because AA phenotypes were unavailable.

Conclusions

The level of LD estimated in the Brazilian Angus population indicates that a 60k SNP panel would be a suitable tool to increase genomic prediction accuracies. The N_e estimated based on pedigree or as a function of past generations using genomic information was sufficiently large to maintain desirable levels of genetic diversity and to increase selection responses for economically important traits. The GEBV from ssGBLUP had higher prediction accuracies than EBV from pedigree BLUP for all evaluated traits, indicating that the implementation of genomic selection in the Brazilian Angus population would be beneficial. The inclusion of genotypes from American Angus sires heavily used in Brazil just marginally increased the GEBV accuracies for selection candidates. A future investigation that shares genotypes, phenotypes, and pedigree across countries could help visualize the benefits of combining Brazilian (or South American) and American Angus data.

Acknowledgments

We gratefully acknowledge the American Angus Association for providing access to the genomic data and Mauricio A. Elzo for editing the manuscript. We also acknowledge funding for the collection of Brazilian Angus data from Conselho Nacional de Desenvolvimento Científico e Tecnológico grants 478780/2013-3, 474829/2013-8, and 305102/2018-4, and from Empresa Brasileira de Pesquisa Agropecuária grants 02.13.10.002 and 30.20.90.018. Gabriel Soares Campos' postdoctoral research at the University of Georgia was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior grant PDE 88881.171892/2018-01.

Conflict of Interest Statement

The authors declare no real or perceived conflicts of interest.

Literature Cited

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743–752. doi:10.3168/jds.2009-2730

Andonov, S., D. Lourenco, B. Fragomeni, Y. Masuda, I. Pocrnic, S. Tsuruta, and I. Misztal. 2017. Accuracy of breeding values in small genotyped populations using different sources of external information—a simulation study. *J. Dairy Sci.* 100:395–401. doi:10.3168/jds.2016-11335

Badke, Y. M., R. O. Bates, C. W. Ernst, C. Schwab, and J. P. Steibel. 2012. Estimation of linkage disequilibrium in four US pig breeds. *BMC Genomics*. 13:24. doi:10.1186/1471-2164-13-24

Bermann, M., A. Legarra, M. K. Hollifield, Y. Masuda, D. Lourenco, and I. Misztal. 2021. Validation of single-step GBLUP genomic predictions from threshold models using the linear regression method: an application in chicken mortality. *J. Anim. Breed. Genet.* 138:4–13. doi:10.1111/jbg.12507

Berry, D. P., J. F. Garcia, and D. J. Garrick. 2016. Development and implementation of genomic predictions in beef cattle. *Anim. Front.* 6:32–38. doi:10.2527/af.2016-0005. Available from https://academic.oup.com/af/article/6/1/32/4638791

Biegelmeyer, P., C. C. Gulias-Gomes, A. R. Caetano, J. P. Steibel, and F. F. Cardoso. 2016. Linkage disequilibrium, persistence of phase and effective population size estimates in Hereford and Braford cattle. *BMC Genet.* 17:32. doi:10.1186/s12863-016-0339-8. Available from http://www.biomedcentral.com/1471-2156/17/32

Campos, G. S., F. A. Reimann, L. L. Cardoso, C. E. R. Ferreira, V. S. Junqueira, P. I. Schmidt, J. Braccini Neto, M. J. I. Yokoo, B. P. Sollero, A. A. Boligon, et al. 2018. Genomic prediction using different estimation methodology, blending and cross-validation techniques for growth traits and visual scores in Hereford and Braford cattle. *J. Anim. Sci.* 96:2579–2595. doi:10.1093/jas/sky175. Available from https://academic.oup.com/jas/advance-article/doi/10.1093/jas/sky175/4993326.

