EULA REGINA CARRARA

GENETIC STUDY OF REPRODUCTIVE, DAIRY AND GROWTH TRAITS IN GUZERÁ CATTLE

Thesis presented to the Animal Science Graduate Program of the Universidade Federal de Viçosa, in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

Adviser: Paulo Sávio Lopes

Co-advisers:
Maria Gabriela Campolina Diniz Peixoto
Fabyano Fonseca e Silva (*in memoriam*)

VIÇOSA – MINAS GERAIS
2022

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

C313g
2022

Carrara, Eula Regina, 1987-
Genetic study of reproductive, dairy and growth traits in Guzerá cattle / Eula Regina Carrara. – Viçosa, MG, 2022.
1 tese eletrônica (103 f.): il.

Texto em inglês.
Inclui apêndices.
Orientador: Paulo Sávio Lopes.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Zootecnia, 2022.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2022.473
Modo de acesso: World Wide Web.

1. Bovinos de corte. 2. Bovinos de leite . 3. Avaliação genética. 4. Avaliação genômica. 5. Endogamia. 6. Zebu. I. Lopes, Paulo Sávio, 1958-. II. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-Graduação em Zootecnia. III. Título.

CDD 22. ed. 636.20821

Bibliotecário(a) responsável: Alice Regina Pinto CRB-6/2523

EULA REGINA CARRARA

GENETIC STUDY OF REPRODUCTIVE, DAIRY AND GROWTH TRAITS IN GUZERÁ CATTLE

Thesis presented to the Animal Science Graduate Program of the Universidade Federal de Viçosa, in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

APPROVED: July 29th, 2022.

Assent:

Eula Regina Carrara

Eula Regina Carrara
Author

Paulo Sávio Lopes

Paulo Sávio Lopes
Adviser

ACKNOWLEDGMENTS

My sincere thanks:

To the Federal University of Viçosa (UFV) and the Postgraduate Program in Animal Science, for the opportunity to do the PhD. Certainly, my best choice.

To my advisor Professor Paulo Sávio Lopes, for sharing knowledge, teachings, support, trust, and for being willing to help me whenever I needed it.

To my co-adviser Dr. Maria Gabriela Campolina Diniz Peixoto, who was an important support in my personal development and in the development of this thesis.

To the memory of my co-adviser Professor Fabyano Fonseca e Silva.

To my family for their support: my parents, Márcio and Regina, my brother Márcio Jr. (Nininho), my sister-in-law Elaine, and the greatest joy of my life: Isabela, the best goddaughter in the world; to Miguel and Maria.

To the Brazilian Center for Genetic Improvement of Guzerá (CBMG^2), Embrapa Dairy Cattle and the Brazilian Association of Zebu Breeders (ABCZ) for providing data sets. And to Professor Luiz Josahkian, Dr. Henrique Ventura, Dr. Frank Bruneli, and Dr. Lenira Zadra for making databases available.

To the Genetics and Animal Breeding Discussion Group (GDMA) for the great and valuable discussions during the group meetings, and of course, for the games, parties, barbecues, "arraiás". Especially to the dear: Priscilla, Layla, Pedro, Téo, Thiago. You guys are too much.

To the professors of the Department of Animal Science and the Department of Statistics, especially the professors in Animal Breeding: Professor Renata, Professor Simone, and my contemporaries and newly hired Professors Daniele e Delvan. "The world may not applaud you, but the clearest knowledge of science must recognize that you are the most important professionals in society." (Augusto Cury)

To Fernanda, Gabriel and Venâncio, collaborators of the Animal Science Department of the UFV, who have always taken care of my bureaucratic needs.

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for granting the scholarship, and to the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support.

To all the “Guzeratistas”. This work is for you.

To all who have contributed in any way to this work. Your affection and support help me to go on and on. Thank you very much!

This research was developed with the support of the UFV cluster.

“ (…)

How will we feed an ever-growing population, provide clean water, generate renewable energy, prevent and cure disease and slow down global climate change?

I hope that science and technology will provide the answers to these questions, but it will take people, human beings with knowledge and understanding to implement the solution.

One of the great revelations of the space age has been a perspective that has given humanity on ourselves. When we see the Earth from space, we see ourselves as a whole; we see the unity and not the divisions. It is such a simple image, with a compelling message: one planet, one human race.

We are here together, and we need to live together with tolerance and respect. We must become global citizens.

(…)

We are all time travellers journeying together into the future. But let us work together to make that future a place we want to visit.

