

ANIMAL GENETICS AND GENOMICS

An assessment of genomic connectedness measures in Nellore cattle

Sabrina T. Amorim,[†] Haipeng Yu,[‡] Mehdi Momen,[‡]
Lúcia Galvão de Albuquerque,[†] Angélica S. Cravo Pereira,^{||} Fernando Baldi,[†]
and Gota Morota^{‡,1}

[†]Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, Departamento de Zootecnia, Via de acesso Prof. Paulo Donato Castellane, CEP 14884-900 Jaboticabal, SP, Brazil, [‡]Department of Animal and Poultry Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, ^{||}Universidade de São Paulo, Faculdade de Zootecnia e Engenharia de Alimentos, Núcleo de Apoio à Pesquisa em Melhoramento Animal, Biotecnologia e Transgenia, Rua Duque de Caxias Norte, 225, CEP 13635-900 Pirassununga, SP, Brazil

¹Corresponding author: morota@vt.edu

ORCID numbers: 0000-0003-4130-2040 (S. T. Amorim); 0000-0002-3567-6911 (G. Morota).

Abstract

An important criterion to consider in genetic evaluations is the extent of genetic connectedness across management units (MU), especially if they differ in their genetic mean. Reliable comparisons of genetic values across MU depend on the degree of connectedness: the higher the connectedness, the more reliable the comparison. Traditionally, genetic connectedness was calculated through pedigree-based methods; however, in the era of genomic selection, this can be better estimated utilizing new approaches based on genomics. Most procedures consider only additive genetic effects, which may not accurately reflect the underlying gene action of the evaluated trait, and little is known about the impact of non-additive gene action on connectedness measures. The objective of this study was to investigate the extent of genomic connectedness measures, for the first time, in Brazilian field data by applying additive and non-additive relationship matrices using a fatty acid profile data set from seven farms located in the three regions of Brazil, which are part of the three breeding programs. Myristic acid (C14:0) was used due to its importance for human health and reported presence of non-additive gene action. The pedigree included 427,740 animals and 925 of them were genotyped using the Bovine high-density genotyping chip. Six relationship matrices were constructed, parametrically and non-parametrically capturing additive and non-additive genetic effects from both pedigree and genomic data. We assessed genome-based connectedness across MU using the prediction error variance of difference (PEVD) and the coefficient of determination (CD). PEVD values ranged from 0.540 to 1.707, and CD from 0.146 to 0.456. Genomic information consistently enhanced the measures of connectedness compared to the numerator relationship matrix by at least 63%. Combining additive and non-additive genomic kernel relationship matrices or a non-parametric relationship matrix increased the capture of connectedness. Overall, the Gaussian kernel yielded the largest measure of connectedness. Our findings showed that connectedness metrics can be extended to incorporate genomic information and non-additive genetic variation using field data. We propose that different genomic relationship matrices can be designed to capture additive and non-additive genetic effects, increase the measures of connectedness, and to more accurately estimate the true state of connectedness in herds.

Key words: genomic connectedness, kernel matrices, Nellore cattle, non-additive gene action

Abbreviations

H^2	broad-sense heritability
h^2	narrow-sense heritability
σ_d^2	dominance genomic variance
σ_e^2	residual variance
σ_g^2	additive genomic variance
$\sigma_{g\#g}^2$	additive by additive epistasis genomic variance
σ_u^2	additive genetic variance
A	numerator relationship matrix
BLUP	best linear unbiased prediction
C14:0	myristic acid
CD	coefficient of determination
FA	fatty acid
G	genomic relationship matrix
G+D	additive and dominance relationship matrix
G+D+G#G	additive, dominance, and additive by additive relationship matrix
GK	Gaussian kernel relationship matrix
MME	mixed model equations
MU	management units
PEV	prediction error variance
PEVD	prediction error variance of difference
SNP	single nucleotide polymorphisms

Introduction

Genetic connectedness is a statistical measurement that allows reliable comparisons of the genetic values across management units (MU) by capturing the linkage among herds. The genetic values of animals from different MU (e.g., contemporary groups, farms, and herds) can be ranked using best linear unbiased prediction (BLUP). However, the accuracy of these comparisons depends on the degree of connectedness among MU: the higher the connectedness, the more reliable the comparison. Genetic connectedness has traditionally been calculated through pedigree-based methods (Lewis et al., 1999b; Kuehn et al., 2007a); however, these methods may underestimate connectedness in production systems such as the beef cattle industry, where commercial herds are poorly registered and multi-sire mating is practiced (Caires et al., 2012; Barbosa et al., 2013; Tonussi et al., 2017; Cavani et al., 2018). A lack of connectedness occurs: when the MU are genetically isolated (or semi-isolated) or there is limited sharing of genetic material; with the use of an incomplete numerator relationship matrix based on poor pedigree data (Carneiro et al., 2001); and with poor use of artificial insemination (AI). Genomic data are expected to more accurately estimate the relationship between individuals using information from genetic markers, such as single nucleotide polymorphisms (SNP), by measuring covariance among relatives and distant relatives previously ignored by a pedigree-based method (Habier et al., 2007).

Yu et al. (2017) evaluated the utility of genome-based connectedness in mice and cattle and noted that genomic relatedness could improve the extent of genetic connectedness measures compared with the pedigree when additive inheritance was assumed. The gain in connectedness measures was later shown to be associated with increased prediction accuracy based on cross-validation (Yu et al., 2018). Genetic connectedness studies were subsequently extended to account for non-additive genetic effects (Momen and Morota, 2018).

Those authors performed a computer simulation and found the increased measures of connectedness using additive and non-additive genomic relationship matrices under non-additive gene action. Collectively, those studies demonstrated that genomics can be used to enhance measures of connectedness. However, evaluations of genetic connectedness from field data remain limited.

In Brazil, cattle herds are often separated by large distances, and the rates of AI are low. A recent study by the Brazilian Association of Artificial Insemination (ASBIA, 2019) showed that only 16% of Brazilian dams are inseminated, with just a few farms available to measure expensive traits that require specific techniques or tests, such as postmortem beef quality traits.

The fatty acid (FA) profile of beef is a trait of interest due to its association with cardiovascular disease in humans (Mensink and Katan, 1992). According to Lawrie (2006), C14:0 is one of the most predominant saturated FA in cattle meat, which interferes with hepatic low-density lipoprotein receptors and consequently increases the amount of circulating low-density lipoprotein cholesterol (Grundy and Denke, 1990; Katan et al., 1994; Katan et al., 1995; Sacks and Katan, 2002). Considering the growing consumer demand for protein sources with a healthy lipid profile, several strategies have been applied to identify and manage the FA profile of beef (Faucitano et al., 2008; Liu et al., 2010; Aboujaoude et al., 2016; Berton et al., 2016; Chiaia et al., 2017).

Non-additive genetic variation for FA has been previously reported in cattle. For example, Malau-Aduli et al. (1998, 2000) reported significant dominance effects in Jersey, Limousin, and Jersey \times Limousin crossbred cattle. Li et al. (2012) detected significant additive and dominant effects for 19 individual FA in commercial beef steers. Kramer et al. (2016) identified epistatic interactions associated with FA concentrations in Angus cattle. Thus, the use of connectedness metrics including additive and non-additive gene effects may help to improve the quality of genetic value comparisons in breeding programs.

To date, few connectedness studies have been performed in Brazil (Carneiro et al., 2001; Pegolo et al., 2012), and the impact of genomic relatedness on connectedness measures in Nellore cattle has not been reported. Assessing connectedness statistics through genomic information may be useful for designing breeding programs and effectively linking units to improve the quality of across unit genetic evaluations, which in turn enhance the genetic improvement of Brazilian beef cattle. Therefore, the aim of this study was to investigate the extent of genomic connectedness measures in Nellore cattle by applying additive and non-additive relationship matrices, and to estimate variance components considering additive, dominance, and epistatic effects for myristic acid (C14:0).

Material and Methods

Data

Ethics statement

This study was approved by the ethics committee in the Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, São Paulo, Brazil.

Fatty acid profile data set

The data set included animals from seven farms located in the southeast, northeast, and mid-west of Brazil, which are part of three beef cattle breeding programs: DeltaGen (F1 [$n = 200$], F2 [$n = 22$], F3 [$n = 80$]); Paint (F4 [$n = 190$], F6 [$n = 292$], F7 [$n = 51$]); and Nellore Qualitas (F5 [$n = 90$]). These seven farms only collect

samples postmortem for analysis of beef quality traits due to the costs of collecting these phenotypes.

The GPS location of each farm was provided by the respective breeding program. The FA profile phenotypes were obtained for Nellore bulls with an average age of 24 mo. The methodology used to determinate FA profiles was consistent with that used in a previous study by Lemos et al. (2016). We analyzed myristic acid (C14:0) because of its importance to human health and high content in animals from feedlots (Zock et al., 1994). The pedigree included 427,740 animals born between 1977 and 2011. A total of $n = 925$ animals having C14:0 phenotype were genotyped using 777,962 SNP (Illumina, San Diego, CA). Following the removal of markers with a minor allele frequency less than 0.05, 505,367 SNP remained for further analysis. Missing genotypes were imputed using allele frequency estimates from a binomial distribution from the data. The seven farms were treated as MU.

