Received: 6 August 2019 | Revised: 3 November 2019 | Accepted: 5 November 2019

DOI: 10.1111/jbg.12458

ORIGINAL ARTICLE

Journal of Animal Breeding and Genetics WILEY

Tag-SNP selection using Bayesian genomewide association study for growth traits in Hereford and Braford cattle

Gabriel Soares Campos[1] | **Bruna Pena Sollero**[2] | **Fernando Antonio Reimann**[1] | **Vinicius Silva Junqueira**[3] | **Leandro Lunardini Cardoso**[1,2] | **Marcos Jun Iti Yokoo**[2] | **Arione Augusti Boligon**[1] | **José Braccini**[4] | **Fernando Flores Cardoso**[1,2]

[1]Departamento de Zootecnia, Universidade Federal de Pelotas, Pelotas, Brazil

[2]Embrapa Pecuária Sul, Bagé, Brazil

[3]Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa, Brazil

[4]Departamento de Zootecnia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Correspondence
Gabriel Soares Campos, Universidade Federal de Pelotas, Departamento de Zootecnia, Pelotas, RS, 96010-900, Brazil.
Email: gabrielsoarescampos@hotmail.com

Funding information
Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number 305102/2018-4; Empresa Brasileira de Pesquisa Agropecuária, Grant/Award Number 12.13.14.014.00.

Abstract

The aim of this study was to perform a Bayesian genomewide association study (GWAS) to identify genomic regions associated with growth traits in Hereford and Braford cattle, and to select Tag-SNPs to represent these regions in low-density panels useful for genomic predictions. In addition, we propose candidate genes through functional enrichment analysis associated with growth traits using Medical Subject Headings (MeSH). Phenotypic data from 126,290 animals and genotypes for 131 sires and 3,545 animals were used. The Tag-SNPs were selected with BayesB ($\pi = 0.995$) method to compose low-density panels. The number of Tag-single nucleotide polymorphism (SNP) ranged between 79 and 103 SNP for the growth traits at weaning and between 78 and 100 SNP for the yearling growth traits. The average proportion of variance explained by Tag-SNP with BayesA was 0.29, 0.23, 0.32 and 0.19 for birthweight (BW), weaning weight (WW205), yearling weight (YW550) and post-weaning gain (PWG345), respectively. For Tag-SNP with BayesA method accuracy values ranged from 0.13 to 0.30 for *k*-means and from 0.30 to 0.65 for random clustering of animals to compose reference and validation groups. Although genomic prediction accuracies were higher with the full marker panel, predictions with low-density panels retained on average 76% of the accuracy obtained with BayesB with full markers for growth traits. The MeSH analysis was able to translate genomic information providing biological meanings of more specific gene products related to the growth traits. The proposed Tag-SNP panels may be useful for future fine mapping studies and for lower-cost commercial genomic prediction applications.

KEYWORDS

beef cattle, genomic prediction, GWAS, low-density panel

1 | INTRODUCTION

The selection for growth traits using phenotypes and pedigree data with best linear unbiased prediction (BLUP) (Henderson, 1975) has been effective over the years. However, this can be improved if DNA polymorphisms affecting growth traits in beef cattle are determined (Snelling et al., 2010). Genome-Wide Association Studies (GWAS) enable the identification of SNPs associated with the quantitative trait loci (QTL) of interest and/or genes related to the phenotypic expression of the trait (Utsunomiya et al., 2013). The assumption in GWAS is that association arise because at least one single nucleotide

polymorphism (SNP) is in LD with a causative mutation, responsible for the observed phenotypic variation. GWAS relies on LD, because this refers to the associations between SNP and the alleles at mutations affecting traits that appear because the SNP and mutations co-occur on chromosome segments of individuals in the current population (Hayes & Goddard, 2010).

Bayesian regression methods, which combine a priori information on marker effects with phenotypes related to the traits of interest, can be used to concurrently estimate all effects from available markers (Meuwissen, Hayes, & Goddard, 2001) and are suitable for GWAS (Fernando & Garrick, 2008, 2013; Habier, Fernando, Kizilkaya, & Garrick, 2011). Recently, GWAS have been carried out for growth traits in beef cattle. Santiago et al. (2017) identified genomic regions associated with carcass and growth traits in Canchim cattle in a Bayesian framework and Terakado et al. (2018) pinpoint regions associated with these kind of traits in Nelore cattle, using single-step GBLUP method with pedigree, phenotype and genotype data. Other GWAS using least-square means, Bayesian methods and genomewide rapid association using mixed model and regression method for different breeds have been performed in beef cattle (Pausch et al., 2011; Peters et al., 2012; Saatchi, Schnabel, Taylor, & Garrick, 2014; Santana et al., 2014; Weng et al., 2016).

One of the main advantages of GWAS is to provide first insights towards identification of genetic variants and biological mechanisms involved in the expression of the phenotype (Wang, Barratt, Clayton, & Todd, 2005). A GWAS can be useful to detect mutations close to QTLs, to select SNP associated with the trait for genomic selection and to construct further gene networks (Eggen, 2012; Snelling et al., 2012). In addition, this methodology may lead to designs of low-density marker panels that explain a substantial part of the genetic variance for these traits. If a good predictive ability is proven, these panels can become highly desirable as lower-cost solutions for commercial applications (Van Eenennaam, Weigel, Young, Cleveland, & Dekkers, 2014).

The aim of this study was to perform a GWAS using Bayesian methodology to identify genomic regions, tag-SNPs and propose candidate genes through functional enrichment analysis associated with growth traits in Hereford and Braford cattle. With that, we proposed low-density panels, verified their prediction accuracies and compared those with results from the full set of SNP markers.

2 | MATERIAL AND METHODS

2.1 | Phenotype, genotype and pedigree data

Phenotypic records of 126,290 animals represented by Hereford (28,392 animals) and Braford (97,898 animals) beef cattle breeds, born between 1991 and 2012 and belonging to Conexão Delta G genetic breeding programme, located in the Brazilian state of Rio Grande do Sul, were used. The growth traits evaluated were as follows: birthweight (BW), weaning weight adjusted to 205 days of age (WW205), yearling weight adjusted to 550 days age (YW550) and postweaning weight gain adjusted for 345 days of age (PWG345).

A total of 3,545 genotyped animals (624 Hereford and 2,921 Braford) with the Illumina BovineSNP50 Bead Chip (Illumina, 2006) and 131 sires (59 Hereford and 72 Braford) genotyped with the Illumina High-Density Beef Chip Array (Illumina, 2006) were used. Genotype quality control was implemented using R/SNPStats package (Clayton, 2014). Samples with Call Rate (CR) bellow 0.90, heterozygosity exceeding three standard deviations above or below the observed mean, mismatching sex and duplicated records were removed. Only SNP mapped to autosomes with CR above 0.98, minor allele frequencies (MAF) higher than 0.03, and not in highly significant deviation Hardy–Weinberg equilibrium ($p > 10^{-7}$) were considered in the analysis. In addition, only the SNP with highest MAF was maintained when SNPs were observed in the same position or the genotypes were highly correlated ($r > .98$). Data from HD chip were filtered to select only SNPs that were also present on the 50K panel. Missing genotypes (0.86% of all genotypes) were imputed using the FImpute software (Sargolzaei, Chesnais, & Schenkel, 2011). Finally, 41,045 SNPs and 3,592 samples were used for further analysis, including 2,934 Braford and 658 Hereford animals.

2.2 | Statistical models and analysis

The dataset is the same used in Campos et al. (2018), and more details can be found in this paper. The (co)variances components and genetic parameters were estimated through Bayesian inference by Gibbs sampling, using the Gibbs2f90 program (Misztal et al., 2015). The following model was applied:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Wm} + \mathbf{Spe} + \mathbf{e} \quad (1)$$

where **y** is the vector of observations, **b** is the vector of fixed effects, **u** is the vector of random direct additive-genetic effects, **m** is the vector of random maternal additive-genetic effects, **pe** is the vector of maternal permanent environmental effects, **e** is the vector of random residual effects, and **X**, **Z**, **W** and **S** are incidence matrices relating the observations to fixed, direct additive-genetic, maternal additive-genetic and maternal permanent environmental effects, respectively. For traits measured at yearling (YW550, PWG345), maternal genetic and maternal permanent environmental effects were not considered.

The fixed effects of contemporary groups (CG—same farm, sex, year and season of birth, management group, and date of

TABLE 1 Description of the dataset used in the analyses for growth and visual scores traits evaluated in Hereford and Braford breeds

Traits[a]	Number of observations	Mean (*SD*)	Number of CG	Number of sires	Number of dams
BW (kg)	105,596	33.00 (±5.32)	1,045	2,162	41,279
WW205 (kg)	113,022	179.53 (±32.87)	2,884	2,226	43,886
WW550 (kg)	60,422	315.87 (±77.37)	3,285	1,971	30,546
PWG345 (kg)	59,617	133.80 (±65.08)	3,261	1,965	30,295

Abbreviation: *SD*: standard deviation; CG: contemporary groups.

[a]BW: birthweight; WW205: weaning weight adjusted to 205 days of age; YW550: yearling weight adjusted to 550 days of age; PWG345: postweaning weight gain adjusted for 345 days of age.

weaning or yearling), age of dam at calving (15 age classes, in which class 1 corresponded to cows with 2 years of age and 15 to cows older than 16 year), linear and quadratic effects of animal age at measurement (in days), and linear covariates for breed and heterozygosity effects were considered. In addition, linear maternal breed and maternal heterozygosity effects were fit only for preweaning traits (BW and WW205). The general structure of the data analysed after editing is shown in Table 1.

2.3 | Genomewide association study (GWAS) using Bayesian framework

After calculating estimated breeding values (EBV) for growth traits using Equation (1), deregressed EBV (DEBV) and corresponding weights (**w**) = {wj} for all animals were generated by applying the procedure proposed by Garrick, Taylor, and Fernando (2009). These DEBVs were analysed in a Bayesian framework model with SNP allelic substitution random effects using GenSel software version 4.0 (Fernando & Garrick, 2008).

BayesA and BayesB methods (Meuwissen et al., 2001) were used to analyse the DEBVs and genotypes simultaneously. Both methods consider that each SNP has a specific variance for each locus, with a scaled inverted chi-square distribution χ^{-2} (v, S) with degrees of freedom (v) and scale (S) determined in GenSel by default. However, BayesB assumes that a proportion of markers has null effect, with fixed probability π, whereas the remaining markers have normally distributed effects with specific locus variance (1-π probability). On the other hand, all SNPs are fit ($\pi = 0$) in each Markov chain Monte Carlo (MCMC) cycle in the BayesA method. The following statistical model was used for the two methodologies:

$$\boldsymbol{y} = \sum_{i=1}^{k=41,045} \boldsymbol{\delta}_i \boldsymbol{z}_i \boldsymbol{a}_i + \boldsymbol{e}, \quad (2)$$

where **y** is the vector of phenotypes (DEBV); $k = 41{,}045$ is the total of SNPs in the panel; δ_i is the indicator that the SNP i was included ($\delta_i = 1$) or excluded ($\delta_i = 0$) from the model for a given iteration of MCMC; $\boldsymbol{z_i}$ is the vector of genotypes for each marker i, coded as −10/0/10; $\boldsymbol{a_i}$ is the random-effect vector of substitution of the adjusted marker i with its own variance $\sigma^2_{a_i}$ and an a priori null effect with probability π or non-null with probability $1 - \pi$; and **e** is the residual random vector with normal distribution. For the BayesA method, δ_i is assumed to be 1. For BayesB, the parameter π was assumed to be equal to 0.995 (Fan et al., 2011; Wolc et al., 2011). A chain size of 42,000 samples was used, where the first 2,000 were discarded as burn-in. The convergence of MCMC was verified by the Geweke test (Geweke, 1992) using the software R in package "boa" (Smith, 1997). Variance components estimated in the same population by Campos et al. (2018) were used as a priori to run the Bayesian GWAS analyses in Gensel.