Be brave, be determined, overcome the odds. It can be done.

It can be done.”

Stephen Hawking

ABSTRACT

CARRARA, Eula Regina, D.Sc., Universidade Federal de Viçosa, July, 2022. **Genetic study of reproductive, dairy and growth traits in Guzerá cattle.** Adviser: Paulo Sávio Lopes. Co-advisers: Maria Gabriela Campolina Diniz Peixoto and Fabyano Fonseca e Silva (*in memoriam*).

The Guzerá breed is an important Brazilian genetic resource and has been widely used as a pure breed and in crossbreeding strategies to produce animals adapted to tropical climatic conditions. In the last 30 years, some breeders have included selection for dairy traits in addition to selection for beef traits to obtain economic benefits from the sale of meat and milk in dual-purpose production systems. Dual-purpose studies in Guzerá cattle are recent and a previous study showed that there is no genetic antagonism between milk production, growth traits, and age at first calving in these animals, allowing joint selection for these traits. Given the advantages of including marker information such as single nucleotide polymorphisms (SNP) in genetic evaluations, efforts were also made to genotype Guzera animals, particularly those selected for milk and dual-purpose. In this sense, the overall objective of this thesis was to perform genetic and genomic evaluations considering the most important traits selected in Guzerá cattle for both dairy and beef in Guzerá cattle. First, the variance components and genetic parameters for the traits were estimated using bi-trait models. The genetic and phenotypic correlations were all favorable. The absence of genetic antagonism between the reproductive, beef and milk traits indicates that joint selection for beef and dairy traits is possible in the studied population without losses in reproductive traits. Genomic predictions were then performed using single-step genomic best linear unbiased prediction (ssGBLUP) using single-trait models. The performance of genomic prediction (in terms of predictive ability, bias, accuracy of breeding values, rank correlations, and genetic trends) using the ssGBLUP approach compared to the traditional BLUP approach was investigated. This study provided interesting insights into the future application of ssGBLUP in genetic evaluation of Guzerá cattle, especially for dairy traits. The results suggest that ssGBLUP is feasible and can be used in national genetic evaluations for reproductive, dairy, and beef traits in Guzera animals to increase the accuracy of breeding values. However, the number of genotyped animals was not yet sufficient to affect the genetic evaluation in terms of predictive ability and bias. In addition, a third study was conducted to determine the runs of homozygosity (ROH) and calculate the inbreeding coefficients of the animals from the ROH (F_{ROH}). The inbreeding coefficients from pedigree information (F_{PED}) were also estimated and the effects of F_{ROH} and F_{PED} on reproductive, dairy and beef traits in

Guzerá animals were evaluated. Although the Brazilian population of Guzerá animals has experienced historical events such as the bottleneck effect, inbreeding coefficients were low compared to other cattle breeds, even in younger animals. There was no inbreeding depression in the evaluated traits. The results described in this thesis will contribute to the definition of selection criteria for dual-purpose cattle. In addition, this thesis provides the first results on the impact of incorporating information from SNP markers into genetic evaluations of dairy, reproductive and growth traits in Guzerá cattle (in terms of predictive ability, bias, accuracy of breeding values, rank correlations and genetic trends). In summary, our results show that joint selection for beef and dairy traits is feasible in the population studied, that the inclusion of SNP marker information increases the accuracy of animal breeding values, and that the population maintains low levels of inbreeding even during a bottleneck event without significant effects on the most important traits subject to selection.

Keywords: Beef cattle. BLUP. *Bos indicus*. Dairy cattle. Genetic evaluation. Inbreeding. ssGBLUP. Zebu.

RESUMO

CARRARA, Eula Regina, D.Sc., Universidade Federal de Viçosa, julho de 2022. **Estudo genético de características reprodutivas, de leite e de peso em bovinos Guzerá**. Orientador: Paulo Sávio Lopes. Coorientadores: Maria Gabriela Campolina Diniz Peixoto and Fabyano Fonseca e Silva (*in memoriam*).