Cardoso, F. F., R. A. Cardellino, and L. T. Campos. 2001. Componentes de (co)variância e parâmetros genéticos para caracteres produtivos à desmama de bezerros Angus criados no Estado do Rio Grande do Sul. *Rev. Bras. Zootec.* 30:41–48. doi:10.1590/s1516-35982001000100008

Cardoso, F. F., R. A. Cardellino, and L. T. Campos. 2004. Componentes de (co)variância e parâmetros genéticos de caracteres pós-desmama em bovinos da raça Angus. *Rev. Bras. Zootec.* 33:313–319. doi:10.1590/S1516-35982004000200006. Available from http://www.scielo.br/j/rbz/a/yxnMvxkWyLS3jfmvsGTndGF/?lang=pt

Cardoso, F. F., C. C. G. Gomes, B. P. Sollero, M. M. Oliveira, V. M. Roso, M. L. Piccoli, R. H. Higa, M. J. Yokoo, A. R. Caetano, and

I. Aguilar. 2015. Genomic prediction for tick resistance in Braford and Hereford cattle1. *J. Anim. Sci*. 93:2693–2705. doi:10.2527/jas.2014-8832. Available from https://academic.oup.com/jas/article/93/6/2693/4703464

Cardoso, D. F., G. A. Fernandes Júnior, D. C. B. Scalez, A. A. C. Alves, A. F. B. Magalhães, T. Bresolin, R. V. Ventura, C. Li, M. C. de Sena Oliveira, L. R. Porto-Neto, et al. 2020. Uncovering sub-structure and genomic profiles in across-countries subpopulations of angus cattle. *Sci. Rep*. 10:1–11. doi:10.1038/s41598-020-65565-1. Available from https://www.nature.com/articles/s41598-020-65565-1

Carvalheiro, R., S. A. Boison, H. H. R. Neves, M. Sargolzaei, F. S. Schenkel, Y. T. Utsunomiya, A. M. P. O'Brien, J. Sölkner, J. C. McEwan, C. P. Van Tassell, et al. 2014. Accuracy of genotype imputation in Nelore cattle. *Genet. Sel. Evol*. 46:69. doi:10.1186/s12711-014-0069-1. Available from http://gsejournal.biomedcentral.com/articles/10.1186/s12711-014-0069-1

Christensen, O. F., and M. S. Lund. 2010. Genomic prediction when some animals are not genotyped. *Genet. Sel. Evol*. 42:2. doi:10.1186/1297-9686-42-2. Available from https://gsejournal.biomedcentral.com/articles/10.1186/1297-9686-42-2

Clayton, D. 2014. Package "snpStats" Title SnpMatrix and XSnpMatrix classes and methods. Available from https://git.bioconductor.org/packages/snpStats

Corbin, L. J., S. C. Blott, J. E. Swinburne, M. Vaudin, S. C. Bishop, and J. A. Woolliams. 2010. Linkage disequilibrium and historical effective population size in the thoroughbred horse. *Anim. Genet*. 41:8–15. doi:10.1111/j.1365-2052.2010.02092.x. Available from http://doi.wiley.com/10.1111/j.1365-2052.2010.02092.x

Daetwyler, H. D., R. Pong-Wong, B. Villanueva, and J. A. Woolliams. 2010. The impact of genetic architecture on genome-wide evaluation methods. *Genetics*. 185:1021–1031. doi:10.1534/genetics.110.116855

De Roo, A. P. W., B. J. Hayes, R. J. Spelman, and M. E. Goddard. 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. *Genetics*. 179:1503–1512. doi:10.1534/genetics.107.084301

Falconer, D. S., and T. F. C. Mackay. 1996. *Introduction to quantitative genetics*. 4th ed. UK: Lo. Harlow.

Fernando, R. L., J. C. Dekkers, and D. J. Garrick. 2014. Selection evolution a class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. *Genet. Sel. Evol*. 46:50. doi:10.1186/1297-9686-46-50

Food and Agriculture Organization of the United Nations (FAO). 1998. *Secondary guidelines for development of national farm animal genetic resources management plans: management of small populations at risk*. Rome, Italy: FAO.