Connectedness statistics

Genetic connectedness statistics are mostly defined as a function of the inverse of the coefficient matrix, which can be obtained from Henderson's mixed model equations (MME) (Henderson, 1984). In this study, we assessed genome-based connectedness across management units by applying the prediction error variance of difference (PEVD) (Kennedy and Trus, 1993) and coefficient of determination (CD) (Laloë, 1993).

Prediction error variance of difference

Prediction error variance (PEV) was obtained by fitting the following standard linear mixed model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon},$$

where \mathbf{y} is a vector of phenotypes, \mathbf{X} is an incidence matrix of systematic effects including management units, \mathbf{b} is a vector of systematic effects (contemporary group including animals born in the same farm, year, and from the same management group at yearling, and age at slaughter as linear covariate), \mathbf{Z} is an incidence matrix relating individuals to phenotypic records, \mathbf{u} is a vector of random additive genetic effects, and $\boldsymbol{\varepsilon}$ is a vector of residuals. The joint distribution of random effects for this model is

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim N \left[\begin{pmatrix} \mathbf{X}\mathbf{b} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{Z}\mathbf{K}\sigma_u^2\mathbf{Z}' + \mathbf{I}\sigma_\varepsilon^2 & \mathbf{Z}\mathbf{K}\sigma_u^2 & \mathbf{I}\sigma_\varepsilon^2 \\ \mathbf{K}\sigma_u^2\mathbf{Z}' & \mathbf{K}\sigma_u^2 & \mathbf{0} \\ \mathbf{I}\sigma_\varepsilon^2 & \mathbf{0} & \mathbf{I}\sigma_\varepsilon^2 \end{pmatrix} \right],$$

where σ_u^2 is the additive genetic variance, σ_ε^2 is the residual variance, and \mathbf{K} is one of the positive (semi) definite relationship matrices defined later.

The inverse of MME coefficient matrix of Henderson (1984) is represented as

$$\begin{aligned} \mathbf{C}^{-1} &= \begin{bmatrix} \mathbf{X}'\mathbf{X}\mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X}\mathbf{X}'\mathbf{Z} + \mathbf{K}^{-1}\lambda \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathbf{C}^{11}\mathbf{C}^{12} \\ \mathbf{C}^{21}\mathbf{C}^{22} \end{bmatrix}, \end{aligned}$$

where λ is the ratio of $\sigma_\varepsilon^2/\sigma_u^2$. The PEV for i th individual (\hat{u}_i) is written as

$$\begin{aligned} \text{PEV}_i &= \text{Var}(\hat{u}_i - u_i) \\ &= \text{Var}(u_i|\hat{u}_i) \\ &= \text{Var}(\hat{u}_i|u_i) \\ &= \mathbf{C}_{ii}^{22}\sigma_\varepsilon^2, \end{aligned}$$

where \mathbf{C}_{ii}^{22} represents the i th diagonal element of the \mathbf{C}^{22} coefficient matrix. Then, the PEVD of genetic values between individuals from different MU (Kennedy and Trus, 1993) is given by

$$\begin{aligned} \text{PEVD}(\hat{u}_i - \hat{u}_j) &= [\text{PEV}(\hat{u}_i) + \text{PEV}(\hat{u}_j) - 2\text{PEC}(\hat{u}_i, \hat{u}_j)] \\ &= (\mathbf{C}_{ii}^{22} - \mathbf{C}_{ij}^{22} - \mathbf{C}_{ji}^{22} + \mathbf{C}_{jj}^{22})\sigma_\varepsilon^2 \\ &= (\mathbf{C}_{ii}^{22} + \mathbf{C}_{jj}^{22} - 2\mathbf{C}_{ij}^{22})\sigma_\varepsilon^2, \end{aligned}$$

where ii and jj refer to the diagonal elements of \mathbf{C}^{22} matrix, corresponding to the i th and j th individuals, respectively, and the off-diagonal elements of \mathbf{C}^{22} are denoted by ij . PEC_{ij} is the prediction error covariance between the errors of genetic values, which is the off-diagonal element of the PEV matrix. Smaller PEVD indicates that the individuals are more connected.

The average PEVD between individuals across two MU was defined as follows:

$$\text{PEVD}_{i'j'} = \frac{1}{n_{i'} \cdot n_{j'}} \sum_{i \in i'} \sum_{j \in j'} \text{PEVD}_{ij},$$

where i is an animal in MU i' , j is an animal in MU j' , $n_{i'}$ and $n_{j'}$ represent the total number of records in i' and j' units, and the sum of all pairwise differences between two units is $\sum_{i \in i'} \sum_{j \in j'} \text{PEVD}_{ij}$.

Coefficient of determination

CD is defined by scaling the inverse of the coefficient matrix by corresponding coefficients from the relationship matrix, in other words, CD accounts for the reduction of connectedness due to relationship variability between individuals under comparison. The extent of CD ranges between 0 and 1, with larger values indicating increased connectedness.

A pairwise CD between individuals i and j is given by (Laloë et al., 1996)

$$\text{CD}_{ij} = 1 - \lambda \frac{\mathbf{C}_{ii}^{22} + \mathbf{C}_{jj}^{22} - 2\mathbf{C}_{ij}^{22}}{\mathbf{K}_{ii} + \mathbf{K}_{jj} - 2\mathbf{K}_{ij}},$$

where \mathbf{K}_{ii} and \mathbf{K}_{jj} are the i th and j th diagonal elements of \mathbf{K} .

The CD between two units can be scaled using the individual average PEVD with the average pairwise relationship differences across individuals to compute the individual average CD as described by Yu and Morota (2020)

$$\begin{aligned} \text{CD}_{i'j'} &= 1 - \lambda \frac{\frac{1}{n_{i'} \cdot n_{j'}} \sum_{i \in i'} \sum_{j \in j'} (\mathbf{C}_{ii}^{22} + \mathbf{C}_{jj}^{22} - 2\mathbf{C}_{ij}^{22})}{\frac{1}{n_{i'} \cdot n_{j'}} \sum_{i \in i'} \sum_{j \in j'} (\mathbf{K}_{ii} + \mathbf{K}_{jj} - 2\mathbf{K}_{ij})} \\ &= 1 - \frac{\frac{1}{n_{i'} \cdot n_{j'}} \sigma_\varepsilon^2 \sum_{i \in i'} \sum_{j \in j'} (\mathbf{C}_{ii}^{22} + \mathbf{C}_{jj}^{22} - 2\mathbf{C}_{ij}^{22})}{\frac{1}{n_{i'} \cdot n_{j'}} \sigma_u^2 \sum_{i \in i'} \sum_{j \in j'} (\mathbf{K}_{ii} + \mathbf{K}_{jj} - 2\mathbf{K}_{ij})} \\ &= 1 - \frac{\frac{1}{n_{i'} \cdot n_{j'}} \sum_{i \in i'} \sum_{j \in j'} \text{PEVD}_{ij}}{\frac{1}{n_{i'} \cdot n_{j'}} \sigma_u^2 \sum_{i \in i'} \sum_{j \in j'} (\mathbf{K}_{ii} + \mathbf{K}_{jj} - 2\mathbf{K}_{ij})} \\ &= 1 - \frac{\sum_{i \in i'} \sum_{j \in j'} \text{PEVD}_{ij}}{\sigma_u^2 \sum_{i \in i'} \sum_{j \in j'} (\mathbf{K}_{ii} + \mathbf{K}_{jj} - 2\mathbf{K}_{ij})}. \end{aligned}$$

Genomic kernel relationship matrices

Parametric relationship matrices

The extent of connectedness measures depends on the choice of relationship matrix K . In this study, we evaluated six K matrices. The pedigree-based relationship matrix ($K = A$) was calculated to obtain the additive numerator relationship, reflecting the probability that alleles are identical by descent inherited from a common ancestor (Wright, 1922). The diagonal element a_{ij} , which is the numerator relationship coefficient between two animals i and j for a population of n individuals is equal to $1 + F_j$, where F_j is the inbreeding coefficient of animal i . The off-diagonals of this matrix are twice the kinship coefficients and are equivalent to the numerators of Wright's correlation coefficients (Wright, 1922; Malécot, 1948). The A matrix among 925 animals was constructed using the pedigree records of 427,740 animals.

The genomic relationship matrix ($K = G$) was used to capture the genomic similarity among individuals, estimating the proportion of the genome between individuals that is identical by state. The G matrix was obtained as a function of the allele content including elements of 0, 1, and 2 representing the copies of reference alleles according to VanRaden (2008) as follows:

$$G = W_a W_a' / \left(2 \sum_{k=1}^m p_k (1 - p_k) \right),$$

where W_a is a centered incidence matrix taking values $0 - 2p_k$ for zero copies of the reference allele; $1 - 2p_k$ for one copy of the reference allele; and $2 - 2p_k$ for two copies of the reference allele. Here p_k is the allele frequency at SNP $k = 1, \dots, m$.

To capture dominance genetic effects, we constructed a dominance relationship matrix ($K = D$) according to Vitezica et al. (2013):

$$D = W_d W_d' / \sum_{k=1}^m (2p_k(1 - p_k))^2,$$

where W_d is the dominance marker incidence matrix, taking values of $-2p_k^2$ for zero copies of the reference allele; $2p_k(1 - p_k)^2$ for one copy of the reference allele; and $-2(1 - p_k)^2$ for two copies of the reference allele.