A raça Guzerá é um importante recurso genético brasileiro e tem sido amplamente utilizada como raça pura e em estratégias de cruzamento para produzir animais adaptados às condições climáticas tropicais. Nos últimos 30 anos, em adição à seleção para produção de carne, alguns produtores têm incluído a seleção para produção de leite visando obter benefícios econômicos da venda de carne e de leite em sistemas de produção de duplo-propósito. Os estudos para duplo-propósito em bovinos Guzerá são recentes e um estudo anterior mostrou que não há antagonismo genético entre produção de leite, características de crescimento e idade ao primeiro parto, nesses animais, viabilizando a seleção conjunta para essas características. Além disso, diante das vantagens da inclusão de informações marcadores do tipo polimorfismos de nucleotídeo único (SNP) em avaliações genéticas, esforços vêm sendo realizados para a genotipagem dos animais Guzerá, principalmente os selecionados para leite e duplo-propósito. Nesse sentido, o objetivo geral dessa tese foi realizar avaliações genéticas e genômicas considerando as principais características sob seleção tanto para leite quanto para corte em bovinos Guzerá. Primeiramente, foi realizada a estimação dos componentes de variância e dos parâmetros genéticos para as características, utilizando-se modelos bicaracterísticos. As correlações genéticas e fenotípicas foram todas favoráveis. O não antagonismo genético entre as características reprodutivas, de corte e de leite demonstra que a seleção conjunta para carne e para leite é viável na população avaliada, sem perdas nas características reprodutivas. Posteriormente, predições genômicas usando melhor predição linear não-viesada genômica em passo único (ssGBLUP) foram feitas utilizando-se modelos unicaracterísticos. Foi investigado o desempenho da predição genômica (em termos de habilidade preditiva, viés, acurácia dos valores genéticos, correlações de *rank* e tendências genéticas) usando abordagem ssGBLUP comparada com a abordagem tradicional sem genômica BLUP. Esse estudo forneceu *insights* interessantes sobre a futura aplicação do ssGBLUP nas avaliações genéticas dos bovinos Guzerá, principalmente para as características leiteiras. Os resultados sugerem que o ssGBLUP é viável e pode ser aplicado em avaliações genéticas nacionais para características reprodutivas, leiteiras e de corte nos animais Guzerá para aumentar a acurácia dos valores genéticos. No entanto, o número de animais genotipados ainda não foi suficiente para influenciar a avaliação genética em termos de habilidade preditiva e viés. Adicionalmente, foi realizado um terceiro estudo para determinar as corridas de homozigose (ROH) e calcular os coeficientes de endogamia dos animais via ROH (F_{ROH}). Foram estimados também os coeficientes de endogamia por meio de informações de genealogia (F_{PED}) e foram avaliados os efeitos de F_{ROH} e F_{PED} sobre características reprodutivas, leiteiras e de corte nos animais Guzerá. Embora a população brasileira de animais Guzerá tenha passado por eventos históricos tais como efeito gargalo, os coeficientes de endogamia foram baixos comparados aos de outras raças bovinas, mesmo em animais mais jovens. Não foi verificada depressão endogâmica nas características avaliadas. Os achados descritos nessa tese contribuirão para a definição dos critérios de seleção de animais com foco no duplo-propósito. Além disso, essa tese fornece os primeiros resultados sobre o impacto da inclusão de informações de marcadores SNP nas avaliações genéticas de características leiteiras, reprodutivas e de corte em bovinos Guzerá (em termos de habilidade preditiva, viés, acurácia dos valores genéticos, correlações de rank e tendências genéticas). Em resumo, nossos resultados demonstram a viabilidade da seleção conjunta para carne e leite na população estudada, que a inclusão de informações de marcadores SNP aumentam a acurácia dos valores genéticos dos animais e que mesmo com um evento de gargalo, a população mantém um baixo nível de endogamia, sem efeitos significativos sobre as principais características em seleção.

Palavras-chave: Avaliação genética. BLUP. Bovinos de corte. Bovinos leiteiros. *Bos indicus*. Endogamia. ssGBLUP. Zebu.

SUMMARY

CHAPTER 1

1.1. General introduction

The Guzerá breed is the third most numerous Zebu breed in Brazil, surpassed only by the Gir and Nelore breeds. This breed is an important genetic resource for tropical animal production because of its good adaptability to different environmental conditions, especially during the dry season and in semi-arid regions, and its use in low-input production systems (Peixoto et al., 2010; Santana et al., 2020).

Initially, Guzerá animals raised in Brazil were genetically selected only for meat production; however, in the last 30 years, breeders have included selection for milk production and have sought production systems for both purposes (Peixoto et al., 2010; dos Santos et al., 2017). In cattle breeds, selection has generally allowed for genetic gain in recent decades, resulting in beef animals with higher and more efficient growth rates and dairy cows with greater capacity for milk production. However, this progress has been accompanied by a decline in animal fertility, making low fertility the primary reason for culling (Wathes et al., 2014). Therefore, it is essential to consider the inclusion of reproductive traits in the selection objectives and to define the selection criteria that allow genetic improvement of the reproductive performance of the herds (Mollah et al., 2015), including the herds of the Guzerá breed.