García-Ruiz, A., J. B. Cole, P. M. VanRaden, G. R. Wiggans, F. J. Ruiz-López, and C. P. Van Tassell. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proc. Natl. Acad. Sci*. 113:E3995–E4004. doi:10.1073/PNAS.1519061113. Available from https://www.pnas.org/content/113/28/E3995

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: A. P. D. and A. F. M. S. J. M. Bernardo, J. O. Berger, editors. *Bayesian statistics*. New York: Oxford University Press. p. 625–631.

Gianola, D., and R. L. Fernando. 1986. Bayesian methods in animal breeding theory. *J. Anim. Sci*. 63:217–244. doi:10.2527/JAS1986.631217X. Available from https://academic.oup.com/jas/article/63/1/217/4658748

Gibbs, R. A., J. F. Taylor, C. P. Van Tassell, W. Barendse, K. A. Eversole, C. A. Gill, R. D. Green, D. L. Hamernik, S. M. Kappes, S. Lien, et al. 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*. 324:528–532. doi:10.1126/science.1167936

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. doi:10.1038/nrg2575. Available from https://www.nature.com/articles/nrg2575

Hayes, B. J., P. J. Bowman, A. J. Chamberlain., and M. E. Goddard. 2009. Invited review: genomic selection in dairy cattle: progress and challenges. *J. Dairy Sci*. 92:433–443. doi:10.3168/jds.2008-1646

Hayes, B. J., P. M. Visscher, H. C. McPartlan, and M. E. Goddard. 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res*. 13:635–643. doi:10.1101/gr.387103

Hickey, J. M., J. Crossa, R. Babu, and G. de los Campos. 2012. Factors affecting the accuracy of genotype imputation in populations from several maize breeding programs. *Crop Sci*. 52:654–663. doi:10.2135/cropsci2011.07.0358

Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. *Theoret. Appl. Genet*. 38:226–231. doi:10.1007/BF01245622

Legarra, A., J. K. Bertrand, T. Strabel, R. L. Sapp, J. P. Sánchez, and I. Misztal. 2007. Multi-breed genetic evaluation in a Gelbvieh population. *J. Anim. Breed. Genet*. 124:286–295. doi:10.1111/j.1439-0388.2007.00671.x. Available from http://doi.wiley.com/10.1111/j.1439-0388.2007.00671.x

Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genet. Sel. Evol*. 50:53. doi:10.1186/s12711-018-0426-6. Available from https://gsejournal.biomedcentral.com/articles/10.1186/s12711-018-0426-6

Lourenco, D. A. L., I. Misztal, H. Wang, I. Aguilar, S. Tsuruta, and J. K. Bertrand. 2013. Prediction accuracy for a simulated maternally affected trait of beef cattle using different genomic evaluation models. *J. Anim. Sci*. 91:4090–4098. doi:10.2527/jas.2012-5826. Available from https://academic.oup.com/jas/article/91/9/4090/4717245

Lourenco, D. A. L., S. Tsuruta, B. O. Fragomeni, Y. Masuda, I. Aguilar, A. Legarra, J. K. Bertrand, T. S. Amen, L. Wang, D. W. Moser, et al. 2015. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *J. Anim. Sci*. 93:2653–2662. doi:10.2527/jas.2014-8836. Available from https://academic.oup.com/jas/article/93/6/2653/4703431

Lu, D., M. Sargolzaei, M. Kelly, C. Li, G. Vander Voort, Z. Wang, G. Plastow, S. Moore, and S. P. Miller. 2012. Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Front. Genet*. 3:152. doi:10.3389/fgene.2012.00152. Available fromhttp://journal.frontiersin.org/article/10.3389/fgene.2012.00152/abstract