The animals genotyped for all the traits considered in this study were divided into four (yearling traits) or five groups (preweaning traits) by two cross-validation techniques using the R software (R Core Team, 2015). The first strategy was according to genomic relationship, by the *k*-means clustering, and secondly at random, according to Saatchi et al. (2011) for the purpose of creating independent training and validation groups. These strategies were used to represent two scenarios of genomic selection so that the animals are more distantly or more closely related to the training set. In order to select subsets of SNPs and to compare genomic prediction, GWAS analyses were performed using the BayesB method ($\pi = 0.995$) considering all animals and for each clustering groups.

2.4 | Top windows and tag-SNPs

Selection of top windows and tag-SNPs were performed according to the methodology proposed by Sollero, Junqueira, Gomes, Caetano, and Cardoso (2017). We selected more informative windows (top windows) that explained at least $5 \times 100\%/ n$ of the trait genetic variance, in which, n is equal the number of 1 Mb total windows in the genome (2,519) (Onteru et al., 2013). Then, assuming an equal contribution of all genomic regions, we selected 1-Mb size windows that explained at least 0.2% of the genetic variance.

In summary, the strategy for selecting Tag-SNPs within the most representative windows considered the Bayesian parameters provided by Gensel program, model frequency (MF—proportion of post-burn-in iterations that included

that particular SNP in the model) and t.like statistic (TL—the absolute value of posterior mean effects divided by the respective standard deviations of those effects), linkage disequilibrium (LD) and the minor allele frequency (MAF). First, we selected SNPs with the maximum MF within each top window as top SNPs and then, select only SNPs included within top windows with MF values above the minimum observed MF for top SNPs. Secondly, we determined the minimum TL value to select the remaining SNPs within top windows that exceeded this minimum TL value among the SNPs selected in the previous step. Finally, the final step to construct the Tag-SNP panel was to avoid redundant SNPs due to observed high LD among subsets of SNPs preselected by MF and TL. So, when two SNPs were observed with r^2 values higher than 0.4 (Badke, Bates, Ernst, Schwab, & Steibel, 2012), only the SNP with the highest MAF was retained (Sollero et al., 2017).

2.5 | Prediction accuracy of low-density SNP panel

In order to verify the efficacy of choosing Tag-SNPs for growth traits in each, BayesA method was applied to cross-validate each cluster independently assuming that all top SNPs have non-zero effects. Sollero et al. (2017) have described additional details of this approach.

The combined prediction accuracy across all groups was derived as the genetic correlation between the observed phenotypes (y) and the DGV for a given trait, estimated in a two-trait animal model, using a modified relationship numerator matrix (**A***), where the covariance between individuals of different groups were zeroed, according to Saatchi et al. (2011). The components of (co)variances and genetic correlations were estimated by Bayesian inference using Gibbs2f90 software (Misztal et al., 2015).

Within a given cluster (c), the predicted accuracy was estimated alternatively as the correlation between phenotype adjusted for fixed effects and DGV, as proposed by Legarra, Robert-Granié, Manfredi, and Elsen (2008). For comparison, DGV prediction was also estimated for the same cross-validation techniques (*k*-means and random clustering) with the BayesB ($\pi = 0.995$) and BayesA methods considering all 41,045 markers.

2.6 | Functional enrichment analysis

The Medical Subject Headings (MeSH) database was used to define functional sets of genes potentially related to the growth traits. Gene information was taken from the Bovine UMD3.1 genome assembly (Zimin et al., 2009) using the Bioconductor R package biomaRt (Durinck, Spellman, Birney, & Huber, 2009). The biomaRt package that accesses and retrieves Esembl data (Entrez IDs, Ensembl gene IDs, HGNC symbols and others) was also used to download all genes from the genome of *Bos taurus* (ORG.MESH.BTA.DB) and map the genes within ±200 kb from the location of each SNPs (Mota et al., 2018). The complete list of selected (mapped) genes was then applied for enrichment analysis in Bioconductor retrieving statistically over-represented annotations (Nelson, Schopen, Savage, Schulman, & Arluk, 2004).

The association of a given MeSH descriptor to the growth traits was analysed using Fisher's exact test, which is a test based on the cumulative hypergeometric distribution (Gambra et al., 2013). This test was performed to search for significant genes in a given functional category among all genes (over-representation analysis [ORA]). In particular, the *p*-value of significant g genes observed was calculated as follows:

$$p-\text{value} = 1 - \sum_{i=0}^{g-1} \frac{\binom{S}{i}\binom{N-S}{K-1}}{\binom{N}{K}}$$

where S is the total number of genes that were considered significantly associated with the traits studied, N is the total number of genes analysed in the study, and K is the total number of genes in the functional category under consideration. The MeSH enrichment analysis was performed using the "Meshr" package (Morota et al., 2015), available in the R environment (R Core Team, 2015).

3 | RESULTS AND DISCUSSION

The additive direct heritability ranged from 0.13 to 0.28 and maternal heritability ranged from 0.07 to 0.09 for growth traits (Table 2). The results were the same as estimated in Campos et al. (2018) and are in concordance with the values reported in the literature for Hereford and Braford cattle.

K-means clustering methods produced groups with unbalanced numbers of animals for preweaning and yearling traits (Campos et al., 2018), as well as the average of genomic relationship (G_{ij}) within a group was higher than the average between groups. For the random clustering, the groups had a balanced number of animals, and there was no difference in G_{ij} values within and between groups, being close to zero. The aim was forming distinct groups to represent scenarios with animals genetically distant (*k*-means) or more closely related (random), between the training and the validation populations.

TABLE 2 Posteriori means and standard deviation (*SD*) of direct additive-genetic variances (σ_a^2), maternal genetic variances (σ_m^2), genetic additive-maternal (co)variances (σ_{am}), maternal permanent environment variances (σ_{pe}^2) and residual variances (σ_e^2), and direct heritability (h_a^2) and maternal heritability (h_m^2)

Trait[a]	σ_a^2	σ_m^2	σ_{am}	σ_{pe}^2	σ_e^2	h_a^2	h_m^2
BW	4.98 (± 0.25)	1.23 (± 0.13)	−0.93 (± 0.14)	0.61 (± 0.08)	12.21 (± 0.18)	0.28 (± 0.01)	0.07 (± 0.01)
WW205	123.39 (± 9.06)	50.29 (± 5.17)	−27.48 (± 5.22)	133.34 (± 3.85)	284.03 (± 6.20)	0.22 (± 0.01)	0.09 (± 0.01)
YW550	273.14 (± 11.69)	–	–		761.30 (± 10.55)	0.26 (± 0.01)	–
PWG345	96.35 (± 7.17)	–	–		649.44 (± 7.30)	0.13 (± 0.01)	–

[a]BW: birthweight; WW205: weaning weight adjusted to 205 days of age; YW550: yearling weight adjusted to 550 days of age; PWG345: postweaning weight gain adjusted for 345 days of age.

3.1 | Bayesian genomewide association study (GWAS)

The BayesB method, including all animals and markers, was used in the analysis with the parameter π of 0.995 (Fan et al., 2011; Wolc et al., 2011), which corresponds to 0.5% of the SNPs adjusted in the model at each iteration. A very large π value was chosen to allow only regions with strongest associations to be identified (Wolc et al., 2011; Zare, Shook, Collins, & Kirkpatrick, 2014). According to Zare et al. (2014), smaller values of π should increase the proportion of genetic variance explained by all SNPs and, therefore, can influence in the identification of larger QTL (Fernando & Garrick, 2013).

Heritabilities (h^2) estimated by Bayesian methodology for BW, WW205, YW550 and PWG345 were, respectively, 0.22, 0.14, 0.28 and 0.16. These h^2 were lower in relation to those estimated by the pedigree (Campos et al., 2018) for BW and WW205 and of greater proportion for YW550 and PWG345, indicating that for yearling traits, the SNPs explained larger amount of the variation as compared to the pedigree model. Peters et al. (2012) analysing data from Brangus cattle and using the BayesC method ($\pi = 0.999$) found lower h^2 values for weaning weight (0.13), weaning weight adjusted for the 205 days of age (0.04) and yearling weight adjusted for 365 days (0.19) compared to those in this present study.

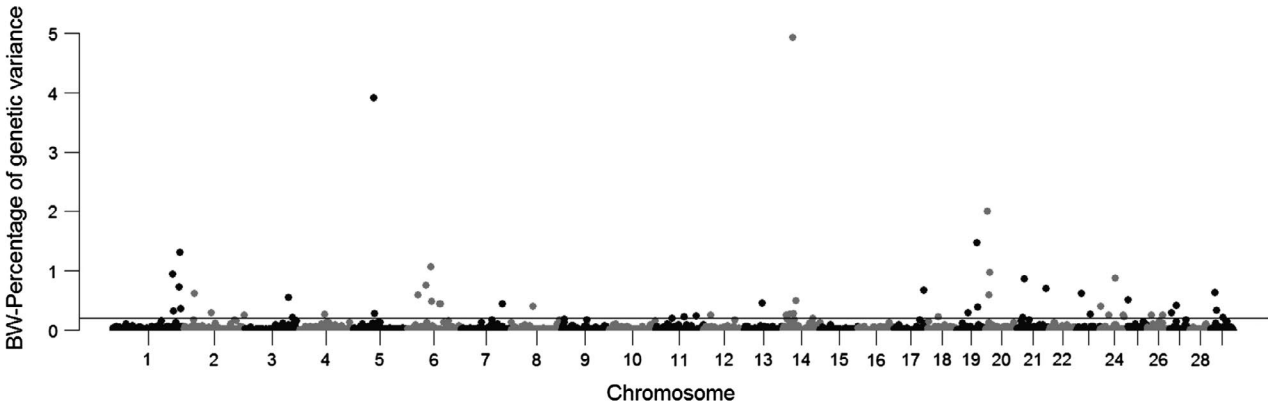
The proportion of the genetic variance explained by the 2.519 of 1-Mb windows using all markers (41,045 SNPs) across the genome for BW, WW205, YW550 and PWG345 are presented in Manhattan plots (Figure 1). According to the GWAS using the BayesB method ($\pi = 0.995$), the most representative windows were selected considering the threshold of 0.2% for the genetic variance explained by each window. The top 1% of 1-Mb windows for growth traits are shown in Table 3.

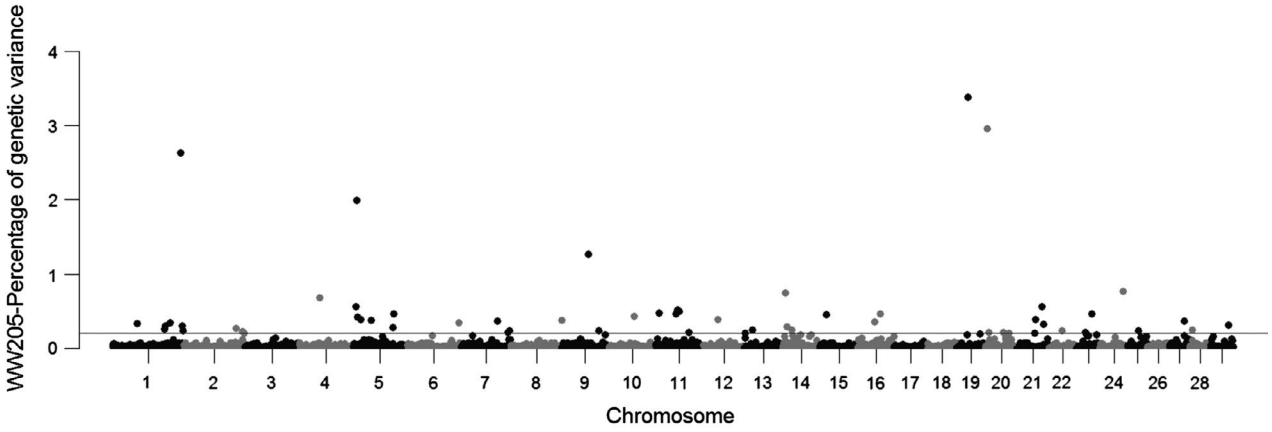
For BW, 55 windows represented by 1,047 SNPs accounted for 36.86% of the total of genetic variance explained by the markers (Table 4). The proportion of total variance explained by 55 windows and 1,060 SNPs for WW205 was equal to 30.71% (Table 4). For the yearling traits, 75 windows (1,369 SNPs) were responded for 37.47% of the genetic variance for YW550 and a total of 59 windows (1,008 SNPs) accounted for 28.8% of the additive variance for the PWG345 trait (Table 4). The number of SNPs in each window ranged from 5 to 28 for growth traits.