In general, traditional genetic evaluations obtain genetic parameters and genetic values based on genealogy data, taking into account Wright's kinship. However, with today's widespread availability of single nucleotide polymorphisms (SNP), genetic similarity between individuals can be more accurately determined because measuring relatedness at the genomic level accounts deviations from Mendelian sampling. Thus, genomic information can be incorporated into genetic evaluations by replacing the traditional relationship matrix (matrix **A**;

Henderson, 1976; Quaas, 1976), which is probabilistically obtained from genealogical information, with a matrix that combines both genealogical and SNP information (matrix **H**; Legarra et al., 2009) and includes animals with and without genotype records in a method called single-step genomic best linear unbiased prediction (ssGBLUP; Misztal et al., 2009; Aguilar et al., 2010).

In addition to genetic evaluations, information from SNP markers can also contribute to understand, for example, the genetic basis of inbreeding depression in a population. For example, genetic markers can be used to calculate a genomic relationship matrix (**G**) between individuals, where the expected value of a diagonal element of **G** is 1 + F, where F corresponds to the inbreeding coefficient of the animal (VanRaden, 2008; Yang et al., 2010). This genomic estimate of inbreeding does not suffer from an incomplete pedigree or from errors. It also measures realized inbreeding, which can vary between animals with the same pedigree, i.e., full siblings (Pryce et al., 2014).

A limitation of the inbreeding determined by the **G** matrix is that it do not distinguish between identity by descent (IBD) and identity by state (IBS). An alternative is runs of homozygosity (ROH), defined as continuous homozygous segments of DNA sequence (Gibson et al., 2006) that can be used to identify the location of specific regions of the genome that are IBD (McQuillan et al., 2008). In this sense, inbreeding depression can be studied as a function of inbreeding obtained from ROH.

Several genetic studies have been conducted with the Guzerá breed in Brazil. Genomic studies have been conducted with the Guzerá breed, identifying genetic variants in candidate genes for some traits and imputation methods (Santos et al., 2017; Rosse et al., 2017). Regarding genetic parameters for reproductive and dairy traits of Guzerá breed cattle, no studies involving genomic information have been conducted yet. Using conventional quantitative methods, Santos et al. (2013) and Peixoto et al. (2014) evaluated milk production of primiparous cows born between 1987 and 2009 using a multiple trait and a random regression model, respectively. The study by Cruz et al. (2015) included animals born between 1987 and 2012 and estimated genetic parameters for milk production. Perez et al. (2016, 2017) estimated heritabilities and genetic correlations for reproductive traits of females born between 2005 and 2013. Brito et al. (2018) performed the evaluation of the reproductive trait age at first calving using censored data, including animals born between 1991 and 2012.

When it comes to estimating genetic parameters for growth and reproductive traits for beef animals, there are few studies (Lima Neto et al., 2009; Grupioni et al., 2015; Ferreira et al., 2017; Tramonte et al., 2019). In Tramonte et al. (2019), animals born between 1987 and 2009 were used to estimate genetic parameters and genetic trends for scrotal perimeter and growth traits (weight at 210, 365, and 450 days old).

To date, the study by Brito et al. (2020) was the only work estimating genetic parameters jointly involving reproductive, dairy and growth traits in Guzerá cattle. This work included measurements of animals born between the years 1991 and 2014, analyzed using a multiple trait method without genomic information.

In this thesis, a study was proposed to predict breeding values and estimate genetic parameters for reproductive, dairy and growth traits in Guzerá cattle, based on an updated dataset for 2019, which includes a larger number of phenotyped animals. Predictions of breeding values and estimation of genetic parameters were also performed for the same traits using a genomic approach incorporating SNP information. The latter is not yet available in the literature on Guzerá cattle. In addition, a study comparing estimates of inbreeding pedigree-based and based on runs of homozygosity, as well as the effects of inbreeding on reproductive, dairy and growth traits in Guzerá cattle was conducted.

1.2. Thesis objectives

The general objective with this thesis was conduct a genetic study that includes the estimation of genetic parameters with and without genomic information for reproductive, dairy and growth traits of Guzerá cattle jointly evaluated, and, similarly, prediction of breeding values but by single-trait analyses. Also, to estimate the inbreeding coefficients with and without genomic information and evaluate their effects on the reproductive, dairy and growth traits of Guzerá cattle.