By combining the aforementioned G and D , we considered the following scenarios: pedigree (A), additive (G), additive and dominance ($G + D$), additive and additive by additive epistasis ($G + G \# G$), and additive, dominance, and additive by additive epistasis ($G + D + G \# G$), where $\#$ denotes the Hadamard product (Henderson, 1985). For a multi-kernel approach, a single kernel matrix was derived by weighting each of these kernels by its relative contribution to the total genetic variation according to Momen and Morota, (2018).

Gaussian kernel

In a Gaussian kernel ($K = GK$), the relationship between individuals is defined as distances on the Euclidean space, creating genetic relatedness in terms of spatial distance (de los Campos et al., 2010). The relationship between a pair of individuals i and j with their genotype vectors $w_i \in (0, 1, 2)$ and $w_j \in (0, 1, 2)$ is given by

$$GK(w_i, w_j) = \exp(-\theta d_{ij}^2)$$

$$= \prod_{k=1}^m \exp(-\theta (w_{ik} - w_{jk})^2),$$

where $d_{ij} = \sqrt{(w_{i1} - w_{j1})^2 + \dots + (w_{ik} - w_{jk})^2 + \dots + (w_{im} - w_{jm})^2}$ is the Euclidean distance between two individuals, and θ is a positive bandwidth parameter that controls the overall smoothness of the kernel function. A small Euclidean distance between two individuals means that their genotypes are similar in state, or in other words, that they have a strong relationship. The parameter θ is what controls the extent of genomic similarity between individuals. As θ increases, the kernel approaches zero (i.e., local kernel), and smaller θ produces entries closer to 1, in other words, two individuals match perfectly (i.e., global kernel) (Morota et al., 2013). We employed a kernel averaging approach (de los Campos et al., 2010) by integrating two kernel matrices ($GK1$ and $GK2$) using two extreme values of θ , so the mean of the off-diagonals elements in each kernel were 0.2 and 0.8, respectively. The averaged GK was obtained by

$$GK = \frac{\sigma_{GK1}^2}{\sigma_{GK1}^2 + \sigma_{GK2}^2} GK1 + \frac{\sigma_{GK2}^2}{\sigma_{GK2}^2 + \sigma_{GK1}^2} GK2,$$

where σ_{GK1}^2 and σ_{GK2}^2 are variance components attached to kernels $GK1$ and $GK2$, respectively. A BLUP model coupled with a Gaussian kernel matrix is known as reproducing Hilbert spaces regression (Morota and Gianola, 2014).

Throughout this study, we used R statistical computing environment (R Development Core Team, 2019) to estimate variance components and connectedness metrics with the following packages: the BGLR package (Pérez and de Los Campos, 2014) to estimate variance components and the GCA package (Yu and Morota, 2020) to compute connectedness measures.

Results

Heritability estimates

Descriptive statistics and genetic parameter estimates for C14:0 are presented in Table 1 for parametric kernel matrices. In this study, we estimated narrow-sense heritability (h^2) accounting for additive effects only ($K = A$ or G) and broad-sense heritability (H^2) accounting for additive, dominance, and epistasis ($K = G + D$, $G + G \# G$, or $G + D + G \# G$).

Narrow- and broad-sense heritability estimates for C14:0 ranged from 0.142 (± 0.095) to 0.462 (± 0.092). Heritability estimates were lower when using the pedigree relationship matrix compared with the other relationship matrices, likely because of incomplete pedigree data. The h^2 estimate from G was approximately two-fold higher than that from A . Moreover, the inclusion of dominance ($G + D$) and epistatic effects ($G + G \# G$ and $G + D + G \# G$) increased the heritability estimates. These findings suggest that C14:0 may be controlled by additive as well as non-additive gene action.

Prediction error variance of difference

Figure 1 shows the PEVD estimates across MU derived from A , G , $G + D$, $G + G \# G$, and $G + D + G \# G$ for C14:0. The smaller the PEVD, the higher the connectedness. The smallest PEVD connectedness measures were found in A , ranging from 1.654 to 1.707. When comparing A with G , enhanced connectedness measures were observed by reduced PEVD (1.387–1.422). The

Table 1. Descriptive statistics and heritability estimates for each gene action scenario

Trait	N	h ²			H ²		
		Mean±SD	A	G	G+D	G+G#G	G+D+G#G
C14:0	925	10.26±0.16	0.142 (0.095)	0.268 (0.081)	0.390 (0.089)	0.420 (0.076)	0.462 (0.092)
Genetic Variance			$\sigma_a^2 = 0.120$ (0.083)	$\sigma_g^2 = 0.225$ (0.073)	$\sigma_d^2 = 0.201$ (0.067) $\sigma_a^2 = 0.150$ (0.056)	$\sigma_g^2 = 0.183$ (0.064) $\sigma_{g\#g}^2 = 0.175$ (0.069)	$\sigma_a^2 = 0.155$ (0.059) $\sigma_d^2 = 0.123$ (0.049) $\sigma_{g\#g}^2 = 0.135$ (0.053)
Residual Variance			$\sigma_e^2 = 0.725$ (0.086)	$\sigma_e^2 = 0.615$ (0.071)	$\sigma_e^2 = 0.550$ (0.078)	$\sigma_e^2 = 0.494$ (0.085)	$\sigma_e^2 = 0.481$ (0.080)

¹The concentration of fatty acids is expressed as a percentage of total fatty acid methyl esters (FAME). A: pedigree. G: additive genomic kernel relationship matrix. G + D: additive × dominance genomic kernel relationship matrix. G + G#G: additive × epistasis genomic kernel relationship matrix. G + D + G#G: additive × dominance × epistasis genomic kernel relationship matrix. σ_a^2 : additive genetic variance; σ_d^2 : dominance genomic variance; $\sigma_{g\#g}^2$: additive by additive epistasis genomic variance; σ_e^2 : residual variance. Posterior standard errors are shown in the parentheses.

MU were more connected when genomic information was used. The inclusion of dominance (G + D) moderately enhanced the PEVD across MU, with estimates ranging from 1.288 to 1.315. Accounting for additive and additive epistasis (G + G#G) in the model marginally increased the measures of connectedness. Including additive, dominance, and additive by additive epistasis (G + D + G#G) resulted in the highest connectedness estimates among the parametric kernel relationship matrices. The use of GK significantly increased estimates of genetic connectedness across all MU. Overall, A yielded the least connected measures, while GK produced most connectedness estimates that were considered to be less connected.

Coefficient of determination

Individual average CD for each of the six scenarios is presented in Figure 2. The higher the CD, the higher the connectedness, and a similar pattern was found as reported for PEVD. The largest measured CD (0.456) was obtained with GK, and the smallest CD (0.146) was obtained with A. The MU presented low levels of connectedness in A, in which the highest estimate was 0.171. Compared to the pedigree-based method, genetic relatedness inferred from G increased the estimates of genetic connectedness across MU, and this trend was enhanced when dominance (G + D) and epistasis (G + G#G) were included. Combining additive, dominance, and epistasis in G + D + G#G resulted in greater measures of connectedness than any of the parametric relationship kernel matrices. The estimates of CD for G + D + G#G ranged from 0.427 to 0.445. GK presented the highest CD estimates, ranging from 0.442 to 0.456. These results demonstrated the importance of accounting for additive, dominance, and epistasis when the trait is also controlled by non-additive gene action.

Connectedness within and across breeding programs

Farms F1–F3 belonged to the DeltaGen breeding program. F1 and F3 farms were found to be the most connected across all scenarios. Farms F4, F6, and F7 belonged to the Paint breeding program. Although all MU presented similar connectedness values for CD, F4 and F6 were found to be the most connected, whereas F7 was well connected with the other two MU. Finally, we investigated the connectedness of the MU between the three different breeding programs and found that only one MU from Nelore Qualitas (F5) was adequately connected with the other MU. This pattern also appeared when we analyzed MU from DeltaGen and Paint, suggesting that these three breeding programs are connected, probably because of the use of AI in recent years.

Discussion

This study presents genome-based connectedness estimates in Nelore cattle using phenotype and genotype samples of commercial herds from the three breeding programs. The phenotype studied in this analysis is not routinely measured in breeding programs. Phenotypic information relating to the C14:0 trait was only available for a small number of contemporary groups in the MU, which could limit the data connection.

