Received: 1 August 2019 | Revised: 18 November 2019 | Accepted: 24 November 2019

DOI: 10.1111/jbg.12463

ORIGINAL ARTICLE

Journal of Animal Breeding and Genetics WILEY

Comparison between haplotype-based and individual snp-based genomic predictions for beef fatty acid profile in Nelore cattle

Fabieli Loise Braga Feitosa[1] | **Angélica Simone Cravo Pereira**[2] | **Sabrina Thaise Amorim**[1] | **Elisa Peripolli**[1] | **Rafael Medeiros de Oliveira Silva**[3] | **Camila Urbano Braz**[1] | **Adrielle Matias Ferrinho**[2] | **Flavio Schramm Schenkel**[4] | **Luiz Fernando Brito**[4] | **Rafael Espigolan**[5] | **Lucia Galvão de Albuquerque**[1] | **Fernando Baldi**[1]

[1]Faculdade de Ciências Agrárias e Veterinárias, Departamento de Zootecnia, Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Jaboticabal, Brazil

[2]Faculdade de Zootecnia e Engenharia de Alimentos, Departamento de Nutrição e Produção Animal, Universidade de São Paulo, Pirassununga, Brazil

[3]Zoetis, Kalamazoo, MI, USA

[4]Department of Animal Sciences, Purdue University, West Lafayette, IN, USA

[5]Faculdade de Zootecnia e Engenharia de Alimentos, Departamento de Medicina Veterinária, Universidade de São Paulo, Pirassununga, Brazil

Correspondence
Fernando Baldi, Faculdade de Ciências Agrárias e Veterinárias, Departamento de Zootecnia, Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Jaboticabal 14884-900 Brazil.
Email: fernandobaldiuy@gmail.com

Funding information
This work was supported by the Sao Paulo Research Foundation (FAPESP), with Grant #2009/16118-5 for financial support and Grants #2015/25304-8 and #2016/24085-3 for the first author.

Abstract

The aim of this study was to evaluate the genomic predictions using the single-step genomic best linear unbiased predictor (ssGBLUP) method based on SNPs and haplotype markers associated with beef fatty acids (FAs) profile in Nelore cattle. The data set contained records from 963 Nelore bulls finished in feedlot (±90 days) and slaughtered with approximately 24 months of age. Meat samples from the *Longissimus dorsi* muscle were taken for FAs profile measurement. FAs were quantified by gas chromatography using a SP-2560 capillary column. Animals were genotyped with the high-density SNP panel (BovineHD BeadChip assay) containing 777,962 markers. SNPs with a minor allele frequency and a call rate lower than 0.05 and 0.90, respectively, monomorphic, located on sex chromosomes, and with unknown position were removed from the data set. After genomic quality control, a total of 469,981 SNPs and 892 samples were available for subsequent analyses. Missing genotypes were imputed and phased using the FImpute software. Haplotype blocks were defined based on linkage disequilibrium using the Haploview software. The model to estimate variance components and genetic parameters and to predict the genomic values included the random genetic additive effects, fixed effects of the contemporary group and the age at slaughter as a linear covariate. Accuracies using the haplotype-based approach ranged from 0.07 to 0.31, and those SNP-based ranged from 0.06 to 0.33. Regression coefficients ranged from 0.07 to 0.74 and from 0.08 to 1.45 using the haplotype- and SNP-based approaches, respectively. Despite the low to moderate accuracies for the genomic values, it is possible to obtain genetic progress trough selection using genomic information based either on SNPs or haplotype markers. The SNP-based approach allows less biased genomic evaluations, and it is more feasible when taking into account the computational and operational cost underlying the haplotypes inference.

KEYWORDS

Bos taurus indicus, genetic markers, heritability, meat quality, ssGBLUP

1 | INTRODUCTION

Beef is a high nutritional food, being a rich source of protein, iron, zinc, complex B vitamins, and essential polyunsaturated fatty acids (PUFA) such as linoleic, linolenic and arachidonic acids (McNeill & Van Elswyk, 2012). It is worth to highlight that beef is also considered a harmful source of saturated fats, which have been associated with coronary heart diseases, diabetes, obesity and cancer (Enser, Hallett, Hewitt, Fursey, & Wood, 1996; Jakobsen, 1999). Consumer's demand for meat nutritional traits has changed in recent years, and an increasing concern about beef fat composition has been recurrent among them. In this regard, studies encompassing beef cattle genomic selection for nutritional quality traits could be a strategy for a beef market differentiation so as to improve competitiveness with other protein sources.

In beef cattle, the additive genetic component plays an important role in beef fat profile for taurine and indicine cattle (Cesar et al., 2014; Feitosa et al., 2017; Nogi, Honda, Mukai, Okagaki, & Oyama, 2011; Pitchford, Deland, Siebert, Malau-Aduliand, & Bottema, 2002; Tait et al., 2007). Several studies identified genomic regions and candidate genes involved in metabolic pathways associated with beef FAs profile in indicine cattle (Berton et al., 2016; Cesar et al., 2014; Lemos et al., 2016), providing better support to elucidate the genetic basis underlying the beef nutritional composition and human health. Nevertheless, there are limited genomic selection studies for beef FAs profile in cattle (Chen et al., 2015; Chiaia et al., 2017; Onogi et al., 2015; Saatchi et al., 2013) and there are some divergences among them, suggesting that differences in the evaluated genomic methods are due to the genetic architecture of the trait, that is accuracy tends to increase as the model fits the genetic architecture of the trait (Lund, Sahana, de Koning, Su, & Carlborg, 2009). Moreover, the mechanisms underlying the associations remain largely unaddressed since the map resolution for genome-wide association studies is limited by the complicated linkage disequilibrium (LD) structure of the genome and the sampling variation in statistical tests due to finite sample sizes (Wu et al., 2018).