The specific objectives were to: i) Estimate heritabilities and genetic correlations for reproductive traits (age at first calving and scrotal perimeter), dairy traits (milk yield, fat, protein, lactose, and total solids), and growth traits (weight at standard age: weight at 210 days, 365 days, and 450 days) of Guzerá cattle using a bi-trait model, using pedigree-based and genomic-based methods; ii) Predict breeding values for animals using pedigree-based and genomic-based methods for the above reproductive, dairy and growth traits using single-trait analysis, and comparate the two methods in terms of bias, prediction accuracy, accuracy of breeding values, ranking of animals and genetic trends; iii) Estimate inbreeding coefficients using pedigree information and runs of homozygosity, and evaluate inbreeding depression using these coefficients.

1.3. References

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S and Lawlor TJ 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score1. Journal of Dairy Science 93, 743–752.

Brito LC, Casellas J, Varona L, Lopes PS, Ventura HT, Peixoto MGCD, Lázaro SF and Silva FF 2018. Genetic evaluation of age at first calving for Guzerá beef cattle using linear, threshold, and survival Bayesian models. Journal of Animal Science 96, 2517–2524.

Brito LC, Peixoto MGCD, Carrara ER, Fonseca e Silva F, Ventura HT, Bruneli FAT and Lopes PS 2020. Genetic parameters for milk, growth, and reproductive traits in Guzerá cattle under tropical conditions. Tropical Animal Health and Production 52, 2251–2257.

Cruz DAC, Peixoto MGCD, Bruneli FAT, Bignardi AB and El Faro L 2015. Genetic parameters of test-day milk yield in Guzerá cattle under tropical conditions. Genetics and Molecular Research 14, 13618–13624.

Ferreira JL, Bresolin T, Lopes FB, Garcia JAS, Nepomuceno LL, Schmidt AB and Lobo RB 2017. Modelos de regressão aleatória para característica de crescimento em bovinos da raça guzerá. Ciencia Animal Brasileira 18, 1–12.

Gibson J, Morton NE and Collins A 2006. Extended tracts of homozygosity in outbred human populations. Human Molecular Genetics 15, 789–795.

Grupioni NV, Guidolin DGF, Venturini GC, Lôbo RB and Munari DP 2015. Parâmetros genéticos e tendências genéticas para características reprodutivas e de crescimento testicular em bovinos Guzerá. Revista Caatinga 28, 152–160.

Henderson CR 1976. A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. Biometrics 32, 69.

Legarra A, Aguilar I and Misztal I 2009. A relationship matrix including full pedigree and genomic information. Journal of Dairy Science 92, 4656–4663.

Lima Neto HR, Bergmann JAG, Gonçalves TM, Araújo FRC, Bezerra LAF, Saiz RD, Lôbo RB and Silva MA 2009. Parâmetros genéticos para características de carcaça avaliadas por ultrassonografia em bovinos da raça Guzerá. Arquivo Brasileiro de Medicina Veterinaria e Zootecnia 61, 251–258.

McQuillan R, Leutenegger A-L, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A, MacLeod AK, Farrington SM, Rudan P, Hayward C, Vitart V, Rudan I, Wild SH, Dunlop MG, Wright AF, Campbell H and Wilson JF 2008. Runs of Homozygosity in European Populations. The American Journal of Human Genetics 83, 359–372.

Misztal I, Legarra A and Aguilar I 2009. Computing procedures for genetic evaluation including phenotypic , full pedigree , and genomic information. Journal of Dairy Science 92, 4648–4655.

Mollah MFK, Gofur MR, Asaduzzama KM and Bhuiyan MMU 2015. Conception Rate of Non-descript Zebu Cows and its Attributing Factors in Bangladesh. Research Journal of Veterinary Sciences 8, 42–51.

Peixoto MGCD, Poggian CF, Verneque RS, Egito AA, Carvalho MRS, Penna VM, Bergmann JAG, Viccini LF and Machado MA 2010. Genetic basis and inbreeding in the Brazilian Guzerat (*Bos indicus*) subpopulation selected for milk production. Livestock Science 131, 168–174.

Peixoto MGCD, Santos DJ de A, Borquis RRA, Bruneli FÂT, Panetto JC do C and Tonhati H 2014. Random regression models to estimate genetic parameters for milk production of Guzerat cows using orthogonal legendre polynomials. Pesquisa Agropecuaria Brasileira 49, 372–383.