In Brazil, the beef cattle industry is concentrated in two main regions: the southeast and the mid-west. The recent expansion of agriculture has introduced animal husbandry to new regions. Figure 3 shows the distance (in kilometers) between the farms evaluated in our study. Farms F1–F3 are part of the DeltaGen

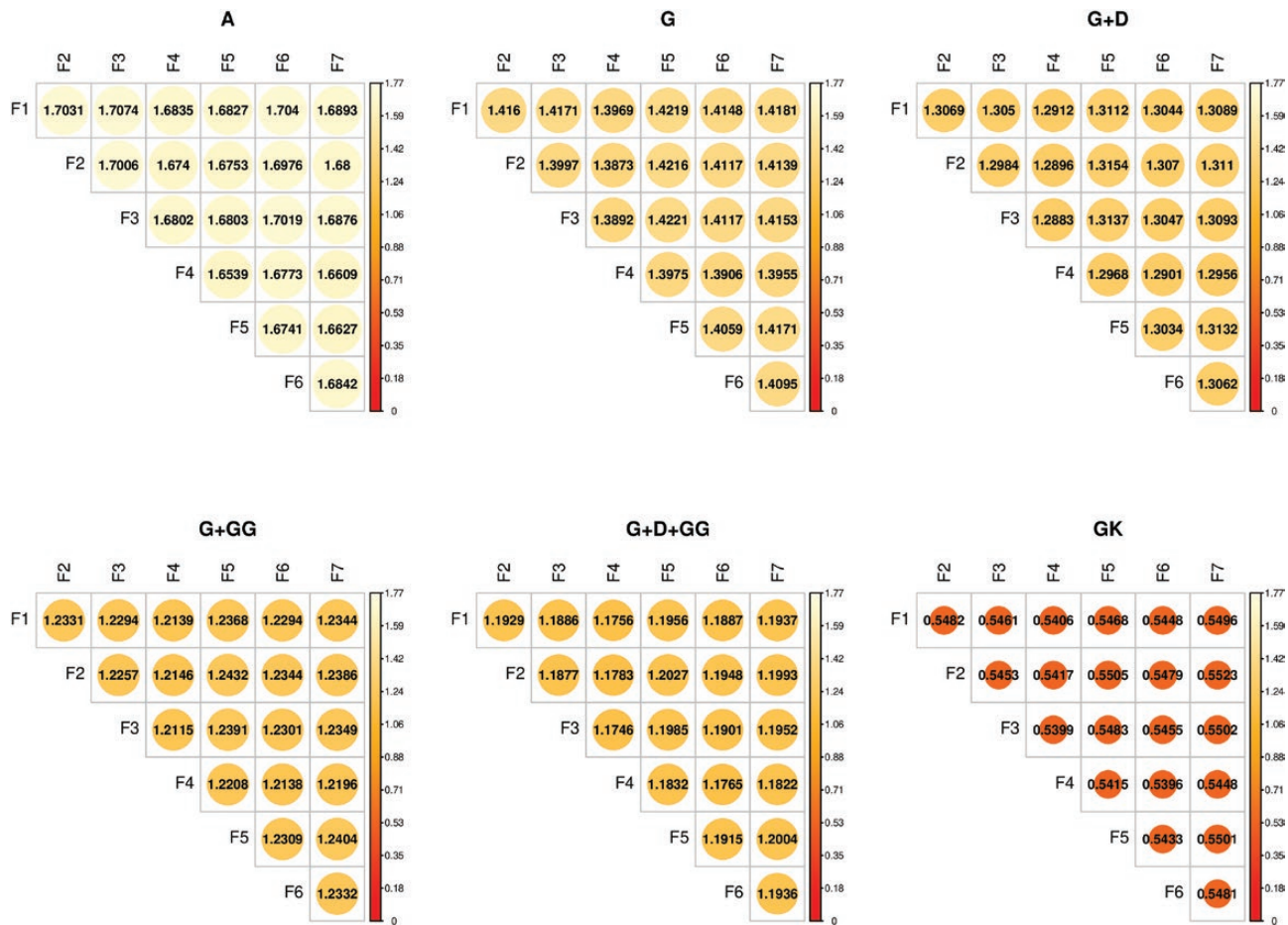


Figure 1. Individual average PEVD for C14:0. A: pedigree kernel relationship matrix. G: additive genomic kernel relationship matrix. G + D: additive and dominance genomic kernel relationship matrices. G + G#G: additive and epistasis genomic kernel relationship matrices. G + D + G#G: additive and dominance and epistasis genomic kernel relationship matrices. GK: Gaussian kernel relationship matrix.

Program and are spread across three regions (mid-west, southeast, and northeast). F1 and F3 are the most distant units in this program. Within the Nellore Paint program, farms F4, F6, and F7 are located in the mid-west and southeast regions in Brazil. The distance between F4 and F6 is the longest, despite being located in the same region.

Overall, considering each farm apart from its breeding program, F4 was the closest to any other farm, while F6 was the most distant. The average distance between two farms was 1,188 km. The extent of connectedness may be partially explained by the geographic position of each farm, since Brazilian herds consist of many subpopulations isolated by geographical distance, limiting the sharing of genetic material across MU. Overall, we did not find a clear relationship between connectedness and distance. Thus, geographical location alone does not explain the extent of connectedness.

The degree of connectedness reflects the reliability of comparisons between animals of different MU (Kennedy and Trus, 1993; Tosh and Wilton, 1994; Hanocq et al., 1999; Carneiro et al., 2001; Mathur et al., 2002; Roso et al., 2004). Low connectedness implies that the reliability of genetic value comparisons and animal rankings across MU are not sufficiently reliable (Lewis et al., 1999a; Lewis and Simm, 2000; Kuehn et al., 2008). Theoretically, the extent of connectedness may be of less concern (Fernando et al., 1983; Kennedy and Trus, 1993;

Fries and Roso, 1997). Kennedy and Trus (1993), and Tosh and Wilton (1994) reported that disconnected MU may not lead to biased predictions if the genetic values of the base animals are randomly and identically distributed throughout the population. However, this is less true for field populations such as beef cattle in Brazil, because of genetic selection, drift, limited use of AI, and low phenotyping rates in hard-to-measure traits, which have an impact on the genetic means and variance component estimates of the MU (Clément et al., 2001; Tosh and Wilton, 1994; Kuehn et al., 2007b; Tarrés et al., 2010).

Estimates of genetic parameters

Connectedness is often used to design or evaluate the effectiveness of breeding programs prior to phenotyping. However, phenotypic information enters the derivation of connectedness through heritability or the ratio of variance components (λ). In the present study, we evaluated the FA C14:0 (myristic acid). In general, the narrow-sense heritability estimates for individual FA profiles in the *Longissimus thoracis* muscle of beef cattle are low to moderate (Tait et al., 2007; Cesar et al., 2014; Lemos et al., 2016; Feitosa et al., 2017), ranging from 0.17 to 0.64. Our estimates for myristic acid using A (0.141) were lower than those reported by Tait et al. (2007), of 0.23, in Angus cattle, but were closer when G (0.274) was used in our study. Using a population of Nellore Cattle, Aboujaoude et al. (2016)

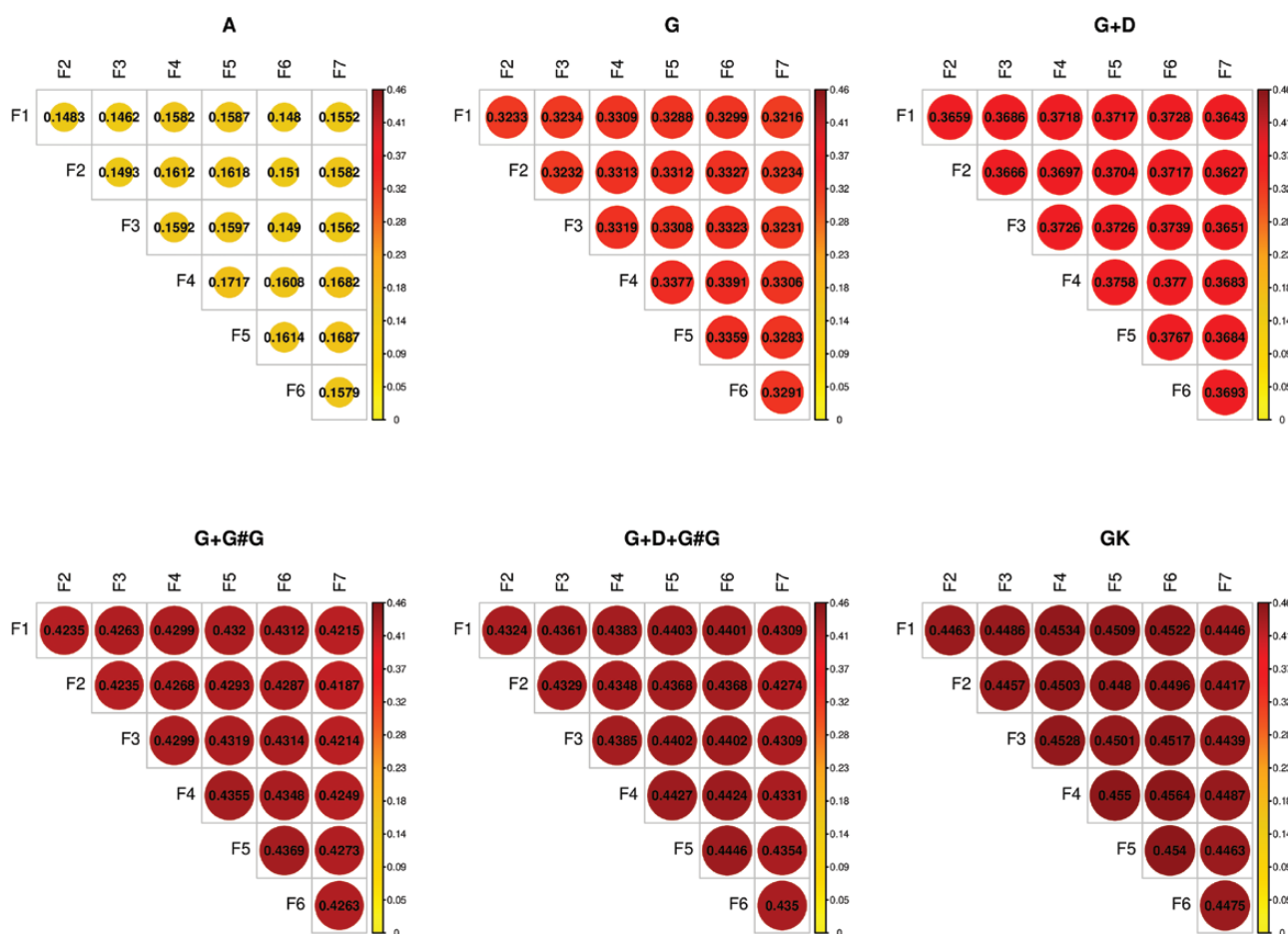


Figure 2. Individual average CD for C14:0. A: pedigree kernel relationship matrix. G: additive genomic kernel relationship matrix. G + D: additive and dominance genomic kernel relationship matrices. G + G#G: additive and epistasis genomic kernel relationship matrices. G + D + G#G: additive and dominance and epistasis genomic kernel relationship matrices. GK: Gaussian kernel relationship matrix.

reported a genomic h^2 estimate of 0.25. Cesar et al. (2014) and Feitosa et al. (2017) reported genomic h^2 estimates of C14:0 in Nellore Cattle of 0.17 and 0.25, respectively.