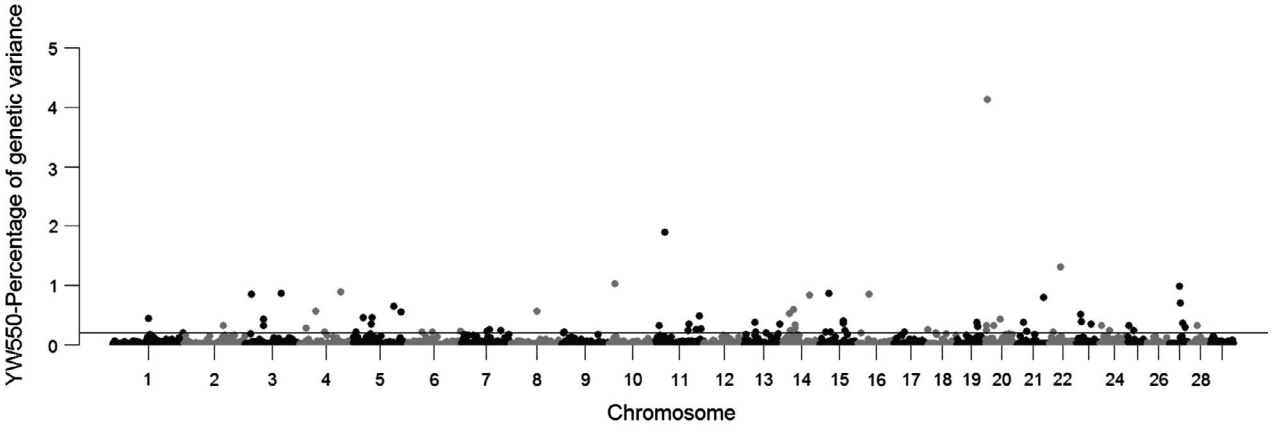
More specifically, for BW, the top 1% of 1-Mb windows (Table 3) were observed on BTA 14, 5, 20, 19, 1 and 6, representing approximately 15% of the additive-genetic variance of the trait. The window located on BTA 14 individually presented the greater proportion of the genetic variance (4.94%), indicating a potential region with greatest impact on the trait. Accordingly, Utsunomiya et al. (2013) and Terakado et al. (2018) also found important SNPs on BTA 14 which explained more than 4.0% of the genetic variance for BW in Nellore cattle. Weng et al. (2016) studying Brangus cattle reported SNPs explaining more than 0.5% of the additive-genetic variance on BTA 6, 5 and 20, similar to the regions found in the present study, but also identified important SNPs on BTA7.

The top 1% of 1-Mb windows for WW205 are listed in order (Table 3) and explained 10.95% additive-genetic variance. The top 1% windows were located on the chromosomes BTA 1, 5, 19, and 20. Weng et al. (2016) found more representative windows on BTA 7, BTA 6, BTA 10, BTA 16 and BTA 29 in Brangus cattle, while Peters et al. (2012) identified those within BTA 7, 6, 20, 12, 5, 14 and 10, considering weaning weight of heifers in the same breed. In the present study, representative SNPs were also found on chromosomes 6, 7, 12, 10, 14 and 16, but the genetic variance explained by these ones was not indicated as potential QTLs in our results. Jahuey-Martínez et al. (2016) detected important SNPs on BTA 4, 11, 18 and 29 for weaning weight in Charolais cattle.

The top 1% 1-Mb windows (Table 3) for traits measured at yearling were found on BTA 10, 11, 20 and 22, which explained 8.39% of the additive-genetic variance for the YW550. Considering the PWG345 trait, the top 1% 1-Mb windows located on BTA 2, 3, 14, 16, 6 and 22 explained 9.16% of the genetic variance. Buzanskas et al. (2014) identified SNPs associated with yearling weight adjusted for 420 days of age in Canchim cattle at BTA7, 22, 25 and 27. Moreover, Peters et al. (2012) found other significant


BW-Percentage of genetic variance
0
1
2
3
4
5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 24 26 28
Chromosome


WW205-Percentage of genetic variance
0
1
2
3
4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 24 26 28
Chromosome


YW550-Percentage of genetic variance
0
1
2
3
4
5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 24 26 28
Chromosome

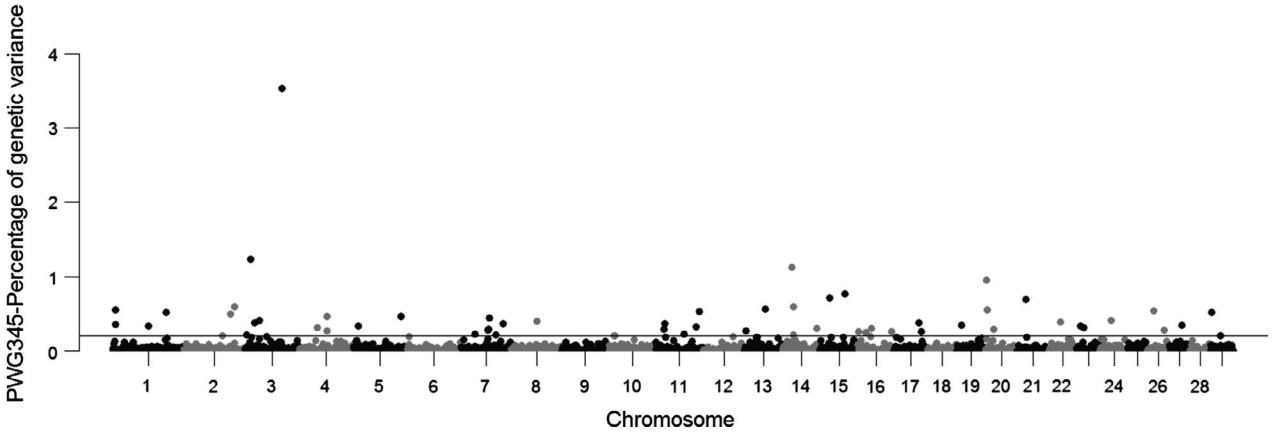

PWG345-Percentage of genetic variance
0
1
2
3
4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 24 26 28
Chromosome

FIGURE 1 Manhattan plot charts showing the results for the genome-wide association study performed by the BayesB method ($\pi = 0.995$) for birth weight, weaning weight adjusted for 205 days of age, yearling weight adjusted for 550 days of age and post-weaning gain for 345 days of age, respectively. The *Y*-axis represents the proportion of the genetic variance explained by 1 Mb windows and the *X*-axis represents the chromosomes that the windows are located.

windows on BTA6, 7, 8 and 9 for weight adjusted at 365 days of age and on chromosomes 5, 6, 7, 8, 9 and 10 for average daily gain from birth to 365 days of age in Brangus breed heifers. Santiago et al. (2017) detected important SNPs on BTA6 using BayesB method for yearling weight in Canchim cattle. It is known that most of those performance or growth traits follow the pattern of quantitative traits spread over the entire genome, but in our study nine top Windows were the same between the preweaning traits (BW and WW205) and 26 top Windows in common between the yearling traits (YW550 and PWG345). Nevertheless, the only common region among all growth traits is located in BTA 20_4 (Chr_Mb). We can infer that the control of the growth traits can be predominantly influenced by these regions in common and, consequently, that one panel with the same SNPs perhaps should be enough to perform genomic prediction for such traits.

Some chromosomal regions associated with the traits of interest in our study have not been previously mentioned in the literature, despite that growth traits are extensively studied in beef cattle. Between the top 1-Mb windows that explained >1.0% of genetic variance, some were located in regions not previously described in the literature for growth traits, like BTA 19 for BW, BTA 19, 1 and 9 for WW205, BTA 11 and 22 for YW550 and BTA 2, 3 and 14 for PWG345. This is, however, the first association study for these traits in taurine populations from Brazil. The differences of allelic frequencies between cattle breeds and the extent of LD between SNPs and causal variants should result in different effects of markers detected in different cattle populations (Biegelmeyer, Gulias-Gomes, Caetano, Steibel, & Cardoso, 2016; O'Brien et al., 2014; Tizioto et al., 2013). The calculated allelic substitution effect of a SNP varies from one dataset to another or between different methodologies applied, while the associations with the phenotype may be significant or not, and ultimately the amount of variance explained by the SNP may also differ substantially (Bolormaa et al., 2013).

Performing GWAS using information of more than one breed is a way to increase the sample size and can improve the power to detect associated variants because the LD is supposed to be conserved in short distances across breeds (van den Berg, Boichard, & Lund, 2016; Raven, Cocks, & Hayes, 2014). Therefore, as more data from different breeds simultaneously analysed, better to understand the mechanisms of genetic control of animals that may be born lighter but grow efficiently.

3.2 | Tag-SNP selection

The top SNPs (TopSNP) within each top windows (TopW) were selected based on the MF and TL parameters provided by Gensel program, to select potential markers to compose a low-density panel. Therefore, LD between SNPs was considered, as well as MAF.

Among all the 52 TopW and 1,032 TopSNP (Table 2) above the variance threshold for BW, we selected 103 Tag-SNPs (Table 4) to the reduced panel representative for this trait. The highest MF value was found to be 0.616 for rs109590980/ARS-BFGL-NGS-73392 marker, indicating that this particular SNP located in the BTA5 within the most informative 1 Mb window for BW genetic variance was fitted in 61.6% of the MCMC.

For WW205, 79 Tag-SNPs out of the TopW that explained more than 0.2% of the variance remained as more representative. The rs110133873/ARS-BFGL-NGS SNP located on BTA19-4904 showed the highest MF value (0.712) for that trait.

The selection of the TopW and TopSNP for YW550 resulted in 112 SNPs selected. Subsequently, 20 out of 112 SNPs were excluded due to LD ($r^2 > .4$) (Badke et al., 2012; Sollero et al., 2017) with another SNP within the same TopW and lower MAF. Therefore, the resulting panel for the YW550 had 100 Tag-SNP. The rs110245673/ARS-BFGL-BAC-11047 SNP, within BTA11, had a higher MF value (0.545), among those selected. Finally, out of the 1,032 TopSNP, there were 77 Tag-SNP remaining for PWG345 after the selection by the MF, TL and LD parameters. The highest MF value observed among those selected SNPs was 0.602 for the rs41567611/Hapmap50215-BTA-111911 marker, located on BTA2.

According to Fernando and Garrick (2013), markers with higher MF value should be prioritized for further inferences. In addition, when $\pi > 0$ in Bayesian analyses, the MF will be highly correlated with the estimate of the absolute value of the variance effect (Fernando & Garrick, 2008). Therefore, SNPs with higher MF within the top windows are potentially associated with the phenotype of interest. Once SNPs of higher MF have been identified, we could exclude (optimize the panels) those in LD with them, as they would already be mapping-related genes (Sollero et al., 2017).