The analysis based on haplotype blocks can potentially identify loci that are not captured by one single marker or identify the joined effect of two or more loci. Haplotypes can be defined as a combination of alleles at adjacent loci belonging to the same chromosome that are transmitted jointly. Haplotype markers have higher LD levels with mutation (QTL, quantitative trait locus) than does SNP markers (Bickel, Kopp, & Nuzhdin, 2011; Fallin & Schork, 2000). According to these authors, the haplotype approach may lead to the identification of a larger aggregate effect due to the combination of several mutations in a chromosomal region, which increases the power to identify loci even if they have minor effects. In this regard, Cuyabano, Su, and Lund (2015) compared the predictive ability of models using individual SNPs and haploblocks and concluded that the haplotypes improved the prediction accuracy, being a potential predictor for complex traits.

The single-step genomic best linear unbiased predictor procedure (ssGBLUP, Aguilar et al., 2010; Legarra, Aguilar, & Misztal, 2009) combines the pedigree-based and genomic relationship matrices into a single matrix (**H**) to predict the genomic estimated breeding values (GEBV). Although the usual multi-step genomic evaluations mostly rely on highly reliable sires as a reference population, the single-step approach includes genomic data into the traditional estimated breeding values (EBV) analysis, which contains all the phenotyped animals (Su et al., 2012). The single-step method does not explicitly divide the population into training group (reference population) and prediction group (validation population), but instead, the genomic data are included along with the phenotypic data and the pedigree relationship information (Aguilar et al., 2010; Christensen & Lund, 2010). Silva et al. (2016) stated that the ssGBLUP provided the most accurate predictions, and it should be considered as an option to simplify the genomic evaluations, especially for low heritability traits and those with unreliable EBV in the training populations. The genomic selection is a potential tool to increase the genetic response for those traits that are difficult to measure, such as beef FAs profile; however, the most suitable marker to evaluate the beef FAs profile whether it is SNP or haplotype-based is still being studied. The objective of this study was to evaluate the feasibility of the genomic selection for beef FAs profile through accuracy and bias of genomic prediction using the ssGBLUP method based on SNPs and haplotype markers in Nelore cattle.

2 | MATERIALS AND METHODS

2.1 | Ethics approval

This study was approved by the Ethics Committee on Animal Use (CEUA) of the Faculty of Agrarian and Veterinary Sciences, Sao Paulo State University (UNESP), Jaboticabal, Brazil, under Protocol Number 18340/16.

2.2 | Local, animals and management

Data from 963 Nelore bulls slaughtered with roughly 24 months of age and finished in the feedlot for 90 days were collected. The animals belonged to eight different farms located in the Southeast, Northeast and Midwest regions of Brazil, which participated in three beef cattle breeding programmes (Nelore Qualitas, Paint CRV Lagoa, and DeltaGen).

Breeding seasons were adopted at different periods on these farms. Therefore, calving seasons concentrated from August to October in some farms and from November to January in others. Bulls were weaned at seven months of age.

Animals were raised on grazing conditions using *Brachiaria sp*. and *Panicum sp*. forages with a density varying from 1.2 to 1.6 animal unit/hectare. Additionally, the animals had free access to mineral salt. After yearling, the breeding animals were selected, and the remaining ones were kept in feedlot conditions. During the feedlot, the forage and concentrate ratio ranged from 50:50 to 70:30, depending on the farm. In general, whole-plant corn or sorghum silage was used as high-quality forage. Grains of corn and/or sorghum, soya beans, soya bean meal or sunflower seeds were used as protein concentrates. The criterion used by farmers for slaughtering was the weight (500 to 550 kg). The slaughtering was performed on commercial slaughterhouses following the Brazilian Federal Inspection Service (SIF), and in accordance with the commercial standard procedures. After stored for 48 hours at 0 to 2°C, beef samples from the *Longissimus dorsi* muscle collected between the 12^{th} and 13^{th} ribs of the left half-carcasses were removed, placed in plastic bags and stored at −80°C for the FAs measurement.

2.3 | Fatty acid profile

The FAs profile was performed at the Meat Science Laboratory (LCC) in the Department of Animal Nutrition and Production at FMVZ/USP. Meat FAs were extracted from intramuscular fat (IMF) of the *Longissimus dorsi* muscle using the methodology described by Folch, Lees, and Stanley (1957), and the methyl esters were formed according to Kramer et al. (1997). The FAs were quantified by gas chromatography (GC-2010 Plus AOC 20i auto-injector; Shimadzu) using a SP-2560 capillary column (100 m × 0.25 mm I.D. × 0.02 mm; Supelco). The initial temperature was set to 70°C and gradually increased up to 175°C (13°C/min) and held for 27 min until a further increase to 215°C (4°C/min) and held again for 31 min. Hydrogen (H_2) was used as the carrier gas flow (40 cm^3/s). The temperature used by the flame ionization detector (FID) was 250°C, H_2 flows of 40 ml/min, air flows of 400 ml/min, make-up of 30 ml/min kPa nitrogen (N_2) and sampling rate of 40 milliseconds. The total running time of each sample (stop time) was 86 min. The FAs were identified by comparing the retention time of sample methyl esters with the FAs standard C4-C24 (F.A.M.E mix Sigma®) and GLC 463 Reference Mixture Nu Check, vaccenic acid C18:1 trans-11 (V038-1G, Sigma®) C18:2 trans-10 cis-12 (UC-61M 100 mg), CLA and C18:2 cis-9, trans-11 (UC-60M 100 mg, Sigma®), tricosanoic acid (Sigma®) and nonadecanoic acid (Sigma®). FAs were quantified by normalizing the area under the curve of methyl esters using the GS 2.42 software. FAs contents were expressed as percentage of the total fatty acid methyl ester quantified.