In the current study, we evaluated six relationship matrices using pedigree and genomic data accounting for additive and non-additive genetic variation into the kernel relationship matrices. We found that the heritability estimates from A were the lowest. In the beef cattle industry, the poor records for herds are problematic, especially for beef quality traits evaluated after slaughter. This affects both the measurement relationships and the genetic parameter estimates (Fouilloux et al., 2008a). Here, approximately 55% of records contained no sire information. The percentage of animals with known sires in each farm was F1 = 13%, F2 = 72%, F3 = 81.25%, F4 = 57.36%, F5 = 63.15%, F6 = 27.39%, F7 = 100%, and for a total 413/925 animals in the data set. This data set is composed by commercial herds, and in Brazil, most farms are under a multiple service sires (MS) mating system, which consists of groups of cows clustered with several bulls in the same paddock during a riding season (Cardoso and Tempelman, 2003). MS improves the conception rate (Lunstra and Laster, 1982), but results in an uncertain paternity scenario. Therefore, incomplete pedigree data due to MS negatively affect the accuracy of genetic evaluations (Cardoso and Tempelman, 2003).

In the breeding programs in our study, the mating season begins with AI in all dams. However, if the dam is not pregnant, the MS mating system is applied. The bulls used in this system are usually from the same farm, and they only produce one generation of progeny. The calf born from the MS mating system is slaughtered because of uncertain paternity. These calves are considered to have unreliable genetic merit, which means that they are not candidates for breeding, while allowing phenotypes to be obtained from the slaughtered animals. The use of genomic information can enhance the feasibility of genetic evaluation and increase the prediction accuracy of these novel traits in the beef cattle industry.

The use of SNP, which capture molecular similarity and Mendelian sampling, determines relationships between individuals at the genomic level, recovering information missing in the pedigree. Cesar et al. (2014) first reported genomic heritability estimates for FA in Nellore cattle and showed that SNP panels are a promising tool for the genetic improvement of Nellore cattle in Brazil, mainly because of the effective cost strategy for their application in breeding programs. We found that the heritability estimates from G were larger compared to those from A, recovering a greater proportion of additive genetic variance. Our observations are consistent with those of Ishii et al. (2013) in Japanese Black cattle and Saatchi et al. (2013) in US Angus cattle.

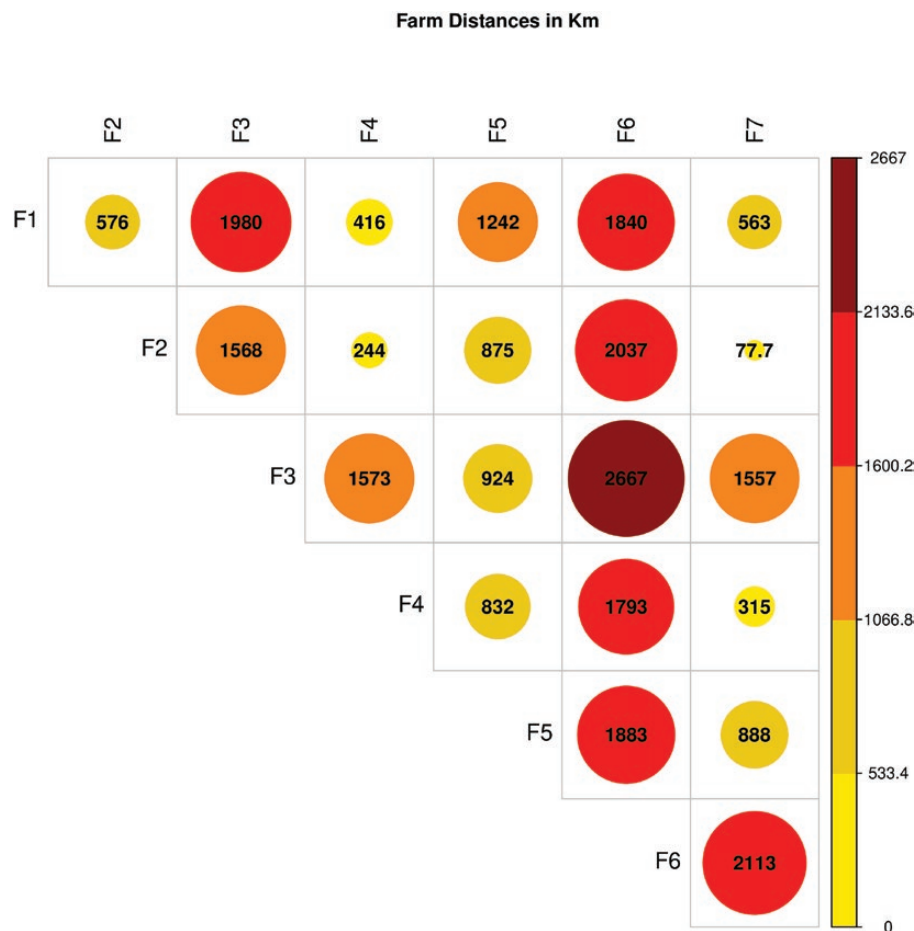


Figure 3. Farm distances in Km. F1—Dourados (MS); F2—Valparaíso (SP); F3—Cotegipe (BA); F4—Água Clara (MS); F5—Goianésia (GO); F6—Juruena (MT); F7—Piacatu (SP).

It is also critical to account for dominance and/or epistatic variation to optimize breeding designs, such as mate allocation. Several animal studies (Serenius et al., 2006; Sun et al., 2014; Moghaddar and van der Werf, 2017; Joshi et al., 2018; Ebrahimi et al., 2019) have reported that dominance heritability is often significantly greater than zero, and its inclusion in prediction models could improve performance in those studies. We observed dominance and epistatic variance for C14:0, suggesting that C14:0 may be controlled by both additive and non-additive genetic variation.

Connectedness statistics

Because of missing pedigree data in the present study, connectedness statistics using the numerator relationship matrix resulted in the lowest connectedness estimates, indicating that A may provide an incorrect picture of connectedness for C14:0 in the population studied. In contrast, G consistently enhanced the measures of connectedness. The results from G highlight how genomic data can help us to better observe the true state of connectedness, particularly when pedigree data are less reliable. This supports the findings of Yu et al. (2017), who reported that genomic relatedness inferred from SNP increases the estimates of genetic connectedness across MU compared to estimates of pedigree information. As noted by Yu et al. (2017, 2018), the availability of genomic information provides an opportunity to improve the quality of genetic value comparisons and revisit a number of critical questions related to connectedness.

The increased interest in non-additive variation (Wolak and Keller, 2014; Varona et al., 2018) suggests that it may be possible to account for such variation in connectedness studies. Momen and Morota (2018) demonstrated that connectedness metrics can be extended to incorporate non-additive genetic variation of complex traits. They showed an increase up to 25% in the capture of connectedness using additive and non-additive genomic kernel relationship matrices when the trait of interest is controlled by non-additive gene action. Our study investigated how the inclusion of such variation could impact connectedness metrics. Increased estimates of connectedness were observed when dominance and epistatic parametric kernels were included. The non-parametric relationship matrix (GK) models higher-order epistatic gene action by taking the Hadamard product between the G matrices when SNP were coded in an additive manner (Jiang and Reif, 2015). In our study, GK was better than all parametric approaches, highlighting the usefulness of GK for incorporating non-additive gene action. Nii et al. (2006) first reported the presence of epistatic quantitative trait loci for perirenal C14:0 in wild boars, and was later reported by Uemoto et al. (2009) on chromosome 16 in swine.

In livestock, the inclusion of dominance effects can be justified by the use of semen from a few genetically superior bulls and reproductive biotechnologies (such as multiple ovulation, embryo transfer, and in vitro fertilization). Consequently, the number of full sibling progenies increases, which increases the relationships within and between generations, as well as

dominance genetic relationships (VanRaden, 1992). The inclusion of such effects was reported by Varona and Misztal (1999), who noted that the inclusion of dominance effects into genetic evaluations enables the determination of specific combinations for mating schemes, and the separation of additive variation from the rest, especially in populations containing many full-sibs in their pedigree. Hayes and Miller (2000) and Ishida and Mukai (2004) stated that ignoring non-additive genetic variance in breeds could result in biased predictions of genetic values, which would affect the animal's classification as well as national and international comparisons.