The correlations (r) between MF and TL, used to distinguish informative and consistent markers of those that generated inconsistent estimates (Fernando & Garrick, 2008) varied from 0.40 to 0.43 for the traits. Once TL refers more

TABLE 3 Top 1-Mb windows[a] that explain >1.0% of genetic variance for growth traits of Hereford and Braford breeds

Traits[b]	Chromosomes	Chromosome_Mb	1 Mb window number	Number of SNPs	% genetic variance
BW	14	14_24	1,523	15	4.94
	5	5_46	585	10	3.91
	20	20_40	1,961	23	2.00
	19	19_44	1,936	20	1.48
	1	1_149	1,49	17	1.31
	6	6_52	713	23	1.07
WW205	19	19_25	1,917	27	3.38
	20	20_40	1,961	23	2.96
	1	1_151	151	19	2.63
	5	5_80	547	26	1.99
	9	9_59	1,067	28	1.27
YW550	20	20_40	1,961	23	4.14
	11	11_21	1,239	24	1.90
	22	22_25	2,126	12	1.32
	10	10_13	1,127	24	1.03
PWG345	2	2_103	262	11	2.15
	3	3_83	379	19	1.80
	14	14_76	1,575	13	1.58
	16	16_60	1,676	9	1.35
	3	3_12	308	8	1.26
	6	6_118	779	18	1.02

[a]Each chromosome was divided into one mega base non-overlapping windows, containing a variable number of SNPs with the BayesB method for $\pi = 0.995$; Chromosome_Mb = position into the chromosome; 1 Mb window number = it is the number of window into the 2,519 windows; number of SNPs = number of SNPs in each window; % genetic variance = percentage of genetic variance explained for each window.

[b]BW: birthweight; WW205: weaning weight adjusted to 205 days of age; YW550: yearling weight adjusted to 550 days of age; PWG345: postweaning weight gain adjusted for 345 days of age.

TABLE 4 Number of top windows (TopW)[a], number of SNPs in each window (TopSNPs), sum of % genetic variance explained all top windows, minimum, mean and maximum values of model frequency and t.like statistics and number of Tag-SNPs for growth traits in genomewide association study

				Model frequency			t.like			
Traits[b]	TopW	TopSNPs	%Var	Min	Mean	Max	Min	Mean	Max	Tag-SNPs
BW	58	1,047	36.86	0.044	0.137	0.616	0.910	0.944	1.198	103
WW205	57	1,060	30.71	0.049	0.150	0.712	0.906	0.948	1.313	79
YW550	79	1,369	37.47	0.058	0.180	0.525	0.917	0.957	1.106	100
PWG345	59	1,032	28.80	0.047	0.145	0.770	0.893	0.940	1.421	78

[a]TopW: represents the number of windows that explained above 0.2% of the genetic variance in the BayesB ($\pi = 0.995$) GWAS analysis in each group; TopSNPs: represents the number of SNPs included in each top window; Model Frequency: proportion of the chain that included a given marker in the model; t.like: ratio of the mean a posteriori effect only of the chains that included the effect on the model on the standard deviation of the distributions of this model; Tag-SNPs: represents the number of SNPs selected according to the parameters model frequency and t-like statistics, linkage disequilibrium and minor allele frequency.

[b]BW: birthweight; WW205: weaning weight adjusted to 205 days of age; YW550: yearling weight adjusted to 550 days of age; PWG345: postweaning weight gain adjusted for 345 days of age.

to the magnitude of the effect (standardized, considering the standard deviation) and due to its moderate correlation with MF, we select informative SNPs using both parameters TL and MF (Fernando & Garrick, 2013; Sollero et al., 2017).

Finally, 385 Tag-SNPs were selected as more informative for all the traits evaluated. Based on the optimized number and once only two SNPs out of those were considered redundant, that is, the same SNPs were indicated for different

TABLE 5 Groups of animals according to the cross-validation strategy, *k*-means (cl) or random (ran), posterior mean proportion of variance explained by markers (h^2) using Bayesian methods, number of top windows (TopW), number of SNPs in each window (TopSNPs) and number of Tag-SNPs for each growth traits

		Groups									
Traits[b]	**Parameters**[a]	**cl1**	**cl2**	**cl3**	**cl4**	**cl5**	**ran1**	**ran2**	**ran3**	**ran4**	**ran5**
BW	h^2_BB995	0.26	0.22	0.22	0.23	0.21	0.21	0.23	0.23	0.24	0.21
	h^2_BA	0.27	0.28	0.29	0.26	0.31	0.29	0.29	0.29	0.29	0.28
	TopW	41	57	57	48	52	60	54	52	47	53
	TopSNPs	753	1,057	1,067	915	982	1,110	985	945	844	1,002
	Tag-SNPs	59	85	72	79	112	91	86	81	75	81
WW205	h^2_BB995	0.17	0.16	0.16	0.15	0.15	0.15	0.16	0.16	0.16	0.14
	h^2_BA	0.24	0.24	0.24	0.23	0.22	0.25	0.24	0.23	0.23	0.21
	TopW	46	46	41	54	52	54	53	46	41	42
	TopSNPs	810	895	762	984	980	995	1,001	864	780	754
	Tag-SNPs	59	69	59	78	69	95	72	62	65	72
YW550	h^2_BB995	0.22	0.24	0.22	0.23	–	0.23	0.22	0.24	0.22	–
	h^2_BA	0.29	0.32	0.34	0.30	–	0.34	0.32	0.34	0.30	–
	TopW	57	54	55	42	–	60	55	56	58	–
	TopSNPs	1,030	955	948	742	–	1,045	1,014	951	1,068	–
	Tag-SNPs	87	84	88	67	–	108	73	80	86	–
PWG345	h^2_BB995	0.13	0.11	0.11	0.12	–	0.11	0.14	0.10	0.12	–
	h^2_BA	0.22	0.21	0.18	0.18	–	0.19	0.18	0.16	0.19	–
	TopW	48	40	31	38	–	39	33	36	42	–
	TopSNPs	891	725	538	661	–	739	575	647	758	–
	Tag-SNPs	83	91	55	53	–	67	44	49	48	–

[a]h^2_BB995: *posterior* mean proportion of variance explained by markers using BayesB method ($\pi = 0.995$); h^2_BA: *posterior* mean proportion of variance explained by markers using BayesA method; TopW: Top windows represents the number of windows that explained above 0.2% of the genetic variance in the BayesB ($\pi = 0.995$) GWAS analysis in each group; TopSNPs: represents the number of SNPs included in each top window; Tag-SNPs: represents the number of SNPs selected according to the parameters *model frequency* and *t-like* statistics, linkage disequilibrium and minor allele frequency.

[b]BW: birthweight; WW205: weaning weight adjusted for 205 days of age; YW550: yearling weight adjusted for 550 days of age; PWG345: postweaning gain adjusted for 345 days of age.

proposed panels (traits), its application is suggested as a general low-density panel for growth traits in Hereford and Braford beef cattle. Habier, Fernando, and Dekkers (2007) emphasized that a major disadvantage of marker selection approaches, would be if different SNPs are selected for different traits and that low-density (LowD-SNP) panel developed for one trait, may not work so well for another. Thus, the optimized design of LowD-SNP panels enabling the test of different traits has to be considered for its proper commercial application, so the costs for development and genotyping may compete with the medium-density panels.

3.3 | Prediction accuracy of Tag-SNP panels

The same Tag-SNP selection strategy using the BayesB ($\pi = 0.995$) method described above was used in each of the *k*-means and random cross-validation groups, generating 10 alternative panels for the preweaning traits and eight for the yearling traits (Table 3). The number of SNPs selected (sublists) varied between all growth traits, from 41 to 108 SNPs in each group (Table 5). The mean a posteriori estimation of h^2 with BayesA method using the Tag-SNPs panel selected in each clustering technique was higher than the h^2 using BayesB with $\pi = 0.995$ applied to the full set of markers (Table 5).

Bayesian methods generally use shrinkage procedures to consider different variations for individual SNPs, in which a priori information on the SNP effect distribution is used to force insignificant effects to zero (Meuwissen et al., 2001). Moreover, some studies have demonstrated the advantage of Bayesian methods to capture the LD between markers and QTL (Cardoso et al., 2015; Habier et al., 2011). The BayesA method was accomplished in our strategy in order to propose the Tag-SNP reduced panel(s) ensuring that only the most significant markers with specific variances and detectable effects were selected.

TABLE 6 Genetic correlations and standard deviation (*SD*) between phenotype observed and direct genomic values (DGV) of cross-validation predictions using different methods for growth traits in Hereford and Braford breeds

		Methods[b]		
Traits[a]	**Cluster**	**BayesATag-SNP**	**BayesA**	**BayesB**
BW	*k*-means	0.30 (±0.08)	0.38 (±0.09)	0.36 (±0.09)
	Random	0.65 (±0.20)	0.77 (±0.17)	0.74 (±0.16)
WW205	*k*-means	0.20 (±0.06)	0.26 (±0.05)	0.25 (±0.05)
	Random	0.30 (±0.08)	0.47 (±0.13)	0.42 (±0.10)
YW550	*k*-means	0.21 (±0.07)	0.33 (±0.10)	0.30 (±0.09)
	Random	0.50 (±0.23)	0.71 (±0.17)	0.67 (±0.15)
PWG345	*k*-means	0.13 (±0.09)	0.23 (±0.11)	0.20 (±0.11)
	Random	0.36 (±0.24)	0.50 (±0.22)	0.46 (±0.23)

[a]BW: birthweight; WW205: weaning weight adjusted for 205 days of age; YW550: yearling weight adjusted for 550 days of age; PWG345: postweaning gain adjusted for 345 days of age.

[b]BayesATag-SNP: Bayesian model with t-distribution and probability ($\pi = 0$), using only more informative markers; BayesA: Bayesian model with t-distribution and probability ($\pi = 0$), using all markers (41,045); BayesB: Bayesian mixture of t-distribution and point mass on zero with probability ($\pi = 0.995$), using all markers (41,045).

The prediction accuracy was estimated by genetic correlations between the observed phenotype and DGV obtained in two-trait analyses from cross-validation with BayesA using each of the Tag-SNP sublist (BayesATag-SNP), and using all markers (41,045) also with BayesA and BayesB ($\pi = 0.995$) methods (Table 6). Accuracies were higher for all traits when using groups formed randomly compared to *k*-means clustering (0.30–0.77 vs. 0.13–0.38, respectively) due to larger genetic relationship distance between training and validation sets imposed by the *k*-means clustering method. Although results showed better accuracy for random compared with *k*-means methodology, as pointed out by one of the anonymous reviewers, in this case random clustering could overestimate the accuracies results, once older animals can be selected in the tested panel and younger animals in the reference panel. On the other hand, the *k*-means could be more precise, because in this case the groups were created in order to reduce the relationship with the tested panel (avoid bias), so lower accuracies are expected for predictions among animals that are less related to each other. It is worth to emphasize that in random grouping, there will be greater relationship between the individuals in reference and test groups, which may increase the accuracies. According to Habier et al. (2007), the prediction accuracy decreases as the additive-genetic relationships between the animals to be predicted and those in training set is also decreased.