Lund, M. S., A. Pw De Roos, A. G. De Vries, T. Druet, V. Ducrocq, S. Fritz, F. Guillaume, B. Guldbrandtsen, Z. Liu, R. Reents, et al. 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. Available from http://www.gsejournal.org/content/43/1/43

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. doi:10.1093/genetics/157.4.1819

Meuwissen, T. H. E., and Z. Luo. 1992. Computing inbreeding coefficients in large populations. *Genet. Sel. Evol*. 24:305–313. doi:10.1186/1297-9686-24-4-305. Available from http://www.gsejournal.org/content/24/4/305

Misztal, I., S. Tsuruta, D. A. L. Lourenco, I. Aguilar, Legarra A., and Vitezica Z. 2014. Manual for BLUPF90 family of programs. Available from http://nce.ads.uga.edu/wiki/doku.php

Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). In: Proc. 7th World Genet. Appl. Livest. Prod. Montpellier, France. p. 21–22. Available from http://www.ozemail.com.au/~milleraj

Piccoli, M. L., J. Braccini, F. F. Cardoso, M. Sargolzaei, S. G. Larmer, and F. S. Schenkel. 2014a. Accuracy of genome-wide imputation in Braford and Hereford beef cattle. *BMC Genet*. 15:157. doi:10.1186/s12863-014-0157-9. Available from https://academic.oup.com/jas/article/92/5/1920/4703370

Piccoli, M. L., J. Braccini Neto, F. V. Brito, L. T. Campos, C. D. Bértoli, G. S. Campos, J. A. Cobuci, C. M. McManus, J. O. J. Barcellos, and L. T. Gama. 2014b. Origins and genetic diversity of British

cattle breeds in Brazil assessed by pedigree analyses. *J. Anim. Sci.* 92:1920–1930. doi:10.2527/jas.2013-7283. Available from https://academic.oup.com/jas/article/92/5/1920/4703370

Piccoli, M. L., L. F. Brito, J. Braccini, H. R. Oliveira, F. F. Cardoso, V. M. Roso, M. Sargolzaei, and F. S. Schenkel. 2020. Comparison of genomic prediction methods for evaluation of adaptation and productive efficiency traits in Braford and Hereford cattle. *Livest. Sci.* 231:103864. doi:10.1016/j.livsci.2019.103864

Pocrnic, I., D. A. L. Lourenco, Y. Masuda, and I. Misztal. 2016. Dimensionality of genomic information and performance of the Algorithm for Proven and Young for different livestock species. *Genet. Sel. Evol.* 48:82. doi:10.1186/s12711-016-0261-6. Available from http://gsejournal.biomedcentral.com/articles/10.1186/s12711-016-0261-6

Promebo®. 2020. Programa de melhoramento de bovinos de carne. Manual do usuário: Associação Nacional de Criadores "Herd Book Collares". Available from http://www.herdbook.org.br/site/arquivosindividuais/index/promebo_manual_criador

Pszczola, M., T. Strabel, H. A. Mulder, and M. P. L. Calus 2012. Reliability of direct genomic values for animals with different relationships within and to the reference population. *J. Dairy Sci.* 95:389–400. doi:10.3168/jds.2011-4338

Qanbari, S., E. C. G. Pimentel, J. Tetens, G. Thaller, P. Lichtner, A. R. Sharifi, and H. Simianer. 2009. The pattern of linkage disequilibrium in German Holstein cattle. *Anim. Genet.* 41:346–356. doi:10.1111/j.1365-2052.2009.02011.x. Available from http://doi.wiley.com/10.1111/j.1365-2052.2009.02011.x

Reimann, F. A., A. A. Boligon, G. S. Campos, L. L. Cardoso, V. S. Junqueira, and F. F. Cardoso. 2018. Genetic parameters and accuracy of traditional and genomic breeding values for eye pigmentation, hair coat and breed standard in Hereford and Braford cattle. *Livest. Sci.* 213:44–50. doi:10.1016/j.livsci.2018.04.007