Fouilloux and Laloë (2001) developed the criterion of admission to the group of connected herds (CACO) method to compare the average CD values of all herds and to cluster them. This method was applied by Pegolo et al. (2012) in Brazilian Nellore cattle using registered animal data from elite herds (National Association of Breeders and Researchers, ANCP, Ribeirão Preto, Brazil). They recorded weight after 210 d ($h^2 = 0.25$) to investigate the trajectory of connectedness from 1999 to 2003 and from 2004 to 2008. They found moderate estimates of pedigree-based connectedness and attributed the increase in connectedness to the use of AI, which increased 47% with the semen sales in that period. According to ASBIA (2011), only 10% of dams were inseminated at reproductive age. AI affects the measurement of genetic connectedness among herds (Fouilloux et al., 2008) and strongly influences the quality of pedigree information.

AI can accelerate genetic improvement in a population as it allows semen from animals with a higher genetic value to be utilized, which are not normally available for use in natural matings. It also increases the number of offspring per sire. Despite the growing use of AI, most commercial beef cattle programs in Brazil still use unproven sires for natural mating. In addition, programs that use AI still need a bull to mate with their dams in the case of AI failure, which is common on commercial farms where animals are destined for slaughter. As the MU evaluated in this study were not part of consolidated breeding programs, the use of AI was more than 50% of that in all farms, because of known paternity scenarios needed for the genetic evaluations. Thus, the use of the MS mating system in these programs incurs the costs of maintaining a non-pregnant cow in the herd.

Changes in connectedness levels reported by Pegolo et al. (2012) showed that herd descriptors, such as the number of animals in the herd, the number of sires used in the herd, the percentage of connecting sires, the percentage of progeny from connecting sires, and the percentage of calves with unknown sires, cannot fully explain how the herds are connected. The use of different types or combinations of relationship matrices to those used in our study may be a viable approach to understand the complexity of genetic connectedness in livestock species.

Here, we describe the first application of genomic connectedness in Brazilian Nellore cattle. This study shows how genomic information can increase connectedness measures when pedigree information is not complete due to multiple sire systems. Collectively, through the use of genomics and by accounting for non-additive gene action, we can better reflect signals of connectedness not captured by pedigree-based counterparts.

Conclusions

Genetic connectedness plays a key role in the quality of genetic value comparisons across MU. We used PEVD and CD to assess genomic connectedness measures in Nellore cattle field data,

accounting for the presence of non-additive gene action. Our findings show that genomic information can capture connectedness signals that may be missed from the pedigree, providing a more precise picture of connectedness. Working with novel traits can be challenging when they are hard to measure, because only a few breeding programs possess the infrastructure and logistics to collect phenotypes such as the FA profile of meat. We show that it is possible to capture connectedness signals from samples of different farms within and between different breeding programs using a specific phenotype (C14:0) with low heritability. We also confirm that the use of AI, even though still used at low levels in the country (16%), has an important role in connecting herds. Furthermore, we observed that the use of additive and non-additive genomic kernel relationship matrices can enhance the capture of connectedness measures compared to purely additive counterparts.

Acknowledgments

We acknowledge FAPESP (São Paulo Research Foundation, grants 2009/16118-5, 2011/21241-0, 2018/19463-4, and 2019/04929-0) for financial support and Nelore Qualitas, Paint, and DeltaGen for providing the data. S.T.A. received a scholarship from the São Paulo Research Foundation in conjunction with the Postgraduate Program on Genetics and Animal Breeding, Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias (FCAV, UNESP). F.B. and L.G.A. held productivity research fellowships from The Brazilian National Council for Scientific and Technological Development (CNPQ).

Conflict of interest statement

The authors declare no real or perceived conflicts of interest.

Literature Cited

- Aboujaoude, C., A. S. C. Pereira, F. L. B. Feitosa, M. V. Antunes de Lemos, H. L. J. Chiaia, M. Piatto Berton, E. Peripolli, R. M. de O. Silva, A. M. Ferrinho, L. F. Mueller, et al. 2016. Genetic parameters for fatty acids in intramuscular fat from feedlot-finished Nelore carcasses. *Anim. Prod. Sci.* 58:234. doi:10.1071/AN16107
- Associação Brasileira De Inseminação Artificial—ASBIA. Relatório estatístico de produção, importação e comercialização de sêmen., 2011a. [accessed on March 13, 2020]. <https://www.slideshare.net/BeefPoint/asbia-ndice-2011>.
- Associação Brasileira De Inseminação Artificial—ASBIA. Relatório estatístico de produção, importação e comercialização de sêmen., 2011b. [accessed on March 13, 2020]. asbia.org.br/wp-content/uploads/2020/02/Index-asbia-1.pdf.
- Barbosa, A. C. B., C. H. M. Malhado, P. L. S. Carneiro, L. M. S. Muniz, D. P. Ambrosini, J. A. Carrillo, and R. Martins-Filho. 2013. Population structure of Nelore cattle in northeastern Brazil. *Rev. Bras. Zootec.* 42:639–644. doi:10.1590/S1516-35982013000900005
- Berton, M. P., L. F. Fonseca, D. F. Gimenez, B. L. Utembergue, A. S. Cesar, L. L. Coutinho, M. V. de Lemos, C. Aboujaoude, A. S. Pereira, R. M. Silva, et al. 2016. Gene expression profile of intramuscular muscle in Nelore cattle with extreme values of fatty acid. *BMC Genomics.* 17:972. doi:10.1186/s12864-016-3232-y
- Caires, D. N., C. H. M. Malhado, L. de A. Souza, M. R. Teixeira Neto, P. L. S. Carneiro, and R. Martins Filho. 2012. Tabapuã breed in Northeastern Brazil: genetic progress and population structure. *Rev. Bras. Zootec.* 41:1858–1865. doi:10.1590/S1516-35982012000800008