For growth traits, genetic correlations ranged from 0.13 to 0.30 for groups generated by *k*-means clustering and from 0.36 to 0.65 for groups formed at random, using Tag-SNPs (Table 6). For preweaning traits, the greatest genetic correlation values were observed for BW: 0.30 for *k*-means and 0.65 for random clustering. For the traits measured at yearling, the

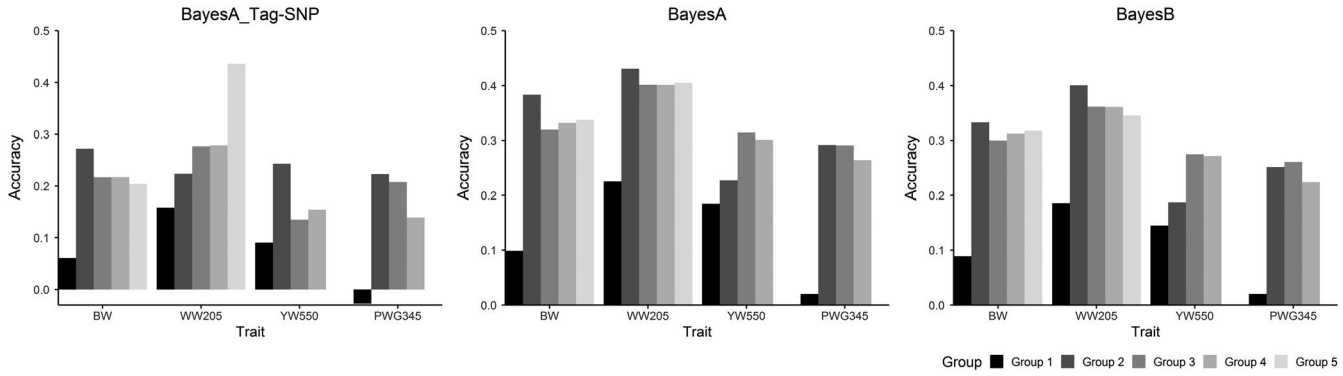


FIGURE 2 Accuracies of direct genomic predictions for each trait with *k*-means clustering cross-validation group according to the prediction method: BayesATag-SNP: Bayesian model with t-distribution and probability ($\pi = 0$), using only the most informative markers; BayesA: Bayesian model with t-distribution and probability ($\pi = 0$) using all markers (41,045); BayesB: Bayesian mixture of t-distribution and point mass on zero with probability ($\pi = 0.995$), using all markers (41,045). BW: birth weight; WW205: weaning weight adjusted for 205 days of age; YW550: yearling weight adjusted for 550 days of age; PWG345: postweaning gain adjusted for 345 days of age.

TABLE 7 Statistically significant MeSH (Medical Subject Headings) terms (ID and name) associated with the selected genes for birthweight (BW) for phenomena and process and anatomy categories

Categories	MeSH term ID	MeSH term name	Significant genes	*p*-value*
Phenomena and process	D013091	Spermatogenesis	FAS, LOC404073, MTNR1A	.011
	D018900	Dinucleotide repeats	FAS, MAP1B	.014
	D042002	Chromatin assembly and disassembly	HIST1H2BN, HIST1H3C	.014
	D005746	Gastric emptying	CARTPT	.019
	D005496	Follicular atresia	FAS, CARTPT	.02
	D002886	Chromosomes, Human, Pair 17	SOST, MYL4	.024
	D001823	Body composition	TG, XKR4, MTNR1A	.031
	D003241	Consanguinity	SOST	.037
	D012143	Respiratory physiological phenomena	TAP	.037
Anatomy	D012661	Semen	PPY, PYY2, PYY	.0058
	D013094	Spermatozoa	PIM1, FAS, MOS, MYL4, ESD	.06
	D006822	Hybrid cells	LOC404073, TEKT1	.016
	D000177	Acrosome	LOC404073, TEKT1	.017
	D009046	Motor neurons	SMN2	.019
	D012670	Seminiferous epithelium	LOC404073	.019
	D016874	Neurofibrillary tangles	PADI2	.019
	D002886	Chromosomes, Human, Pair 17	SOST, MYL4	.024
	D013687	Telencephalon	MTNR1A	.037
	D018515	Plant leaves	MTIF2	.037

*Significance declared at $p < .05$.

greatest value was 0.21 and 0.50 for YW550, with *k*-means and random clustering, respectively. For most traits, higher accuracy values were found using the full marker panel versus Tag-SNPs. Prediction accuracies are expected to decrease as marker panel density decreases because of the lower expected LD between highly dispersed markers and QTLs affecting the trait of interest (Habier, Fernando, & Dekkers, 2009).

Comparing the results using the proposed low-density panel by Tag-SNPs, for BW the accuracy corresponded to at least 79% if compared to the 41,045 SNPs using BayesA method. With the same base of comparison, for the Tag-SNP for WW205, the prediction accuracy corresponded to 77% of the one in the full panel, and 63% and 58% for YW550 and PWG345 (yearling traits), respectively. These results demonstrate the possibility of using only most informative SNPs to obtaining low-cost predictions with tolerable accuracy loss in the commercial application of low-density panels for growth traits. Campos et al. (2018) tested several genomic selection and blending methods for growth traits and visual scores and concluded that the best method was ssGBLUP using a full panel of SNPs in Hereford and Braford cattle. Therefore, combining pedigree and historical data can be suitable alternative to improve the prediction accuracy of GEBV with LowD panels for commercial application with lower genotyping cost.

Considering results between breeds, the prediction accuracy was also estimated within each cluster, *k*-means and random, as proposed by Legarra et al. (2008). For all growth traits, the group composed mostly of Hereford breed animals (Figure 2—group 1) presented lower accuracies in the *k*-means clustering, possibly due to the fact that there were fewer animals genotyped for this breed and presented the largest genetic distance to the other groups. These results were expected as demonstrated by Cardoso et al. (2015); Campos et al. (2018), Sollero et al. (2017) analysing the same dataset, but for tick resistance, growth traits and visual scores, which also found lower prediction accuracy for the Hereford breed group.

Further investigations should be carried out to improve the prediction accuracy of LowD panels. One straightforward alternative can be using blending methods to combine LowD panel with historical data such as single-step framework like ssGBLUP, as mentioned above, or Bayesian regression single-step (Aguilar et al., 2010; Fernando, Dekkers, & Garrick, 2014).

TABLE 8 Statistically significant MeSH (Medical Subject Headings) terms (ID and name) associated with the selected genes for weaning weight adjusted for 205 days of age for phenomena and process and anatomy categories

Categories	MeSH term ID	MeSH term name	Significant genes	*p*-value*
Phenomena and process	D016385	TATA box	CHGA, MYF5	.0012
	D012038	Regeneration	MYF5, MYF6	.0043
	D004357	Drug synergism	CHGA, IL2RA	.0081
	D013318	Stroke volume	CHGA	.0092
	D053140	Organelle shape	MFF	.0092
	D012091	Repetitive sequences, Nucleic acid	MYF5, COL4A3, PGA5	.013
	D016897	Respiratory burst	CHGA	.018
	D000375	Aging	MYF5, ADRB3	.035
	D001724	Birth weight	MYF5	.036
	D058507	Host specificity	IL2RA	.036
Anatomy	D005753	Gastric mucosa	CHGA, PGA5	.00082
	D016874	Neurofibrillary tangles	PADI2	.0092
	D020769	Posterior cerebral artery	CHGA	.0092
	D001479	Basal ganglia	ARPP19	.018
	D009504	Neutrophils	CHGA, IL2RA, PLCB1	.018
	D001921	Brain	CHGA, RELN, PLCB1, ASAP1, PFKFB3, FKBP3	.022
	D002837	Chromaffin granules	CHGA, ATP6V0E1	.029
	D007687	Kidney tubules, Proximal	LGMN	.036
	D002001	Adipose tissue, Brown	ADRB3	.045
	D002421	Caudate nucleus	ARPP19	.045
	D007678	Kidney glomerulus	COL4A3	.045

*Significance declared at $p < .05$.

3.4 | Functional enrichment analysis

After mapping genes through the list of target SNPs, the MeSH ORA was evaluated to discern biological properties of the identified genes suggested to play a role in the growth traits under study. Each MeSH term was clustered into two categories: "phenomena and process" and "anatomy."

Nine MeSH terms were classified in the "phenomena and processes" category and 10 in the "anatomy," representing genes statistically associated with BW (Table 7). Listed in the "body composition" MeSH term, the MTNR1A gene on BTA27 is a negative regulator of adipocyte biology, energy and lipid metabolism (Majumdar, Giri, & Pai, 2014). For these reasons, it has been used to improve the production performance and meat quality of pigs (Yang, Wang, Fu, & Zan, 2015). The TG gene, also related with the "body composition" MeSH term, is considered as a functional and positional candidate of fat thickness, postweaning average daily gain, BW and weaning weight traits (Maltecca, Weigel, Khatib, Cowan, & Bagnato, 2009; McClure et al., 2010; Zhang et al., 2015). It is the precursor of thyroid hormones being involved with metabolism regulation and therefore has effects on adipocyte growth, differentiation and homeostasis of fat depots (Ailhaud, Grimaldi, & Negrel, 1992). The Tag-SNP identified on BTA14, mapped to the MOS gene ("body composition" MeSH term) explained 4.94% of the genetic variance for BW. This gene is related to stature in cattle and humans (Gudbjartsson et al., 2008; Pryce, Hayes, Bolormaa, & Goddard, 2011). Another important gene in the same region (BTA14) was XKR4, a protein of the Kell blood group complex subunit-related family, member 4 associated with rump fat thickness (Porto Neto, Bunch, Harrison, & Barendse, 2012) and BW in Nellore cattle (Terakado et al., 2018). In addition, Utsunomiya et al. (2013) demonstrated that the same QTL region on BTA14 affects BW and size in Nellore cattle.

CARTPT ("Gastric Emptying" term) on BTA20 are neurotransmitters found to be significant regulators of body weight via both feeding behaviour and control of digestion and metabolism (Aja, Sahandy, Ladenheim, Schwartz, & Moran, 2001; Wortley, Chang, Davydova, Fried, & Leibowitz, 2004), and therefore, it is involved with traits related to body weight in cattle (Zhang et al., 2008).

Two genes, SOST and MYL4, located on BTA 19 were mapped close to the SNP that seems to explain 1.48% of the genetic variance for BW. SOST gene is a protein related to bone formation (Brunkow et al., 2001), and MYL4 gene is a

TABLE 9 Statistically significant MeSH (Medical Subject Headings) terms (ID and name) associated with the selected genes for yearling weight adjusted for 550 days of age for phenomena and process and anatomy categories

Categories	MeSH term ID	MeSH Term Name	Significant genes	*p*-value*
Phenomena and process	D000327	Adsorption	CEL, RASA1	.006
	D018919	Neovascularization, Physiologic	CDK1, S1PR1, SPARC, DPM1	.0096
	D001794	Blood Pressure	CHGA, P2RY2	.011
	D013318	Stroke Volume	CHGA	.015
	D014664	Vasodilation	CHGA, P2RY2	.03
	D002872	Chromosome Deletion	ASPH	.03
	D005720	Gamma Rays	CEL	.03
	D009006	Monosomy	RPS4Y1	.03
	D016391	Genes, src	RASA1	.03
	D016897	Respiratory Burst	CHGA	.03
	D040681	Structural Homology, Protein	MRPL24, FOLR1	.036
	D007408	Intestinal Absorption	FOLR3	.045
	D017344	Genes, Insect	MOCOS	.045
	D020239	Receptor Cross-Talk	S1PR1	.045
	D006706	Homeostasis	CHGA, PPARD	.047
Anatomy	D002837	Chromaffin Granules	CHGA, PENK, ATP6V1F, ATP6V0E1, ATP6V0C	.000067
	D014617	Vacuoles	ATP6V1F, ATP6V0C	.011
	D020769	Posterior Cerebral Artery	CHGA	.015
	D015672	Erythroid Precursor Cells	RHBG	.03
	D007425	Intracellular Membranes	CHGA, ATP6V1F, ATP6V0C	.036
	D012270	Ribosomes	IMPDH1, MRPS11, MRPL24	.044

*Significance declared at $p < .05$.

protein having fundamental roles in embryonic muscle development (Cassar-Malek, Boby, Picard, Reverter, & Hudson, 2017). Both genes can be novel candidate genes associated with BW.