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* 15:478. doi:10.1186/1471-2164-15-478

Schaeffer, L. R. 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218–223. doi:10.1111/j.1439-0388.2006.00595.x. Available from https://onlinelibrary.wiley.com/doi/full/10.1111/j.1439-0388.2006.00595.x

Smith, B. J. 2007. boa: an R package for MCMC output convergence assessment and posterior inference. *J. Stat. Softw.* 21:1–37. doi:10.18637/jss.v021.i11

Stam, P. 1980. The distribution of the fraction of the genome identical by descent in finite random mating populations. *Genet. Res.* 35:131–155. doi:10.1017/S0016672300014002

Strandén, I., and M. Lidauer. 1999. Solving large mixed linear models using preconditioned conjugate gradient iteration. *J. Dairy Sci.* 82:2779–2787. doi:10.3168/jds.S0022-0302(99)75535-9

Sved, J. A. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theor. Popul. Biol.* 2:125–141. doi:10.1016/0040-5809(71)90011-6

Teixeira, B. B. M., M. D. MacNeil, R. F. da Costa, N. J. L. Dionello, M. J. Yokoo, and F. F. Cardoso. 2018. Genetic parameters and trends for traits of the Hereford and Braford breeds in Brazil. *Livest. Sci.* 208:60–66. doi:10.1016/J.LIVSCI.2017.12.008

Tsuruta, S., D. A. L. Lourenco, Y. Masuda, I. Misztal, and T. J. Lawlor. 2019. Controlling bias in genomic breeding values for young genotyped bulls. *J. Dairy. Sci.* 102:9956–9970. doi:10.3168/jds.2019-16789

Tsuruta, S., I. Misztal, and I. Strandén. 2001. Use of the preconditioned conjugate gradient algorithm as a generic solver for mixed-model equations in animal breeding applications. *J. Anim. Sci.* 79:1166. doi:10.2527/2001.7951166x. Available from https://academic.oup.com/jas/article/79/5/1166-1172/4625922

Vandenplas, J., F. G. Colinet, and N. Gengler. 2014. Unified method to integrate and blend several, potentially related, sources of information for genetic evaluation. *Genet. Sel. Evol.* 46:59. doi:10.1186/s12711-014-0059-3. Available from http://gsejournal.biomedcentral.com/articles/10.1186/s12711-014-0059-3

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. doi:10.3168/jds.2007-0980

VanRaden, P. M., J. R. O'Connell, G. R. Wiggans, and K. A. Weigel. 2011. Genomic evaluations with many more genotypes. *Genet. Sel. Evol.* 43:1–11. doi:10.1186/1297-9686-43-10. Available from https://link.springer.com/articles/10.1186/1297-9686-43-10

Ventura, R. V., D. Lu, F. S. Schenkel, Z. Wang, C. Li, and S. P. Miller. 2014. Impact of reference population on accuracy of imputation from 6K to 50K single nucleotide polymorphism chips in purebred and crossbreed beef cattle. *J. Anim. Sci.* 92:1433–1444. doi:10.2527/jas.2013-6638. Available from https://academic.oup.com/jas/article/92/4/1433/4703266

Villa-Angulo, R., L. K. Matukumalli, C. A. Gill, J. Choi, C. P. Van Tassell, and J. J. Grefenstette. 2009. High-resolution haplotype block structure in the cattle genome. *BMC Genet.* 10:19. doi:10.1186/1471-2156-10-19. Available from https://bmcgenet.biomedcentral.com/articles/10.1186/1471-2156-10-19

Wharton, R., K. UtechH. Turner. 1970. Resistance to the cattle tick, *Boophilus microplus* in a herd of Australian Illawarra Shorthorn cattle: its assessment and heritability. *Aust. J. Agric. Res.* 21:163–181. doi:10.1071/AR9700163. Available from https://www.publish.csiro.au/ar/ar9700163