- Cardoso, F. F., and R. J. Tempelman. 2003. Bayesian inference on genetic merit under uncertain paternity. *Genet. Sel. Evol.* 35:469–487. doi:[10.1186/1297-9686-35-6-469](https://doi.org/10.1186/1297-9686-35-6-469)
- Carneiro, A. P. S., R. de A. Torres, R. F. Euclides, M. de A. e Silva, P. S. Lopes, P. L. S. Carneiro, and R. de A. Torres Filho. 2001. Efeito da conexão de dados sobre o valor fenotípico médio e a variância genética aditiva. *Rev. Bras. Zootec.* 30:336–341. doi:[10.1590/s1516-35982001000200006](https://doi.org/10.1590/s1516-35982001000200006)
- Cavani, L., R. M. de O. Silva, L. O. D. Carreño, R. K. Ono, T. S. Bertipaglia, M. M. Farah, D. D. Millen, R. da Fonseca, L. Cavani, R. M. de O. Silva, et al. 2018. Genetic diversity of Brazilian Brahman cattle by pedigree analysis. *Pesqui. Agropecuária Bras.* 53:74–79. doi:[10.1590/s0100-204x2018000100008](https://doi.org/10.1590/s0100-204x2018000100008)
- Cesar, A. S., L. C. Regitano, G. B. Mourão, R. R. Tullio, D. P. Lanna, R. T. Nassu, M. A. Mudado, P. S. Oliveira, M. L. do Nascimento, A. S. Chaves, et al. 2014. Genome-wide association study for intramuscular fat deposition and composition in Nellore cattle. *BMC Genet.* 15:39. doi:[10.1186/1471-2156-15-39](https://doi.org/10.1186/1471-2156-15-39)
- Chiaia, H. L. J., E. Peripoli, R. M. O. Silva, C. Aboujaoude, F. L. B. Feitosa, M. V. A. Lemos, M. P. Berton, B. F. Olivieri, R. Espigolan, R. L. Tonussi, et al. 2017. Genomic prediction for beef fatty acid profile in Nellore cattle. *Meat Sci.* 128:60–67. doi:[10.1016/j.meatsci.2017.02.007](https://doi.org/10.1016/j.meatsci.2017.02.007)
- Clément, V., B. Bibé, E. Verrier, J. M. Elsen, E. Manfredi, J. Bouix, and E. Hanocq. 2001. Simulation analysis to test the influence of model adequacy and data structure on the estimation of genetic parameters for traits with direct and maternal effects. *Genet. Sel. Evol.* 33:369–95. doi:[10.1051/gse:2001123](https://doi.org/10.1051/gse:2001123)
- De los Campos, G., D. Gianola, G. J. Rosa, K. A. Weigel, and J. Crossa. 2010. Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. *Genet. Res. (Camb.)* 92:295–308. doi:[10.1017/S0016672310000285](https://doi.org/10.1017/S0016672310000285)
- Ebrahimi, K., G. R. Dashab, H. Faraji-Arough, and M. Rokouei. 2019. Estimation of additive and non-additive genetic variances of body weight in crossbred populations of the Japanese quail. *Poult. Sci.* 98:46–55. doi:[10.3382/ps/pey357](https://doi.org/10.3382/ps/pey357)
- Faucitano, L., P. Y. Chouinard, J. Fortin, I. B. Mandell, C. Lafrenière, C. L. Girard, and R. Berthiaume. 2008. Comparison of alternative beef production systems based on forage finishing or grain-forage diets with or without growth promotants: 2. Meat quality, fatty acid composition, and overall palatability. *J. Anim. Sci.* 86:1678–1689. doi:[10.2527/jas.2007-0756](https://doi.org/10.2527/jas.2007-0756)
- Feitosa, F. L., B. F. Olivieri, C. Aboujaoude, A. S. Pereira, M. V. de Lemos, H. L. Chiaia, M. P. Berton, E. Peripoli, A. M. Ferrinho, L. F. Mueller, et al. 2017. Genetic correlation estimates between beef fatty acid profile with meat and carcass traits in Nellore cattle finished in feedlot. *J. Appl. Genet.* 58:123–132. doi:[10.1007/s13353-016-0360-7](https://doi.org/10.1007/s13353-016-0360-7)
- Fernando, R. L., D. Gianola, and M. Grossman. 1983. Identifying all connected subsets in a two-way classification without interaction. *J. Dairy Sci.* 66:1399–1402. doi:[10.3168/jds.S0022-0302\(83\)81951-1](https://doi.org/10.3168/jds.S0022-0302(83)81951-1)
- Fouilloux, M. N., V. Clément, and D. Laloë. 2008. Measuring connectedness among herds in mixed linear models: from theory to practice in large-sized genetic evaluations. *Genet. Sel. Evol.* 40:145–159. doi:[10.1186/1297-9686-40-2-145](https://doi.org/10.1186/1297-9686-40-2-145)
- Fouilloux, M. N., and D. Laloë. 2001. A sampling method for estimating the accuracy of predicted breeding values in genetic evaluation. *Genet. Sel. Evol.* 33:473–486. doi:[10.1186/1297-9686-33-5-473](https://doi.org/10.1186/1297-9686-33-5-473)
- Fries, L. A., and V. M. Roso. 1997. Conectabilidade em avaliações genéticas de gado de corte: uma proposta heurística. In: *Proceedings of the Reunião Anual Da Sociedade Brasileira De Zootecnia*, constitute summaries of papers presented at the Reunião Anual Da Sociedade Brasileira De Zootecnia in Juiz de Fora in May 1997. SBZ. p. 159–161.
- Grundy, S. M., and M. A. Denke. 1990. Dietary influences on serum lipids and lipoproteins. *J. Lipid Res.* 31:1149–1172.
- Habier, D., R. L. Fernando, and J. C. M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177:2389–2397. doi:[10.1534/genetics.107.081190](https://doi.org/10.1534/genetics.107.081190)
- Hanocq, E. D., L. Tiphine, and B. Bibé. 1999. Le point sur la notion de connexion en génétique animale. *INRA Prod. Anim.* 12:101–111.
- Hayes, B. J., and S. P. Miller. 2000. Mate selection strategies to exploit across- and within-breed dominance variation. *J. Anim. Breed. Genet.* 117:347–359. doi:[10.1046/j.1439-0388.2000.00252.x](https://doi.org/10.1046/j.1439-0388.2000.00252.x)
- Henderson, C. R. 1984. *Applications of linear models in animal breeding*. Guelph, Canada: University of Guelph.
- Henderson, C. R. 1985. Best linear unbiased prediction of nonadditive genetic merits in noninbred populations. *J. Anim. Sci.* 60:111–117. doi:[10.2527/jas1985.601111x](https://doi.org/10.2527/jas1985.601111x)
- Ishida, T., and F. Mukai. 2004. Estimation of dominance genetic variances for reproductive traits and growth traits of calves in Japanese Black cattle. *Anim. Sci. J.* 75:285–294. doi:[10.1111/j.1740-0929.2004.00188.x](https://doi.org/10.1111/j.1740-0929.2004.00188.x)
- Ishii, A., K. Yamaji, Y. Uemoto, N. Sasago, E. Kobayashi, N. Kobayashi, T. Matsuhashi, S. Maruyama, H. Matsumoto, S. Sasazaki, et al. 2013. Genome-wide association study for fatty acid composition in Japanese Black cattle. *Anim. Sci. J.* 84:675–682. doi:[10.1111/asj.12063](https://doi.org/10.1111/asj.12063)
- Jiang, Y., and J. C. Reif. 2015. Modeling epistasis in genomic selection. *Genetics* 201:759–768. doi:[10.1534/genetics.115.177907](https://doi.org/10.1534/genetics.115.177907)
- Joshi, R., J. A. Woolliams, T. Meuwissen, and H. M. Gjølén. 2018. Maternal, dominance and additive genetic effects in Nile tilapia; influence on growth, fillet yield and body size traits. *Heredity (Edinb.)* 120:452–462. doi:[10.1038/s41437-017-0046-x](https://doi.org/10.1038/s41437-017-0046-x)
- Katan, M. B., P. L. Zock, and R. P. Mensink. 1994. Effects of fats and fatty acids on blood lipids in humans: an overview. *Am. J. Clin. Nutr.* 60(6 Suppl):1017S–1022S. doi:[10.1093/ajcn/60.6.1017S](https://doi.org/10.1093/ajcn/60.6.1017S)
- Katan, M. B., P. L. Zock, and R. P. Mensink. 1995. Dietary oils, serum lipoproteins, and coronary heart disease. *Am. J. Clin. Nutr.* 61(6 Suppl):1368S–1373S. doi:[10.1093/ajcn/61.6.1368S](https://doi.org/10.1093/ajcn/61.6.1368S)
- Kennedy, B. W., and D. Trus. 1993. Considerations on genetic connectedness between management units under an animal model. *J. Anim. Sci.* 71:2341–2352. doi:[10.2527/1993.7192341x](https://doi.org/10.2527/1993.7192341x)
- Kramer, L. M., M. A. Gaffar, J. E. Koltes, E. R. Fritz-Waters, M. S. Mayes, A. D. Sewell, N. T. Weeks, D. J. Garrick, R. L. Fernando, L. Ma, et al. 2016. Epistatic interactions associated with fatty acid concentrations of beef from angus sired beef cattle. *BMC Genomics* 17:891. doi:[10.1186/s12864-016-3235-8](https://doi.org/10.1186/s12864-016-3235-8)
- Kuehn, L. A., R. M. Lewis, and D. R. Notter. 2007a. Managing the risk of comparing estimated breeding values across flocks or herds through connectedness: a review and application. *Genet. Sel. Evol.* 39:225–247. doi:[10.1186/1297-9686-39-3-225](https://doi.org/10.1186/1297-9686-39-3-225)
- Kuehn, L. A., R. M. Lewis, and D. R. Notter. 2007b. Managing the risk of comparing estimated breeding values across flocks or herds through connectedness: a review and application. *Genet. Sel. Evol.* 39. doi:[10.1186/1297-9686-39-3-225](https://doi.org/10.1186/1297-9686-39-3-225)
- Kuehn, L. A., D. R. Notter, and R. M. Lewis. 2008. Assessing genetic gain, inbreeding, and bias attributable to different flock genetic means in alternative sheep sire referencing schemes. *J. Anim. Sci.* 86:526–535. doi:[10.2527/jas.2007-0255](https://doi.org/10.2527/jas.2007-0255)
- Laloë, D. 1993. Precision and information in linear models of genetic evaluation. *Genet. Sel. Evol.* 25:557. doi:[10.1186/1297-9686-25-6-557](https://doi.org/10.1186/1297-9686-25-6-557)
- Laloë, D., F. Phocas, and F. Ménissier. 1996. Considerations on measures of precision and connectedness in mixed linear models of genetic evaluation. *Genet. Sel. Evol.* 28:359. doi:[10.1186/1297-9686-28-4-359](https://doi.org/10.1186/1297-9686-28-4-359)