Associated with WW205 trait, nine MeSH terms classified in "phenomena and processes" and 10 in "anatomy" categories also presented statistically significant genes and relevant to describe along with this trait (Table 8). Development of muscle fibres is regulated by the MyoD gene family consisting of myogenin gene (MYOG), MYF3In (MYOD1), MYF5 and MYF6 ("TATA box," "Regeneration," "Repetitive" and "Aging" MeSH terms). In livestock producing animals like cattle, myofibre numbers have been related to growth capacity (Soumillion, Erkens, Lenstra, Rettenberger, & Pas, 1997). Li et al. (2004) reported that the SNP close to or associated with the MYF5 showed a significant association with preweaning average daily gain and with average daily gain on feed in beef cattle. Another important gene found associated with the MeSH term "brain" and "chromaffin granules" for WW205 and also YW550 was the ATP6V0E1, which encodes a component of vacuolar ATPase (V-ATPase) located on BTA20 and is related to adiposity, obesity and skeletal muscle in humans and mice (Linnemann et al., 2010; Sharma et al., 2008). This gene was associated with mature metabolic weight in two populations of beef cattle in the United States, Hereford and Simmental-Angus (Seabury et al., 2017).

Although related with the "brain" MeSH term, the FKBP3 gene on BTA21 is a member of the immunophilin protein family, involved in controlling cellular function and organisation as well as in protein synthesis (Keogh et al., 2016). This gene was differentially expressed with compensatory gain (Keogh et al., 2016) and may be suggested to be a novel gene related to growth traits in beef cattle.

For the yearling traits, 17 and 25 MeSH terms were classified within "phenomena and processes" and 10 and 18 in "Anatomy" categories associated with YW550 and PWG345, respectively (Tables 9 and 10). The ASPH gene on BTA14 is involved with tissue morphology and skeletal and muscle development (Ramayo-Caldas et al., 2014), and the PPARD gene on BTA23 is a poorly studied nuclear receptor with a potentially important role in skeletal muscle metabolism, which therefore may influence growth and carcass traits (Graugnard et al., 2009; Ramayo-Caldas et al., 2014). The ATP6V0E1 gene on BTA20 associated with the SNP ARS-BFGL-NGS-102165 explained 4.14% of genetic variance for YW550. This gene has been previously described as related

TABLE 10 Statistically significant MeSH (Medical Subject Headings) terms (ID and name) associated with the selected genes for postweaning gain adjusted for 345 days of age for phenomena and process and anatomy categories

Categories	MeSH term ID	MeSH term name	Significant genes	*p*-value*
Phenomena and process	D005809	Genes, Regulator	DSG1, PRNP	.0025
	D024363	Transcription initiation site	PRND, PRNP	.0035
	D025461	Feedback, Physiological	PRNP, DUSP1	.0073
	D004058	Diffusion	PRNP, CRYAB	.0088
	D006570	Heterochromatin	H3F3A	.013
	D030762	Estrous cycle	ADRB2, SPARC, PTGER2, HSD11B1	.016
	D048430	Cell shape	ARRB1, MIR143	.018
	D009690	Nucleic acid conformation	LOX, PRNP, MRPS11, MRPL2	.024
	D015966	Gene expression regulation, Fungal	H3F3A	.026
	D054337	Cell dedifferentiation	MIR143	.026
	D014018	Tissue distribution	PRNP, SPARC, ENTPD1, PTGER2, LAMTOR5, SLC16A4, STK10, MIR145	.027
	D005822	Genetic vectors	PRNP, CRYAB, H3F3A	.033
	D020413	3' untranslated regions	PRND, PRNP, RASSF2	.033
	D059748	Proteolysis	LOX, PRNP	.036
	D005809	Disease susceptibility	PRNP	.039
	D004198	Trinucleotide repeats	PRNP	.039
	D018911	Neutron diffraction	CRYAB	.039
	D033363	Term birth	PTGER2	.039
	D047929	Substrate specificity	NEU3, PRNP, CRYAB, MTFMT, H3F3A, STK10	.046
	D013379	Protons	PRNP, ATP6V0A2	.046
Anatomy	D045744	Cell line, Tumor	ACAN, SERPINE1, PRND, PRNP, CRYAB, H3F3A	.001
	D014407	Tumor cells, Cultured	ACAN, CCNA2, HYAL2, MFGE8, MEA1, ACHE	.0024
	D016716	PC12 cells	SERPINE1, PRNP, ACHE	.0055
	D041681	NIH 3T3 cells	FN1, HTR4, H3F3A	.0062
	D006570	Heterochromatin	H3F3A	.016
	D008526	Medulla oblongata	PRNP	.016
	D008592	Menisci, Tibial	ACAN	.016
	D008650	Mesonephros	PRNP	.016
	D012501	Saphenous vein	SERPINE1	.016
	D017933	Peripheral nervous system	PRNP	.016
	D005347	Fibroblasts	FN1, ACAN, PRNP, HYAL2, TUBA1A	.02
	D007422	Intestines	PRNP, HTR4, CALB1	.03
	D011650	Pulmonary alveoli	FN1	.031
	D022162	Cytoplasmic vesicles	ACAN	.031
	D005854	Germ cells	FN1, H3F3A	.032
	D015496	CD4-Positive T lymphocytes	ACAN, LOC100335205	.039
	D005269	Femur	CCNA2	.047
	D035264	Flowers	H3F3A	.047

*Significance declared at $p < .05$.

to WW205. Related to the "Ribossome" MeSH term, the IMPDH1 gene located on BTA14 was associated with the BTB-00203359 SNP, which explained 0.90% of genetic variance for YW550. This gene is known for regulating intracellular fat accumulation, preadipocyte maturation in humans (Pazik et al., 2018).

The DUSP1 gene on BTA20 is significantly associated with the mitogen-activated protein kinase (MAPK) signalling pathway and is related to the several MeSH terms (Table 10). The MAPK cascade is a highly conserved module that controls many cellular events from complex programs, such as embryogenesis, cell differentiation, cell proliferation, cell death, to short-term changes required for homeostasis and acute hormonal responses and is related to several economically important traits such as growth, and carcass in different cattle breeds (Saatchi et al., 2014). The gene STK10 ("Tissue distribution" MeSH term) also on BTA20 acts in cell cycle arrest, protein amino acid phosphorylation and regulation of fatty acid oxidation and were associated with slaughter weight and carcass traits (Karisa et al., 2013). The CRYAB gene associated with some MeSH terms like "Genetic vectors," "Neutron diffraction," "Substrate specificity" and "Cell line Tumor" and MIR143 gene ("Cell differentiation" MeSH term), was related with quality meat and ultrassom carcass traits (Jin et al., 2012; Wang, Zheng, Wang, & Li, 2013). The NEU3 gene on BTA15 related to "Substrate specificity" MeSH term was associated with daughter pregnancy rate, cow conception rate and productive life (Cochran, Cole, Null, & Hansen, 2013).

The ARRB1 gene on BTA15 ("Cell shape" MeSH term) associated with the ARS-BFGL-NGS-116843 SNP explained 0.77% of genetic variance for PWG345 takes part in agonist-mediated desensitization of G protein-coupled receptors and cause-specific dampening of cellular responses to stimuli such as hormones, neurotransmitters, or sensory signals (Berglund, Schober, Statnick, McDonald, & Gehlert, 2003). This gene was identified acting through certain pathways of differentially expressed genes for residual feed intake (RFI) in dairy cows (Salleh et al., 2017). RFI is used to describe feed efficiency and it is related to energy support for growth and body maintenance. The ACAN gene on BTA21 related to various MeSH terms ("Cell line Tumor," "Tumor Cells Cultured," "Menisci Tibial" and "Fibroblasts) encoding an essential component of cartilage extracellular matrix aggrecan. Mutations in this gene may be involved in skeletal dysplasia and spinal degeneration (Hattori et al., 2017). The ACAN has been shown to be associated with height in humans (Fariello et al., 2014) and therefore a potential candidate gene for growth traits.

There are several already identified genes that are involved in the genetic orchestra that determines the expression of growth traits in different cattle breeds (Pausch et al., 2011; Peters et al., 2012; Saatchi et al., 2014; Santiago et al., 2017; Seabury et al., 2017; Terakado et al., 2018).

However, the current ORA through MeSH was able to translate genomic information summarized by the Tag-SNP selection strategy, providing biological meanings of more specific gene products related to the traits.

MeSH can be an interesting alternative to gene ontology (GO), which is a mainstay in functional genomics studies and ORA, because it includes categories not covered by GO and it has the potential to refine and enrich the representation of biological knowledge in complex traits (Morota et al., 2015). Some candidate genes detected in the present study by MeSH were found to be associated with growth traits in previous research and other genes are novel for BW, WW205, PWG345 and YW550. The identification of these genes and chromosome regions that affect growth traits in Hereford and Braford cattle is essential to understand biological mechanisms and genetic architecture of these traits in addition to being able to include the molecular markers information in genetic evaluations of beef cattle breeding programmes.

4 | CONCLUSION

The BayesB and BayesA methods were suitable for selecting Tag-SNPs from GWAS analyses of growth traits in Hereford and Braford cattle breeds. Based on the current strategy applied, markers that explained greatest part of the genetic variation of the traits under study were able to be defined, as well as the gene products and functions related to them through Medical Subject Headings.

Genomic predictions obtained with the proposed low-density panels using BayesA presented moderate prediction accuracies for growth traits. These panels may be useful for future fine mapping studies or gene expression to unravel the differences between the growth traits at weaning and yearling. In addition, low-density panels can be utilized to develop lower-cost commercial applications of genomic prediction.

ACKNOWLEDGEMENTS

Research supported by CNPq—National Council for Scientific and Technological Development grant 305102/2018-4 and Embrapa—Brazilian Agricultural Research Corporation grants 12.13.14.014.00. F.F. Cardoso, A. A. Boligon, and J. Braccini are CNPq research fellows. Authors acknowledge the Delta G Connection for providing performance data for this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ORCID

Gabriel Soares Campos https://orcid.org/0000-0002-7459-824X
Marcos Jun Iti Yokoo https://orcid.org/0000-0003-3821-1826

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the Conexao Delta G breeding programme. Restrictions apply to the availability of these data, which were used under licence for this study. Data may be obtained from the authors upon request with the permission of the owner breeders.