The sum of the saturated FAs (SFA, C10:0 + C11:0 + C12:0 + C13:0 + C14:0 + C15:0 + C16:0 + C17:0 + C18:0 + C21:0 + C24:0), monounsaturated FAs (MUFA, C16:1 + C17:1 c10 + C18:1 t11 + C15:1 c10 + C20:1 c11 + C24:1 + C22:1 n9 + C18:1n9c + C14:1 + C18:1 n9t), polyunsaturated FAs (PUFA, C18:2 n6 + C18:3 n3 + C18:3 n6 + C20:3 n3 cis-11, 14, 17 + C20:3 n6 cis-8, 11, 14 + C20:4 n6 + C20:5 n3 + C22:6 n3), omega-6 (n6, C18:3 n6 + C20:3 n6 c8, c11, c14 + C18:2 n6 + C20:4 n6) and omega-3 (n3, C18:3 n3 + C20:3 n3 c11, c14, c17 + C22:6 n3 + C20:5 n3) was calculated. The PUFA:SFA and n6:n3 ratios were also calculated.

2.4 | Genotyping

A total of 963 animals were genotyped for 777,962 SNPs using the high-density SNP panel (BovineHD BeadChip assay; Illumina Inc.). The quality control of the markers consisted of excluding those with unknown genomic position, located on sex chromosomes, with a call rate lower than 0.90 and minor allele frequency lower than 0.05, out of Hardy–Weinberg equilibrium ($p < 10^{-6}$), and with excess of heterozygosity. Additionally, samples with a call rate lower than 0.90 were also removed from the data set. After the genomic quality control, a total of 469,981 SNPs and 893 samples were retained for the subsequent analyses. The PREGSF90 software (Misztal et al., 2002) was used for SNP quality control.

2.5 | Haplotype construction and selection

Missing genotypes were imputed and then phased to haplotypes using the FImpute software version 2.2 (Sargolzaei, Chesnais, & Schenkel, 2014). Haplotype blocks were defined using the Haploview software (Barrett, Fry, Maller, & Daly, 2005) adopting a method based on LD (Gabriel et al., 2002), in which a 95% of confidence bounds based on D' are generated and each comparison is called "strong LD" (D' between 0.70 and 0.98). Only haplotypes alleles with a frequency higher than 1% were considered for the analyses.

2.6 | Genomic prediction analyses

The contemporary groups (CG) included animals born in the same year and farm and raised in the same management group at yearling. The CG with less than three animals and CG with less than two known sires were deleted. Records spanning over three standard deviations (plus/minus) from the CG mean were excluded. After data editing, a total of 801, 803, 798, 799, 797, 795 and 802 animals with records for SFA, MUFA, PUFA, n3, n6, PUFA:SFA ratio and n6:n3 ratio remained in the data set, respectively.

TABLE 1 Descriptive statistics and heritability estimates (h^2) for beef fatty acid profile in Nelore cattle using haplotypes

Trait	$h^2 \pm SD$	*N*	Mean	Minimum	Maximum
Sum of SFA	0.09 ± 0.05	801	41.00	22.32	52.05
Sum of MUFA	0.35 ± 0.09	803	38.00	17.43	61.78
Sum of PUFA	0.11 ± 0.06	798	13.19	2.07	30.12
Sum of n3	0.06 ± 0.03	799	3.77	0.24	8.38
Sum of n6	0.12 ± 0.05	797	9.10	0.74	22.22
PUFA:SFA ratio	0.07 ± 0.04	795	0.33	0.04	0.94
n6:n3 ratio	0.02 ± 0.02	802	2.46	0.12	5.12

Abbreviations: N, number of animals with records; SD, standard deviation; Mean, concentration of the fatty acids expressed as a percentage of total fatty acid methyl esters quantified; SFA, saturated fatty acid; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid; n3, omega-3; n6, omega-6.

The variances and genetic parameters were estimated applying the restricted maximum likelihood (REML) method using the AIREMLF90 software (Misztal et al., 2015, 2002), and considering an animal linear model. The genomic predictions were performed using the ssGBLUP (Aguilar et al., 2010). The model used to estimate variance components, genetic parameters and to predict the genomic values for the beef FAs profile included the random genetic additive effects, the fixed effects of the CG and the age at slaughter as a linear covariable. The effect of dam breeding season and dam age was tested in the model; however, non-significant effects were observed. For all traits, the model can be represented by the following matrix form:

$$y = Xb + Zu + e$$

where y is the vector of phenotypes, b is the vector of the fixed effects, u is the vector of the additive genetic effects, **X** and **Z** are the incidence matrices, and e is the vector of random residuals. An infinitesimal model was assumed, $var(u) = H\sigma_u^2$, in which **H** is the numerator relationship matrix obtained from the pedigree and genomic information and σ_u^2 is the variance of the additive genetic effects. Applying the ssGBLUP, the inverse of the numerator relationship matrix ($\mathbf{A}^{-1}$) was replaced by $\mathbf{H}^{-1}$ that combines pedigree and genomic information. The $\mathbf{H}^{-1}$ was constructed according to Aguilar et al. (2010), as shown below:

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

where, $\mathbf{H^{-1}}$ is the inverse of the relationship matrix that incorporates the inverse of the genomic relationship matrix ($\mathbf{G^{-1}}$) and the inverse of the pedigree relationship matrix among the genotyped animals $\left(\mathbf{A_{22}^{-1}}\right)$. The **G** matrix was created according to VanRaden (2008):

$$\mathbf{G} = \frac{(\mathbf{M} - \mathbf{P})(\mathbf{M} - \mathbf{P})'}{2\sum_{j=1}^{m} p_j (1 - p_j)},$$

where, **M** is a matrix of marker alleles with *m* columns (*m* is the total number of haplotypes or SNPs alleles) and *n* rows (*n* is the total number of the genotyped individuals), and **P** is a matrix containing the frequency of the second allele (p_j), expressed as $2p_j$. When using the haplotype-based approach, $\mathbf{M_{ij}}$ was 0 if the haplotype of individual *i* for haplotype *j* was homozygous for one allele, 1 if heterozygous or 2 if the haplotype was homozygous for both alleles. When the SNP-based approach was used, $\mathbf{M_{ij}}$ was 0 if the genotype of individual *i* for SNP *j* was homozygous for the first allele, 1 if heterozygous or 2 if the genotype was homozygous for the second allele.

Forward validation was used in the analyses. The youngest animals according to month and year of birth were set as the validation population ($N = 160$), and the remaining ones were used as the training population (varied from 635 to 642 animals). The prediction ability was computed as the correlation between the EBV and GEBV. The regression of the EBV on GEBV was used as an inflation measurement of the prediction method, in which a regression coefficient equal to one denotes no inflation.

After pruning, 2,989 animals were used to compute the relationship matrix and the data set for the beef FA profile contained records from 606 sires and 1,200 dams parents of progenies with phenotypic records.

3 | RESULTS AND DISCUSSION

3.1 | Heritability estimates

The heritability estimates for the SFA, MUFA, PUFA, n3, n6, n6:n3 ratio and PUFA:SFA ratio varied from low to moderate magnitude (Table 1). Lemos et al. (2016) working with the same data set applied the ssGBLUP method with a SNP-based approach and reported similar heritability estimates for the SFA (0.12), PUFA (0.08), n3 (0.11), n6:n3 ratio (0.07), and PUFA:SFA (0.11) ratio. Dissimilar values for the MUFA (0.20) and n6 (0.23) were also described by the authors. Cesar et al. (2014) using the GBLUP

method with a SNP-based approach obtained low heritability estimates in Nelore cattle for the SFA (0.11) and moderate for MUFA (0.14), PUFA (0.15), n3 (0.17) and n6 (0.15). Studies in taurine breeds (Enser et al., 1996; Inoue, Kobayashi, Shoji, & Kato, 2011) reported higher heritability estimates for these FA groups, with values of 0.47 for PUFA, 0.35 to 0.66 for SFA and 0.35 to 0.68 for MUFA. Kelly, Tume, Newman, and Thompson (2013) estimated higher heritability estimates for the SFA (0.54) and MUFA (0.54) and concluded that there is enough genetic variation to respond to selection regarding the FA profile of subcutaneous fat in cattle. The results of this study pointed out that selection to improve the beef FA profile in Nelore cattle is feasible. However, the elevated costs to obtain the phenotypic records together with the fact that this trait can be only obtained after the slaughter limits the genetic improvement through traditional selection.

3.2 | Haplotype blocks and genomic predictions

A total of 84,395 haplotypes blocks were constructed, and the number of haplotypes blocks and alleles decreased with chromosome length (Table 2). Cuyabano, Su, and Lund (2014) reported a similar number of haplotype blocks with a D' = 0.75 (n = 84,395) and alleles (n = 325,268) than did our results. Hess, Druet, Hess, and Garrick (2017) evaluated fixed-length haplotype alleles (from 125 kb to 2 Mb) with varying allele frequency thresholds (from 1% to 10%) in an admixed New Zealand dairy cattle population genotyped with the Illumina BovineSNP50 BeadChip and reported lower number of haplotype blocks (n = 17,452) with a haploblock length of 125 kb. Using simulated data from dairy cattle, Villumsen, Janss, and Lund (2009) reported higher reliabilities for GEBV when the haploblock length was set to 1 cM (10 SNP haplotypes), with a heritability ranging from 0.02 to 0.30. According to Hess et al. (2017), the optimal haplotype length needs to be evaluated independently for each data set and it should be taken into consideration the purpose of the analysis, that is shorter length haplotype for QTL mapping or longer for genomic prediction.

The predictive ability of the ssGBLUP using the haplotype-based approach was low to moderate, ranging from 7 to 31 percentage points (Table 3). The lowest predictive ability was described for the SFA and the highest for the n6:n3 ratio. For the MUFA, PUFA, PUFA:SFA ratio, n3 and n6, the prediction ability was 12, 18, 20, 19 and 26 percentage points, respectively. The results were similar using the SNP-based approach, with values ranging from 6 to 33 percentage points. Interestingly, the use of the haplotype-based approach did not improve the prediction ability for the beef FA profile. However, Cuyabano et al. (2014) compared the genomic predictions between the haplotype-(constructed based on LD and using high-density panels) and SNP-based approach for milk protein, fertility, and mastitis in dairy cattle and reported higher prediction ability using the haplotype-based approach. Moreover, Hess et al. (2017) performed genomic predictions fitting covariates for either SNP or haplotypes using BayesA, BayesB and BayesN, and concluded that fitting covariates for haplotype alleles rather than SNPs may increase the prediction accuracy up to 5.5% although it decreased drastically for long (>500 kb) haplotypes.