- Lawrie, R. A. 2006. Meat and human nutrition. In: R. A. B. T.-L. M. S. (Seventh E. Lawrie, editor. *Lawrie's Meat Science*. UK: Woodhead Publishing. p. 342–357.
- Lemos, Marcos V. A., H. L. J. Chiaia, M. P. Berton, F. L. B. Feitosa, C. Aboujaoud, G. M. F. Camargo, A. S. C. Pereira, L. G. Albuquerque, A. M. Ferrinho, L. F. Mueller, et al. 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. doi:10.1186/s12864-016-2511-y
- Lewis, R. M., R. E. Crump, G. Simm, and R. Thompson. 1999a. Assessing connectedness in across-flock genetic evaluations. *Proc. Br. Soc. Anim. Sci.* 1999:121–121. doi:10.1017/s1752756200002763.
- Lewis, R. M., R. E. Crump, G. Simm, R. Thompson, and B. S. of A. Science. 1999b. Proceedings of the British Society of Animal Science: constitute summaries of papers presented at the Society's Annual Meeting in Scarborough in March 1999. BSAS.
- Lewis, R., and G. Simm. 2000. Selection strategies in sire referencing schemes in sheep. *Livest. Prod. Sci.* 67:129–141. doi:10.1016/S0301-6226(00)00182-2
- Li, C., N. Aldai, M. Vinsky, M. E. Dugan, and T. A. McAllister. 2012. Association analyses of single nucleotide polymorphisms in bovine stearoyl-CoA desaturase and fatty acid synthase genes with fatty acid composition in commercial cross-bred beef steers. *Anim. Genet.* 43:93–97. doi:10.1111/j.1365-2052.2011.02217.x
- Liu, G. E., Y. Hou, B. Zhu, M. F. Cardone, L. Jiang, A. Cellamare, A. Mitra, L. J. Alexander, L. L. Coutinho, M. E. Dell'Aquila, et al. 2010. Analysis of copy number variations among diverse cattle breeds. *Genome Res.* 20:693–703. doi:10.1101/gr.105403.110
- Lunstra, D. D., and D. B. Laster. 1982. Influence of single-sire and multiple-sire natural mating on pregnancy rate of beef heifers. *Theriogenology* 18:373–382. doi:10.1016/0093-691x(82)90159-5
- Malau-Aduli, A. E. O., M. A. Edriss, B. D. Siebert, C. D. K. Bottema, and W. S. Pitchford. 2000. Breed differences and heterosis in triacylglycerol fatty acid composition of bovine adipose tissue. *J. Anim. Physiol. Anim. Nutr. (Berl.)* 83:106–112. doi:10.1046/j.1439-0396.2000.00257.x
- Malau-Aduli, A. E. O., B. D. Siebert, C. D. K. Bottema, and W. S. Pitchford. 1998. Mode of inheritance of triacylglycerol fatty acids in beef adipose tissue. *J. Anim. Sci.* (76 Suppl 1). 593:153.
- Malécot, G. 1948. Mathematics of heredity. [accessed on March 13, 2020]. cabdirect.org.
- Mathur, P. K., B. P. Sullivan, and J. P. Chesnais. 2002. Measuring connectedness: concept and application to a large industry breeding program. *Proc. 7th World Congr. Genet. Appl. to Livest. production, Montpellier, Fr.* Communication No: 20-13.
- Mensink, R. P., and M. B. Katan. 1992. Effect of dietary fatty acids on serum lipids and lipoproteins. A meta-analysis of 27 trials. *Arterioscler. Thromb.* 12:911–919. doi:10.1161/01.atv.12.8.911
- Moghaddar, N., and J. H. J. van der Werf. 2017. Genomic estimation of additive and dominance effects and impact of accounting for dominance on accuracy of genomic evaluation in sheep populations. *J. Anim. Breed. Genet.* 134:453–462. doi:10.1111/jbg.12287
- Momen, M., and G. Morota. 2018. Quantifying genomic connectedness and prediction accuracy from additive and non-additive gene actions. *Genet. Sel. Evol.* 50:45. doi:10.1186/s12711-018-0415-9
- Morota, G., and D. Gianola. 2014. Kernel-based whole-genome prediction of complex traits: a review. *Front. Genet.* 5:363. doi:10.3389/fgene.2014.00363
- Morota, G., M. Koyama, G. J. Rosa, K. A. Weigel, and D. Gianola. 2013. Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. *Genet. Sel. Evol.* 45:17. doi:10.1186/1297-9686-45-17
- Nii, M., T. Hayashi, F. Tani, A. Niki, N. Mori, N. Fujishima-Kanaya, M. Komatsu, K. Aikawa, T. Awata, and S. Mikawa. 2006. Quantitative trait loci mapping for fatty acid composition traits in perirenal and back fat using a Japanese wild boar x Large White intercross. *Anim. Genet.* 37:342–347. doi:10.1111/j.1365-2052.2006.01485.x
- Pegolo, N. T., D. Laloë, H. N. de Oliveira, R. B. Lôbo, and M. N. Fouilloux. 2012. Trends of the genetic connectedness measures among Nellore beef cattle herds. *J. Anim. Breed. Genet.* 129:20–29. doi:10.1111/j.1439-0388.2011.00934.x
- Pérez, P., and G. De Los Campos. 2014. Genome-wide regression and prediction with the BGLR statistical package. *Genetics*. 198:483–495. doi:10.1534/genetics.114.164442
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [accessed on March 13, 2020]. <https://www.R-project.org/>.
- Roso, V. M., F. S. Schenkel, and S. P. Miller. 2004. Degree of connectedness among groups of centrally tested beef bulls. For personal use only. *Can. J. Anim. Sci.* 84:37–47. doi:10.4141/A02-094.
- Saatchi, M., D. J. Garrick, R. G. Tait, Jr, M. S. Mayes, M. Drewnoski, J. Schoonmaker, C. Diaz, D. C. Beitz, and J. M. Reecy. 2013. Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle. *BMC Genomics* 14:730. doi:10.1186/1471-2164-14-730
- Sacks, F. M., and M. Katan. 2002. Randomized clinical trials on the effects of dietary fat and carbohydrate on plasma lipoproteins and cardiovascular disease. *Am. J. Med.* 113(Suppl 9B):13S–24S. doi:10.1016/S0002-9343(01)00987-1
- Serenius, T., K. J. Stalder, and M. Puonti. 2006. Impact of dominance effects on sow longevity. *J. Anim. Breed. Genet.* 123:355–361. doi:10.1111/j.1439-0388.2006.00614.x
- Sun, C., P. M. VanRaden, J. B. Cole, and J. R. O'Connell. 2014. Improvement of prediction ability for genomic selection of dairy cattle by including dominance effects. *PLoS One* 9:e103934. doi:10.1371/journal.pone.0103934
- Tait, R. G., S. Zhang, T. Knight, J. M. Bormann, D. R. Strohbehn, R. G. Tait, S. Zhang, T. Knight, J. M. Bormann, D. R. Strohbehn, et al., 2007. Heritability estimates for fatty acid concentration in angus beef.
- Tarrés, J., M. Fina, and J. Piedrafit. 2010. Connectedness among herds of beef cattle bred under natural service. *Genet. Sel. Evol.* 42:6. doi:10.1186/1297-9686-42-6
- Tonussi, R. L., R. M. O. Silva, A. F. B. Magalhães, R. Espigolan, E. Peripolli, B. F. Olivieri, F. L. B. Feitosa, M. V. A. Lemos, M. P. Berton, H. L. J. Chiaia, et al. 2017. Application of single step genomic BLUP under different uncertain paternity scenarios using simulated data. *PLoS One* 12:e0181752. doi:10.1371/journal.pone.0181752
- Tosh, J. J., and J. W. Wilton. 1994. Effects of data structure on variance of prediction error and accuracy of genetic evaluation. *J. Anim. Sci.* 72:2568–2577. doi:10.2527/1994.72102568x.
- Uemoto, Y., S. Sato, C. Ohnishi, S. Terai, A. Komatsuda, and E. Kobayashi. 2009. The effects of single and epistatic quantitative trait loci for fatty acid composition in a Meishan x Duroc crossbred population. *J. Anim. Sci.* 87:3470–3476. doi:10.2527/jas.2009-1917
- VanRaden, P. M. 1992. Accounting for inbreeding and crossbreeding in genetic evaluation of large populations. *J. Dairy Sci.* 75:3136–3144. doi:10.3168/jds.S0022-0302(92)78077-1
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. doi:10.3168/jds.2007-0980

- Varona, L., A. Legarra, M. A. Toro, and Z. G. Vitezica. 2018. Non-additive effects in genomic selection. *Front. Genet.* 9:78. doi:[10.3389/fgene.2018.00078](https://doi.org/10.3389/fgene.2018.00078)
- Varona, L., and I. Misztal. 1999. Prediction of parental dominance combinations for planned matings, methodology, and simulation results. *J. Dairy Sci.* 82:2186–2191. doi:[10.3168/jds.S0022-0302\(99\)75463-9](https://doi.org/10.3168/jds.S0022-0302(99)75463-9)
- Vitezica, Z. G., L. Varona, and A. Legarra. 2013. On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics* 195:1223–1230. doi:[10.1534/genetics.113.155176](https://doi.org/10.1534/genetics.113.155176)
- Wolak, M. E., and L. F. Keller. 2014. Dominance genetic variance and inbreeding in natural populations. In: *Quantitative genetics in the wild*. Oxford, UK: Oxford University Press. p. 104–127.
- Wright, S. 1922. Coefficients of inbreeding and relationship. *Am. Nat.* 56:330–338. doi:[10.1086/279872](https://doi.org/10.1086/279872)
- Yu, H., and G. Morota. 2020. GCA: an R package for genetic connectedness analysis using pedigree and genomic data. bioRxiv. 696419. doi:[10.1101/696419](https://doi.org/10.1101/696419).
- Yu, H., M. L. Spangler, R. M. Lewis, and G. Morota. 2017. Genomic relatedness strengthens genetic connectedness across management units. *G3 (Bethesda)*. 7:3543–3556. doi:[10.1534/g3.117.300151](https://doi.org/10.1534/g3.117.300151)
- Yu, H., M. L. Spangler, R. M. Lewis, and G. Morota. 2018. Do stronger measures of genomic connectedness enhance prediction accuracies across management units? *J. Anim. Sci.* 96:4490–4500. doi:[10.1093/jas/sky316](https://doi.org/10.1093/jas/sky316)
- Zock, P. L., J. H. M. De Vries, and M. B. Katan. 1994. Impact of myristic acid versus palmitic acid on serum lipid and lipoprotein levels in healthy women and men. *Arterioscler. Thromb. Vasc. Biol.* 14:567–575. doi:[10.1161/01.ATV.14.4.567](https://doi.org/10.1161/01.ATV.14.4.567)