REFERENCES

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (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. https://doi.org/10.3168/jds.2009-2730

Ailhaud, G., Grimaldi, P., & Negrel, R. (1992). Cellular and molecular aspects of adipose tissue development. *Annual Review of Nutrition*, *12*(1), 207–233. https://doi.org/10.1146/annurev.nu.12.070192.001231

Aja, S., Sahandy, S., Ladenheim, E. E., Schwartz, G. J., & Moran, T. H. (2001). Intracerebroventricular CART peptide reduces food intake and alters motor behavior at a hindbrain site. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *281*, R1862–R1867. https://doi.org/10.1152/ajpregu.2001.281.6.R1862

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

Berglund, M. M., Schober, D. A., Statnick, M. A., McDonald, P. H., & Gehlert, D. R. (2003). The use of bioluminescence resonance energy transfer 2 to study neuropeptide Y receptor agonist-induced beta-arrestin 2 interaction. *Journal of Pharmacology and Experimental Therapeutics*, *306*, 147–156. https://doi.org/10.1124/jpet.103.051227

Biegelmeyer, P., Gulias-Gomes, C. C., Caetano, A. R., Steibel, J. P., & Cardoso, F. F. (2016). Linkage disequilibrium, persistence of phase and effective population size estimates in Hereford and Braford cattle. *BMC Genetics*, *17*(1), 32. https://doi.org/10.1186/s12863-016-0339-8

Bolormaa, S., Pryce, J. E., Kemper, K. E., Hayes, B. J., Zhang, Y., Tier, B., … Goddard, M. E. (2013). Detection of quantitative trait loci in *Bos indicus* and *Bos taurus* cattle using genome-wide association studies. *Genetics Selection Evolution*, *45*(1), 43. https://doi.org/10.1186/1297-9686-45-43

Brunkow, M. E., Gardner, J. C., Van Ness, J., Paeper, B. W., Kovacevich, B. R., Proll, S., … Mulligan, J. T. (2001). Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot–containing protein. *The American Journal of Human Genetics*, *68*(3), 577–589. https://doi.org/10.1086/318811

Buzanskas, M. E., Grossi, D. A., Ventura, R. V., Schenkel, F. S., Sargolzaei, M., Meirelles, S. L. C., … Munari, D. P. (2014). Genome-wide association for growth traits in Canchim beef cattle. *PLoS ONE*, *9*(4), e94802. https://doi.org/10.1371/journal.pone.0094802

Campos, G. S., Reimann, F. A., Cardoso, L. L., Ferreira, C. E. R., Junqueira, V. S., Schmidt, P. I., … Cardoso, F. F. (2018). Genomic prediction using different estimation methodology, blending and cross-validation techniques for growth traits and visual scores in Hereford and Braford cattle. *Journal of Animal Science*, *96*(7), 2579–2595. https://doi.org/10.1093/jas/sky175

Cardoso, F. F., Gomes, C. C. G., Sollero, B. P., Oliveira, M. M., Roso, V. M., Piccoli, M. L., … Aguilar, I. (2015). Genomic prediction for tick resistance in Braford and Hereford cattle. *Journal of Animal Science*, *93*(6), 2693–2705. https://doi.org/10.2527/jas.2014-8832

Cassar-Malek, I., Boby, C., Picard, B., Reverter, A., & Hudson, N. J. (2017). Molecular regulation of high muscle mass in developing Blonde d'Aquitaine cattle foetuses. *Biology Open*, *6*(10), 1483–1492. https://doi.org/10.1242/bio.024950

Clayton, D. (2014) *snpStats: SnpMatrix and XSnpMatrix classes and methods. R package version, 1.18*. Retrieved from http://www.bioconductor.org/packages/release/bioc/html/snpStats.html

Cochran, S. D., Cole, J. B., Null, D. J., & Hansen, P. J. (2013). Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genetics*, *14*(1), 49. https://doi.org/10.1186/1471-2156-14-49

Durinck, S., Spellman, P. T., Birney, E., & Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols*, *4*(8), 1184. https://doi.org/10.1038/nprot.2009.97

Eggen, A. (2012). The development and application of genomic selection as a new breeding paradigm. *Animal Frontiers*, *2*(1), 10–15. https://doi.org/10.2527/af.2011-0027

Fan, B., Onteru, S. K., Du, Z. Q., Garrick, D. J., Stalder, K. J., & Rothschild, M. F. (2011). Genome-wide association study identifies loci for body composition and structural soundness traits in pigs. *PLoS ONE*, *6*(2), e14726. https://doi.org/10.1371/journal.pone.0014726

Fariello, M. I., Servin, B., Tosser-Klopp, G., Rupp, R., Moreno, C., San Cristobal, M., … International Sheep Genomics Consortium (2014). Selection signatures in worldwide sheep populations. *PLoS ONE*, *9*(8), e103813. https://doi.org/10.1371/journal.pone.0103813

Fernando, R. L., Dekkers, J. C., & Garrick, D. J. (2014). A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. *Genetics Selection Evolution*, *46*(1), 50. https://doi.org/10.1186/1297-9686-46-50

Fernando, R. L., & Garrick, D. J. (2008). *GenSel - User manual for a portfolio of genomic selection related analyses*. Ames, IA: Animal Breeding and Genetics, Iowa State University. Retrieved from http://bigs.ansci.iastate.edu/bigsgui/login.html

Fernando, R. L., & Garrick, D. J. (2013). Bayesian methods applied to GWAS. In C. Condro, J. H. J. Van der Werf, & B. J. Hayes (Eds.), *Genome-wide association studies and genomic predictions* (pp. 237–274) Berlin, Germany: Springer.

Gambra, R., Peñagaricano, F., Kropp, J., Khateeb, K., Weigel, K. A., Lucey, J., & Khatib, H. (2013). Genomic architecture of bovine κ-casein and β-lactoglobulin. *Journal of Dairy Science*, *96*(8), 5333–5343. https://doi.org/10.3168/jds.2012-6324

Garrick, D. J., Taylor, J. F., & Fernando, R. L. (2009). Deregressing estimated breeding values and weighting information for genomic

regression analyses. *Genetics Selection Evolution*, *41*(1), 55. https://doi.org/10.1186/1297-9686-41-55

Geweke, J. (1992). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*. Minneapolis, MN: Federal Reserve Bank of Minneapolis, Research Department.

Graugnard, D. E., Piantoni, P., Bionaz, M., Berger, L. L., Faulkner, D. B., & Loor, J. J. (2009). Adipogenic and energy metabolism gene networks in longissimus lumborum during rapid post-weaning growth in Angus and Angus× Simmental cattle fed high-starch or low-starch diets. *BMC Genomics*, *10*(1), 142. https://doi.org/10.1186/1471-2164-10-142

Gudbjartsson, D. F., Walters, G. B., Thorleifsson, G., Stefansson, H., Halldorsson, B. V., Zusmanovich, P., … Stefansson, K. (2008). Many sequence variants affecting diversity of adult human height. *Nature Genetics*, *40*(5), 609. https://doi.org/10.1038/ng.122

Habier, D., Fernando, R. L., & Dekkers, J. C. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, *177*(4), 2389–2397. https://doi.org/10.1534/genetics.107.081190

Habier, D., Fernando, R. L., & Dekkers, J. C. (2009). Genomic selection using low-density marker panels. *Genetics*, *182*(1), 343–353. https://doi.org/10.1534/genetics.108.100289

Habier, D., Fernando, R. L., Kizilkaya, K., & Garrick, D. J. (2011). Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics*, *12*(1), 186. https://doi.org/10.1186/1471-2105-12-186

Hattori, A., Katoh-Fukui, Y., Nakamura, A., Matsubara, K., Kamimaki, T., Tanaka, H., Ida, S.. (2017). Next generation sequencing-based mutation screening of 86 patients with idiopathic short stature. *Endocrine Journal*, *64*, 947–954. https://doi.org/10.1507/endocrj.EJ17-0150

Hayes, B., & Goddard, M. (2010). Genome-wide association and genomic selection in animal breeding. *Genome*, *53*(11), 876–883. https://doi.org/10.1139/G10-076

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, *31*, 423–447. https://doi.org/10.2307/2529430

.Illumina (2006). *"TOP/BOT" strand and "A/B" allele: A guide to Illumina's method for determining strand and allele for the GoldenGate and Infinium assays*. Technical note. Retrieved from http://res.illumina.com/documents/products/technotes/technote_topbot.pdf

Jahuey-Martínez, F. J., Parra-Bracamonte, G. M., Sifuentes-Rincón, A. M., Martínez-González, J. C., Gondro, C., García-Pérez, C. A., & López-Bustamante, L. A. (2016). Genomewide association analysis of growth traits in Charolais beef cattle. *Journal of Animal Science*, *94*(11), 4570–4582. https://doi.org/10.2527/jas.2016-0359

Jin, W., Olson, E. N., Moore, S. S., Basarab, J. A., Basu, U., & Guan, L. L. (2012). Transcriptome analysis of subcutaneous adipose tissues in beef cattle using 3′ digital gene expression-tag profiling. *Journal of Animal Science*, *90*(1), 171–183. https://doi.org/10.2527/jas.2011-4229

Karisa, B. K., Thomson, J., Wang, Z., Bruce, H. L., Plastow, G. S., & Moore, A. S. (2013). Candidate genes and biological pathways associated with carcass quality traits in beef cattle. *Canadian Journal of Animal Science*, *93*(3), 295–306. https://doi.org/10.4141/cjas2012-136

Keogh, K., Kenny, D. A., Cormican, P., McCabe, M. S., Kelly, A. K., & Waters, S. M. (2016). Effect of dietary restriction and subsequent re-alimentation on the transcriptional profile of bovine skeletal muscle. *PLoS ONE*, *11*(2), e0149373. https://doi.org/10.1371/journal.pone.0149373

Legarra, A., Robert-Granié, C., Manfredi, E., & Elsen, J. M. (2008). Performance of genomic selection in mice. *Genetics*, *180*(1), 611–618. https://10.1534/genetics.108.088575 .

Li, C., Basarab, J., Snelling, W. M., Benkel, B., Murdoch, B., Hansen, C., & Moore, S. S. (2004). Assessment of positional candidate genes myf 5 and igf 1 for growth on bovine chromosome 5 in commercial lines of *Bos taurus*. *Journal of Animal Science*, *82*(1), 1–7. https://doi.org/10.2527/2004.8211

Linnemann, A., van der Ven, P. F. M., Vakeel, P., Albinus, B., Simonis, D., Bendas, G., … Fürst, D. O. (2010). The sarcomeric Z-disc component myopodin is a multiadapter protein that interacts with filamin and α-actinin. *European Journal of Cell Biology*, *89*(9), 681–692. https://doi.org/10.1016/j.ejcb.2010.04.004

Majumdar, A. S., Giri, P. R., & Pai, S. A. (2014). Resveratrol-and melatonin-abated ovariectomy and fructose diet–induced obesity and metabolic alterations in female rats. *Menopause*, *21*(8), 876–885. https://doi.org/10.1097/GME.0000000000000187

Maltecca, C., Weigel, K. A., Khatib, H., Cowan, M., & Bagnato, A. (2009). Whole-genome scan for quantitative trait loci associated with birth weight, gestation length and passive immune transfer in a Holstein× Jersey crossbred population. *Animal Genetics*, *40*(1), 27–34. https://doi.org/10.1111/j.1365-2052.2008.01793.x

McClure, M. C., Morsci, N. S., Schnabel, R. D., Kim, J. W., Yao, P., Rolf, M. M., … Taylor, J. F. (2010). A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Animal Genetics*, *41*(6), 597–607. https://doi.org/10.1111/j.1365-2052.2010.02063.x

Meuwissen, T. H., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*, 1819–1829.

Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2015). *Manual for BLUPF90 family of programs*. Retrieved from https://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf

Morota, G., Peñagaricano, F., Petersen, J. L., Ciobanu, D. C., Tsuyuzaki, K., & Nikaido, I. (2015). An application of Me SH enrichment analysis in livestock. *Animal Genetics*, *46*(4), 381–387. https://doi.org/10.1111/age.12307

Mota, R. R., Silva, F. F., Lopes, P. S., Tempelman, R. J., Sollero, B. P., Aguilar, I., & Cardoso, F. F. (2018). Analyses of reaction norms reveal new chromosome regions associated with tick resistance in cattle. *Animal*, *12*(2), 205–214. https://doi.org/10.1017/S1751731117001562

Nelson, S. J., Schopen, M., Savage, A. G., Schulman, J. L. A., & Arluk, N. (2004). The MeSH translation maintenance system: structure, interface design, and implementation. In *Medinfo* (pp. 67–69).