Recently, Karimi, Sargolzaei, Robinson, and Schenkel (2018) carried out a study to assess the possible advantages of using haplotype-based GBLUP compared with individual SNP-based GBLUP in terms of reliability and bias of genomic predictions in Holstein cattle. For most of the 57 traits analysed, the predictive reliability of the haplotype-based GBLUP was found to be equal to or slightly better than did the SNP-based. According to these authors, the results revealed that the effect of the alternate relationship matrices had a significant interaction with the level of the trait's heritability since higher reliability gains by using the haplotype-based GBLUP were obtained for high-heritable traits compared with low-heritable traits. Moreover, previous studies described the predictive abilities of genomic selection for beef FA profile using the SNP-based approach ranging from −0.06 to 0.57 in American Angus cattle (Saatchi et al., 2013), −0.05 to 0.73 in Canadian beef cattle (Chen et al., 2015), 0.44 to 0.72 in Japanese Black cattle (Onogi et al., 2015) and 0.03 to 0.51 in Chinese Simmental cattle (Zhu et al., 2017). In most of these studies, except for the Japanese Black cattle, the accuracies of the genomic prediction were low to moderate concurring with those obtained in the present study. In addition, Chiaia et al. (2017) who studied the same data set as did the present study, evaluated the prediction ability of four methods using SNP markers (SNP-BLUP, Bayesian Lasso, BayesC and BayesCπ) and reported genomic prediction accuracies ranging from 7 to 50 percentage points. Some results obtained by Chiaia et al. (2017) were similar to those displayed in this study, such as those for MUFA (0.07), PUFA (0.24), PUFA:SFA ratio (0.32) and n6 (0.22).

The regression coefficient encompassing the observed and predicted values is a practical measure regarding the method's ability to perform unbiased predictions. In this regard, a coefficient less than one indicates that the genetic values are overestimated and exhibit higher variability than expected, whereas coefficients higher than one indicate that the estimated genetic values exhibit less than expected variability (Wiggans, Cooper, VanRaden, & Cole, 2011). The estimated regression coefficients (bias) ranged from 0.07 to 0.74 for the haplotype-based approach (Table 3), indicating an inflation of the genomic prediction variance.

TABLE 2 Descriptive statistics for the number of haplotypes blocks and alleles per chromosome

BTA	Haplotypes blocks	Haplotypes alleles	Mean (kb)	Minimum (kb)	Maximum (kb)	Coverage (%)
1	5,239	20,259	17.000	0.350	603.047	56.25
2	4,499	17,413	17.058	0.454	299.121	55.99
3	4,167	16,397	17.011	0.044	257.827	58.37
4	4,012	15,259	16.157	0.470	613.343	53.65
5	3,626	14,058	19.354	0.151	547.956	57.91
6	4,075	16,414	18.157	0.458	277.842	61.94
7	3,650	14,343	17.762	0.004	662.438	57.56
8	3,993	16,020	17.006	0.420	543.382	59.89
9	3,713	14,705	17.176	0.467	487.080	60.33
10	3,249	12,274	16.297	0.464	418.942	50.76
11	3,457	13,065	16.556	0.376	352.740	53.34
12	2,894	11,157	16.740	0.443	392.517	53.14
13	2,808	10,744	16.786	0.035	405.403	55.95
14	3,103	12,511	16.538	0.092	335.174	60.63
15	2,877	10,973	15.612	0.098	282.611	52.66
16	2,743	10,563	15.856	0.229	233.669	53.22
17	2,721	10,425	15.161	0.452	665.442	54.89
18	2,108	8,142	16.593	0.225	363.815	52.99
19	2,069	7,606	14.268	0.140	314.829	46.08
20	2,417	9,143	15.305	0.442	250.540	51.35
21	2,448	9,368	16.285	0.272	999.176	55.68
22	2,003	7,503	15.125	0.474	258.122	49.31
23	1,902	7,188	13.097	0.167	739.807	47.42
24	2,148	8,258	15.402	0.316	481.986	52.75
25	1,508	5,620	15.292	0.658	253.137	53.75
26	1,892	7,171	14.284	0.466	432.115	52.29
27	1,653	6,202	13.714	0.130	504.407	49.92
28	1,653	6,088	12.901	0.486	181.162	46.05
29	1,768	6,400	13.508	0.126	289.930	46.37
Total	84,395	325,269	15.931	0.307	429.226	53.81

Abbreviation: BTA, *Bos taurus* autosome.

For the SNP-based approach, the regression coefficients ranged from 0.10 to 1.45, indicating an inflation of the variance of genomic prediction for all FA but PUFA:SFA ratio. When comparing the regression coefficients obtained with the ssGBLUP using either the haplotype or SNP-based approach, the results were very close except for PUFA:SFA ratio. For the haplotype and SNP-based approaches, the less biased results were for PUFA:SFA ratio, n3 and PUFA, and the more biased ones were for MUFA, SFA and n6. Hess et al. (2017), using haploblocks longer than 500 kb in length, tended to decrease the accuracy and increase the bias of the haplotype model compared with the SNP model, especially when using a higher haplotype allele frequency threshold (>5%). Chiaia et al. (2017) also showed biased results for MUFA (0.37), PUFA (2.28) and PUFA:SFA ratio (2.35).

The low to moderate genomic prediction ability and biased regression coefficients obtained for the beef FA profile are justified mainly by the low heritability estimates obtained for most of the FA as well as the small training population and the pedigree records with missing information (multiple-sires mating system) used in this study. Cuyabano et al. (2014) reported similar prediction ability considering a haplotype and SNP-based approaches. Hess et al. (2017) observed that long haploblocks, that is ≥1 Mb in length with at least 15 SNPs per haploblock on average, showed lower accuracies than did the SNP model for milk dairy traits in New Zealand admixture dairy cattle population.

Overall, the beef FA profile of intramuscular fat is a high-cost trait and it is difficult to evaluate since it

TABLE 3 Prediction ability (r) and bias (b) for beef fatty acid profile using the haplotype or SNP-based approach and the ssGBLUP method

Trait	$r_{(HAPLO)}$	$b_{(HAPLO)}$	$r_{(SNP)}$	$b_{(SNP)}$
Sum of SFA	0.072	0.126	0.062	0.107
Sum of MUFA	0.119	0.077	0.100	0.080
Sum of PUFA	0.186	0.467	0.181	0.549
Sum of n3	0.207	0.645	0.213	0.596
Sum of n6	0.197	0.254	0.189	0.311
PUFA:SFA ratio	0.263	0.745	0.272	1.453
n6:n3 ratio	0.311	0.255	0.333	0.538

Abbreviations: $r_{(HAPLO)}$, correlation between EBV and GEBV using the haplotype-based approach; $r_{(SNP)}$, correlation between EBV and GEBV using the SNP-based approach; $b_{(HAPLO)}$, regression coefficient of the linear regression of EBV on GEBV using the haplotype-based approach; $b_{(SNP)}$, regression coefficient of the linear regression of EBV on GEBV using the SNP-based approach; SFA, saturated fatty acid; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid; n3, omega-3; n6, omega-6.

requires a progeny test and complexes laboratory analyses. Therefore, the use of genomic information to obtain direct genomic value predictions for each FA can be a viable alternative from a technical and economical point of view. In general, the models evaluated in our study to perform the genomic predictions displayed close results in terms of prediction ability for the beef FA profile. Studies with simulated and real data have shown that evaluations with haplotype blocks are more reliable than using individual markers, especially when the marker density is low (Calus, De Roos, & Veerkamp, 2008; De Roos, Hayes, Spelman, & Goddard, 2008), since the ancestral haplotype segments capture higher LD with QTL than does the individual markers. Probably, the marker density used in the present study was enough to capture sufficient LD between the SNP or haplotype markers with the QTL region. Moreover, genomic selection studies that reported higher prediction ability with the haplotype-based approach (Cuyabano et al., 2014; Hess et al., 2017) adjusted directly the SNPs or haplotype effects in the model. In this study, the SNPs or haplotype markers were used to infer the actual genomic relationships. With the ssGBLUP method, the relationships reflect the actual proportion of marker alleles shared by identity-by-state (IBS) as a deviation from the expected proportion of alleles shared in the population. Probably, differences in the approach used to model the marker effects, marker density, population background, heritability and trait architecture could explain the differences obtained among the different studies.

Considering the operational cost to infer the haplotypes and the similar results obtained between the haplotype and SNP-based approaches, the SNP-based ssGBLUP would be a viable tool in genomic evaluations for beef FA profile in indicine cattle. Although accuracies estimated have not been high, the use of the genomic information to predict the genomic values for the beef FA profile is a reasonable technical strategy contributing to the improvement of beef quality. Furthermore, the increase prediction ability of the genetic markers together with the reliability predictions' improvement is fundamental to increase the size of the reference population, with a higher number of phenotyped and genotyped animals. However, this will require greater investment in time, financial resources and logistics from breeding programmes and research groups.

Despite the low to moderate accuracies for the beef FA profile genomic values, it is possible to obtain genetic progress through selection using genomic information based either on SNPs or haplotype markers. None of the genetic marker approaches evaluated (SNPs or haplotypes) excelled in terms of accuracy; however, we identified that the SNP-based approach allows obtaining less biased genomic evaluations; thereby, this method is more feasible when taking into account the computational and operational cost to infer the haplotypes.

ACKNOWLEDGEMENTS

The first author thanks the Sao Paulo Research Foundation (FAPESP) for financial support and the University of Guelph by the reception and host during the exchange doctorate. The authors acknowledge DeltaGen, Nelore Qualitas, CRV Lagoa/PAINT, Cia. De Melhoramento and GenSys for providing the data used in this study.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest associated with this research.

DATA AVAILABILITY STATEMENT

The data sets generated and/or analysed during the current study are not publicly available due they are databases that belong to the private commercial farms, but are available from the corresponding author on reasonable request.