Onteru, S. K., Gorbach, D. M., Young, J. M., Garrick, D. J., Dekkers, J. C., & Rothschild, M. F. (2013). Whole genome association studies of residual feed intake and related traits in the pig. *PLoS ONE*, *8*(6), e61756. https://doi.org/10.1371/journal.pone.0061756

Pausch, H., Flisikowski, K., Jung, S., Emmerling, R., Edel, C., Götz, K. U., & Fries, R. (2011). Genome-wide association study identifies two major loci affecting calving ease and growth-related traits in cattle. *Genetics*, *187*(1), 289–297. https://doi.org/10.1534/genetics.110.124057

Pazik, J., Lewandowski, Z., Nowacka Cieciura, E., Ołdak, M., Podgórska, M., Sadowska, A., … Durlik, M. (2018). Malnutrition risk in kidney recipients treated with mycophenolate mofetil is associated with IMPDH1 rs2278294 polymorphism. In *Transplantation proceedings* (Vol. *50*, No. 6, pp. 1794–1797). Elsevier.

Pérez O'Brien, A. M., Utsunomiya, Y. T., Mészáros, G., Bickhart, D. M., Liu, G. E., Van Tassell, C. P., … Sölkner, J. (2014). Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genetics Selection Evolution*, *46*(1), 19. https://doi.org/10.1186/1297-9686-46-19

Peters, S. O., Kizilkaya, K., Garrick, D. J., Fernando, R. L., Reecy, J. M., Weaber, R. L., … Thomas, M. G. (2012). Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. *Journal of Animal Science*, *90*(10), 3398–3409. https://doi.org/10.2527/jas.2011-4507

Porto Neto, L. R., Bunch, R. J., Harrison, B. E., & Barendse, W. (2012). Variation in the XKR4 gene was significantly associated with subcutaneous rump fat thickness in indicine and composite cattle. *Animal Genetics*, *43*(6), 785–789. https://doi.org/10.1111/j.1365-2052.2012.02330.x

Pryce, J. E., Hayes, B. J., Bolormaa, S., & Goddard, M. E. (2011). Polymorphic regions affecting human height also control stature in cattle. *Genetics*, *187*(3), 981–984. https://doi.org/10.1534/genetics.110.123943

R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ramayo-Caldas, Y., Fortes, M. R. S., Hudson, N. J., Porto-Neto, L. R., Bolormaa, S., Barendse, W., … Reverter, A. (2014). A marker-derived gene network reveals the regulatory role of PPARGC1A, HNF4G, and FOXP3 in intramuscular fat deposition of beef cattle. *Journal of Animal Science*, *92*(7), 2832–2845. https://doi.org/10.2527/jas.2013-7484

Raven, L. A., Cocks, B. G., & Hayes, B. J. (2014). Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics*, *15*(1), 62. https://doi.org/10.1186/1471-2164-15-62

Saatchi, M., McClure, M. C., McKay, S. D., Rolf, M. M., Kim, J. W., Decker, J. E., … Taylor, J. F. (2011). Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. *Genetics Selection Evolution*, *43*(1), 40. https://doi.org/10.1186/1297-9686-43-40

Saatchi, M., Schnabel, R. D., Taylor, J. F., & Garrick, D. J. (2014). Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics*, *15*(1), 442. https://doi.org/10.1186/1471-2164-15-442

Salleh, M. S., Mazzoni, G., Höglund, J. K., Olijhoek, D. W., Lund, P., Løvendahl, P., & Kadarmideen, H. N. (2017). RNA-Seq transcriptomics and pathway analyses reveal potential regulatory genes and molecular mechanisms in high-and low-residual feed intake in Nordic dairy cattle. *BMC Genomics*, *18*(1), 258. https://doi.org/10.1186/s12864-017-3622-9

Santana, M. H. D. A., Utsunomiya, Y. T., Neves, H. H., Gomes, R. D. C., Garcia, J. F., Fukumasu, H., … Ferraz, J. B. S. (2014). Genome-wide association study for feedlot average daily gain in Nellore cattle (*Bos indicus*). *Journal of Animal Breeding and Genetics*, *131*(3), 210–216. https://doi.org/10.1111/jbg.12084

Santiago, G. G., Siqueira, F., Cardoso, F. F., Regitano, L. C. A., Ventura, R., Sollero, B. P., … Torres, R. A. A. (2017). Genomewide association study for production and meat quality traits in Canchim beef cattle. *Journal of Animal Science*, *95*(8), 3381–3390. https://doi.org/10.2527/jas.2017.1570

Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2011). FImpute: An efficient imputation algorithm for dairy cattle populations. *Journal of Dairy Science*, *94*, 421.

Seabury, C. M., Oldeschulte, D. L., Saatchi, M., Beever, J. E., Decker, J. E., Halley, Y. A., … Taylor, J. F. (2017). Genome-wide association study for feed efficiency and growth traits in US beef cattle. *BMC Genomics*, *18*(1), 386. https://doi.org/10.1186/s12864-017-3754-y

Sharma, N. K., Das, S. K., Mondal, A. K., Hackney, O. G., Chu, W. S., Kern, P. A., … Elbein, S. C. (2008). Endoplasmic reticulum stress markers are associated with obesity in nondiabetic subjects. *The Journal of Clinical Endocrinology and Metabolism*, *93*(11), 4532–4541. https://doi.org/10.1210/jc.2008-1001

Smith, B. J. (1997). BOA: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistics Software*, *21*, 1–37. https://doi.org/10.18637/jss.v021.i11

Snelling, W. M., Allan, M. F., Keele, J. W., Kuehn, L. A., Mcdaneld, T., Smith, T. P. L., … Bennett, G. L. (2010). Genome-wide association study of growth in crossbred beef cattle. *Journal of Animal Science*, *88*(3), 837–848. https://doi.org/10.2527/jas.2009-2257

Snelling, W. M., Cushman, R. A., Fortes, M. R. S., Reverter, A., Bennett, G. L., Keele, J. W., … Thomas, M. G. (2012). How single nucleotide polymorphism chips will advance our knowledge of factors controlling puberty and aid in selecting replacement beef females. *Journal of Animal Science*, *90*(4), 1152–1165. https://doi.org/10.2527/jas.2011-4581

Sollero, B. P., Junqueira, V. S., Gomes, C. C., Caetano, A. R., & Cardoso, F. F. (2017). Tag SNP selection for prediction of tick resistance in Brazilian Braford and Hereford cattle breeds using Bayesian methods. *Genetics Selection Evolution*, *49*(1), 49. https://doi.org/10.1186/s12711-017-0325-2

Soumillion, A., Erkens, J. H., Lenstra, J. A., Rettenberger, G., & te Pas, M. F. (1997). Genetic variation in the porcine myogenin gene locus. *Mammalian Genome*, *8*(8), 564–568. https://doi.org/10.1007/s003359900504

Terakado, A. P. N., Costa, R. B., de Camargo, G. M. F., Irano, N., Bresolin, T., Takada, L., … de Albuquerque, L. G. (2018). Genome-wide association study for growth traits in Nelore cattle. *Animal*, *12*(7), 1358–1362. https://doi.org/10.1017/S1751731117003068

Tizioto, P. C., Decker, J. E., Taylor, J. F., Schnabel, R. D., Mudadu, M. A., Silva, F. L., … Regitano, L. C. A. (2013). Genome scan for meat quality traits in Nelore beef cattle. *Physiological Genomics*, *45*(21), 1012–1020. https://doi.org/10.1152/physiolgenomics.00066.2013

Utsunomiya, Y. T., do Carmo, A. S., Carvalheiro, R., Neves, H. H. R., Matos, M. C., Zavarez, L. B., … Garcia, J. F. (2013). Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. *BMC Genetics*, *14*(1), 52. https://doi.org/10.1186/1471-2156-14-52

van den Berg, I., Boichard, D., & Lund, M. S. (2016). Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. *Journal of Dairy Science*, *99*, 8932–8945. https://doi.org/10.3168/jds.2016-11073

Van Eenennaam, A. L., Weigel, K. A., Young, A. E., Cleveland, M. A., & Dekkers, J. C. (2014). Applied animal genomics: Results from the field. *Annual Review of Animal Biosciences*, *2*(1), 105–139. https://doi.org/10.1146/annurev-animal-022513-114119

Wang, H., Zheng, Y., Wang, G., & Li, H. (2013). Identification of microRNA and bioinformatics target gene analysis in beef cattle intramuscular fat and subcutaneous fat. *Molecular bioSystems*, *9*(8), 2154–2162. https://doi.org/10.1039/c3mb70084d

Wang, W. Y., Barratt, B. J., Clayton, D. G., & Todd, J. A. (2005). Genome-wide association studies: Theoretical and practical concerns. *Nature Reviews Genetics*, *6*(2), 109. https://doi.org/10.1038/nrg1522

Weng, Z., Su, H., Saatchi, M., Lee, J., Thomas, M. G., Dunkelberger, J. R., & Garrick, D. J. (2016). Genome-wide association study of growth and body composition traits in Brangus beef cattle. *Livestock Science*, *183*, 4–11. https://doi.org/10.1016/j.livsci.2015.11.011

Wolc, A., Arango, J., Settar, P., Fulton, J. E., O'Sullivan, N. P., Preisinger, R., … Dekkers, J. C. M. (2011). Persistence of accuracy of genomic estimated breeding values over generations in layer chickens. *Genetics Selection Evolution*, *43*(1), 23. https://doi.org/10.1186/1297-9686-43-23

Wortley, K. E., Chang, G. Q., Davydova, Z., Fried, S. K., & Leibowitz, S. F. (2004). Cocaine-and amphetamine-regulated transcript in the arcuate nucleus stimulates lipid metabolism to control body fat accrual on a high-fat diet. *Regulatory Peptides*, *117*(2), 89–99. https://doi.org/10.1016/j.regpep.2003.08.005

Yang, W., Wang, Y., Fu, C., & Zan, L. S. (2015). Association study and expression analysis of MTNR1A as a candidate gene for body measurement and meat quality traits in Qinchuan cattle. *Gene*, *570*(2), 199–204. https://doi.org/10.1016/j.gene.2015.06.012

Zare, Y., Shook, G. E., Collins, M. T., & Kirkpatrick, B. W. (2014). Genome-wide association analysis and genomic prediction of Mycobacterium avium subspecies paratuberculosis infection in US Jersey cattle. *PLoS ONE*, *9*(2), e88380. https://doi.org/10.1371/journal.pone.0088380

Zhang, C. L., Chen, H., Wang, Y. H., Lan, X. Y., Lei, C. Z., & Fang, X. T. (2008). The polymorphisms of bovine cocaine-and amphetamine-regulated transcripts and their associations with cattle (*Bos taurus*) growth traits. *Journal of Biosciences*, *33*(3), 365–370. https://doi.org/10.1007/s12038-008-0056-1

Zhang, L. P., Gan, Q. F., Hou, G. Y., Gao, H. J., Li, J. Y., & Xu, S. Z. (2015). Investigation of TG gene variants and their effects on growth, carcass composition, and meat quality traits in Chinese steers. *Genetics and Molecular Research*, *14*(2), 5320–5326. https://doi.org/10.4238/2015.May.22.2

Zimin, A. V., Delcher, A. L., Florea, L., Kelley, D. R., Schatz, M. C., Puiu, D., … Salzberg, S. L. (2009). A whole-genome assembly of the domestic cow, *Bos Taurus*. *Genome Biology*, *10*(4), R42. https://doi.org/10.1186/gb-2009-10-4-r42

How to cite this article: Campos GS, Sollero BP, Reimann FA, et al. Tag-SNP selection using Bayesian genomewide association study for growth traits in Hereford and Braford cattle. *J Anim Breed Genet*. 2019;00:1–19. https://doi.org/10.1111/jbg.12458