ORCID

Fabieli Loise Braga Feitosa https://orcid.org/0000-0003-4218-0348
Elisa Peripolli https://orcid.org/0000-0002-0962-6603
Luiz Fernando Brito https://orcid.org/0000-0002-5819-0922
Rafael Espigolan https://orcid.org/0000-0003-2586-1643

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*(2), 743–752. https://doi.org/10.3168/jds.2009-2730

Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics*, *21*(2), 263–265. https://doi.org/10.1093/bioinformatics/bth457

Berton, M. P., Fonseca, L. F. S., Gimenez, D. F. J., Utembergue, B. L., Cesar, A. S. M., Coutinho, L. L., … Baldi, F. (2016). Gene expression profile of intramuscular muscle in Nellore cattle with extreme values of fatty acid. *BMC Genomics*, *17*(1), 972. https://doi.org/10.1186/s12864-016-3232-y

Bickel, R. D., Kopp, A., & Nuzhdin, S. V. (2011). Composite effects of polymorphisms near multiple regulatory elements create a major-effect QTL. *PLoS Genetics*, *7*(1), e1001275. https://doi.org/10.1371/journal.pgen.1001275

Calus, M. P. L., Meuwissen, T. H. E., de Roos, A. P. W., & Veerkamp, R. F. (2008). Accuracy of genomic selection using different methods to define haplotypes. *Genetics*, *178*(1), 553–561. https://doi.org/10.1534/genetics.107.080838

Cesar, A. S. M., Regitano, L. C. A., Mourão, G. B., Tullio, R. R., Lanna, D. P. D., Nassu, R. T., … Coutinho, L. L. (2014). Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. *BMC Genetics*, *15*(1), 39. https://doi.org/10.1186/1471-2156-15-39

Chen, L., Ekine-Dzivenu, C., Vinsky, M., Basarab, J., Aalhus, J., Dugan, M. E. R., … Li, C. (2015). Genome-wide association and genomic prediction of breeding values for fatty acid composition in subcutaneous adipose and longissimus lumborum muscle of beef cattle. *BMC Genetics*, *16*(1), 135. https://doi.org/10.1186/s12863-015-0290-0

Chiaia, H. L. J., Peripoli, E., Silva, R. M. D. O., Aboujaoude, C., Feitosa, F. L. B., Lemos, M. V. A. D., … Baldi, F. (2017). Genomic prediction for beef fatty acid profile in Nellore cattle. *Meat Science*, *128*, 60–67. https://doi.org/10.1016/j.meatsci.2017.02.007

Christensen, O. F., & Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*, *42*(1), 2. https://doi.org/10.1186/1297-9686-42-2

Cuyabano, B. C., Su, G., & Lund, M. S. (2014). Genomic prediction of genetic merit using LD-based haplotypes in the Nordic Holstein population. *BMC Genomics*, *15*(1), 1171. https://doi.org/10.1186/1471-2164-15-1171

Cuyabano, B. C., Su, G., & Lund, M. S. (2015). Selection of haplotype variables from a high-density marker map for genomic prediction. *Genetics Selection Evolution*, *47*(1), 61. https://doi.org/10.1186/s12711-015-0143-3

De Roos, A. P. W., Hayes, B. J., Spelman, R. J., & Goddard, M. E. (2008). Linkage disequilibrium and persistence of phase in Holstein–Friesian, Jersey and Angus Cattle. *Genetics*, *179*(3), 1503–1512. https://doi.org/10.1534/genetics.107.084301

Enser, M., Hallett, K., Hewitt, B., Fursey, G. A., & Wood, J. D. (1996). Fatty acid content and composition of english beef, lamb and pork at retail. *Meat Science*, *42*(4), 443–456. https://doi.org/10.1016/0309-1740(95)00037-2

Fallin, D., & Schork, N. J. (2000). Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. *American Journal of Human Genetics*, *67*(4), 947–959. https://doi.org/10.1086/303069

Feitosa, F. L. B., Olivieri, B. F., Aboujaoude, C., Pereira, A. S. C., de Lemos, M. V. A., Chiaia, H. L. J., … Baldi, F. (2017). Genetic correlation estimates between beef fatty acid profile with meat and carcass traits in Nellore cattle finished in feedlot. *Journal of Applied Genetics*, *58*(1), 123–132. https://doi.org/10.1007/s13353-016-0360-7

Folch, J., Lees, M., & Stanley, G. S. (1957). A simple method for the isolation and purification of total lipides from animal tissues. *Journal of Biological Chemistry*, *226*(1), 497–509.

Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., … Altshuler, D. (2002). The structure of haplotype blocks in the human genome. *Science*, *296*(5576), 2225–2229. https://doi.org/10.1126/science.1069424

Hess, M., Druet, T., Hess, A., & Garrick, D. (2017). Fixed-length haplotypes can improve genomic prediction accuracy in an admixed dairy cattle population. *Genetics Selection Evolution*, *49*(1), 54. https://doi.org/10.1186/s12711-017-0329-y

Inoue, K., Kobayashi, M., Shoji, N., & Kato, K. (2011). Genetic parameters for fatty acid composition and feed efficiency traits in Japanese Black cattle. *Animal*, *5*(7), 987–994. https://doi.org/10.1017/S1751731111000012

Jakobsen, K. (1999). Dietary modifications of animal fats: Status and future perspectives. *European Journal of Lipid Science and Technology*, *101*(12), 475–483. https://doi.org/10.1002/(SICI)1521-4133(199912)101:12<475:AID-LIPI475>3.0.CO;2-H

Karimi, Z., Sargolzaei, M., Robinson, J. A. B., & Schenkel, F. S. (2018). Assessing haplotype-based models for genomic evaluation in Holstein cattle. *Canadian Journal of Animal Science*, *98*(4), 750–759. https://doi.org/10.1139/cjas-2018-0009

Kelly, M. J., Tume, R. K., Newman, S., & Thompson, J. M. (2013). Genetic variation in fatty acid composition of subcutaneous fat in cattle. *Animal Production Science*, *53*(2), 129. https://doi.org/10.1071/AN12154

Kramer, J. K., Fellner, V., Dugan, M. E., Sauer, F. D., Mossoba, M. M., & Yurawecz, M. P. (1997). Evaluating acid and base catalysts in the methylation of milk and rumen fatty acids with special emphasis on conjugated dienes and total trans fatty acids. *Lipids*, *32*(11), 1219–1228. https://doi.org/10.1007/s11745-997-0156-3

Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, *92*(9), 4656–4663. https://doi.org/10.3168/jds.2009-2061

Lemos, M. V. A., Chiaia, H. L. J., Berton, M. P., Feitosa, F. L. B., Aboujaoud, C., Camargo, G. M. F., … Baldi, F. (2016). Genome-wide association between single nucleotide polymorphisms with beef fatty acid profile in Nellore cattle using the single step procedure. *BMC Genomics*, *17*, 213. https://doi.org/10.1186/s12864-016-2511-y

Lund, M. S., Sahana, G., de Koning, D. J., Su, G., & Carlborg, Ö. (2009). Comparison of analyses of the QTLMAS XII common dataset. I: Genomic selection. *BioMed Central*, *3*(Suppl 1): S1. https://doi.org/10.1186/1753-6561-3-s1-s1

McNeill, S., & Van Elswyk, M. E. (2012). Red meat in global nutrition. *Meat Science*, *92*(3), 166–173. https://doi.org/10.1016/j.meatsci.2012.03.014

Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2015). Manual for BLUPF90 family of programs. Retrieved from http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf. Accessed May 29, 2019.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., & Lee, D. H. (2002). BLUPF90 and related programs (BGF90). In Proceedings of the 7th world congress on genetics applied to livestock production. Retrieved from http://www.wcgalp.org/system/files/proceedings/2002/blupf90-and-related-programs-bgf90.pdf. Accessed April 15, 2019.

Nogi, T., Honda, T., Mukai, F., Okagaki, T., & Oyama, K. (2011). Heritabilities and genetic correlations of fatty acid compositions in

longissimus muscle lipid with carcass traits in Japanese Black cattle. *Journal of Animal Science*, *89*(3), 615–621. https://doi.org/10.2527/jas.2009-2300

Onogi, A., Ogino, A., Komatsu, T., Shoji, N., Shimizu, K., Kurogi, K., … Iwata, H. (2015). Whole-genome prediction of fatty acid composition in meat of Japanese Black cattle. *Animal Genetics*, *46*(5), 557–559. https://doi.org/10.1111/age.12300

Pitchford, W. S., Deland, M. P. B., Siebert, B. D., Malau-Aduliand, A. E. O., & Bottema, C. D. K. (2002). Genetic variation in fatness and fatty acid composition of crossbred cattle. *Journal of Animal Science*, *80*(11), 2825–2832. https://doi.org/10.2527/2002.80112825x

Saatchi, M., Garrick, D. J., Tait, R. G., Mayes, M. S., Drewnoski, M., Schoonmaker, J., … Reecy, J. M. (2013). Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle. *BMC Genomics*, *14*, 730. https://doi.org/10.1186/1471-2164-14-730

Sargolzaei, M., Chesnais, J. P., & Schenkel, F. S. (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*, *15*(1), 478. https://doi.org/10.1186/1471-2164-15-478

Silva, R. M. O., Fragomeni, B. O., Lourenco, D. A. L., Magalhães, A. F. B., Irano, N., Carvalheiro, R., … Albuquerque, L. G. (2016). Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population1. *Journal of Animal Science*, *94*(9), 3613–3623. https://doi.org/10.2527/jas.2016-0401

Su, G., Brøndum, R. F., Ma, P., Guldbrandtsen, B., Aamand, G. P., & Lund, M. S. (2012). Comparison of genomic predictions using medium-density (~54,000) and high-density (~777,000) single nucleotide polymorphism marker panels in Nordic Holstein and Red Dairy Cattle populations. *Journal of Dairy Science*, *95*(8), 4657–4665. https://doi.org/10.3168/jds.2012-5379

Tait, R. G., Zhang, S., Knight, T., Bormann, J. M., Strohbehn, D. R., Tait, R. G., & Beitz, D. C. (2007). Heritability estimates for fatty acid concentration in angus beef. https://doi.org/10.31274/ans_air-180814-524

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Villumsen, T. M., Janss, L., & Lund, M. S. (2009). The importance of haplotype length and heritability using genomic selection in dairy cattle. *Journal of Animal Breeding and Genetics*, *126*(1), 3–13. https://doi.org/10.1111/j.1439-0388.2008.00747.x

Wiggans, G. R., Cooper, T. A., VanRaden, P. M., & Cole, J. B. (2011). Technical note: Adjustment of traditional cow evaluations to improve accuracy of genomic predictions. *Journal of Dairy Science*, *94*(12), 6188–6193. https://doi.org/10.3168/JDS.2011-4481

Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., …, Yang, J. (2018). Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nature Communications*, *9*(1), 918. https://doi.org/10.1038/s41467-018-03371-0

Zhu, B. O., Niu, H., Zhang, W., Wang, Z., Liang, Y., Guan, L., … Li, J. (2017). Genome wide association study and genomic prediction for fatty acid composition in Chinese Simmental beef cattle using high density SNP array. *BMC Genomics*, *18*(1), 464. https://doi.org/10.1186/s12864-017-3847-7

How to cite this article: Feitosa FLB, Pereira ASC, Amorim ST, et al. Comparison between haplotype-based and individual snp-based genomic predictions for beef fatty acid profile in Nelore cattle. *J Anim Breed Genet*. 2019;00:1–9. https://doi.org/10.1111/jbg.12463