Perez BC, Peixoto MGCD, Bruneli FT, Ramos PVB and Balieiro JCC 2016. Genetic analysis of oocyte and embryo production traits in Guzerá breed donors and their associations with age at first calving. Genetics and Molecular Research 15, 1–9.

Perez BC, Silva FF, Ventura R V., Bruneli FAT, Balieiro JCC and Peixoto MGDC 2017. Count Bayesian models for genetic analysis of in vitro embryo production traits in Guzerá cattle. animal 11, 1440–1448.

Pryce JE, Haile-Mariam M, Goddard ME and Hayes BJ 2014. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genetics Selection Evolution 46, 71.

Quaas RL 1976. Computing the Diagonal Elements and Inverse of a Large Numerator Relationship Matrix. Biometrics 32, 949.

Rosse IC, Assis JG, Oliveira FS, Leite LR, Araujo F, Zerlotini A, Volpini A, Dominitini AJ, Lopes BC, Arbex WA, Machado MA, Peixoto MGCD, Verneque RS, Martins MF, Coimbra RS, Silva MVGB, Oliveira G and Carvalho MRS 2017. Whole genome sequencing of Guzerá cattle reveals genetic variants in candidate genes for production, disease resistance, and heat tolerance. Mammalian Genome 28, 66–80.

Santana ML, Pereira RJ, Bignardi AB, El Faro L, Pires MFÁ, Andrade RG, Perez BC, Bruneli FAT and Peixoto MGCD 2020. Dual-purpose Guzerá cattle exhibit high dairy performance under heat stress. Journal of Animal Breeding and Genetics 137, 486–494.

Santos FC, Peixoto MGCD, Fonseca PA de S, Pires M de FÁ, Ventura RV, Rosse I da C, Bruneli FAT, Machado MA and Carvalho MRS 2017. Identification of Candidate Genes for Reactivity in Guzerat (*Bos indicus*) Cattle: A Genome-Wide Association Study. PLOS ONE 12, e0169163.

Santos DJA, Peixoto MGCD, Aspilcueta Borquis RR, Verneque RS, Panetto JCC and Tonhati H 2013. Comparison of random regression models to estimate genetic parameters for milk production in Guzerat (*Bos indicus*) cows. Genetics and Molecular Research 12, 143–153.

Tramonte NC, Grupioni NV, Stafuzza NB, Guidolin DGF, Savegnago RP, Bezerra LAF, Lôbo RB and Munari DP 2019. Genetic parameters, genetic trends, and principal component analysis for productive and reproductive traits of Guzera beef cattle. Revista Brasileira de Zootecnia 48.

VanRaden PM 2008. Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science 91, 4414–4423.

Wathes DC, Pollott GE, Johnson KF, Richardson H and Cooke JS 2014. Heifer fertility and carry over consequences for life time production in dairy and beef cattle. animal 8, 91–104.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME and Visscher PM 2010. Common SNPs explain a large proportion of the heritability for human height. Nature Genetics 42, 565–569.

CHAPTER 2

Genetic study of quantitative traits supports use of Guzerá as dual-purpose cattle

Carrara, E.R.[1], Peixoto, M.G.C.D.[2], Veroneze, R.[1], Silva, F.F[1], Ramos, P.V.B.[1], Bruneli, F.A.T.[2], Zadra, L.E.F[3], Ventura, H.T.[4], Josahkian, L.A.[4], and Lopes, P.S.[1]

[1] Department of Animal Sciences, Universidade Federal de Viçosa, Viçosa, MG, Brazil

[2] Embrapa Dairy Cattle, Juiz de Fora, MG, Brazil.

[3] Brazilian Center for the Genetic Improvement of Guzerá, Belo Horizonte, MG, Brazil.

[4] Brazilian Association of Zebu Cattle, Uberaba, MG, Brazil.

2.1. Abstract

The aim of this study was to estimate genetic parameters for 305-day cumulative milk yield and components, growth, and reproductive traits in Guzerá cattle. The evaluated traits were 305-day first-lactation cumulative yields (kg) of milk (MY305), fat (FY305), protein (PY305), lactose (LY305), and total solids (SY305); age at first calving (AFC) in days; adjusted scrotal perimeter (cm) at the ages of 365 (SP365) and 450 (SP450) days; and adjusted body weight (kg) at the ages of 210 days (W210), 365 (W365) and 450 (W450) days. The (co)variance components were estimated using the restricted maximum likelihood method regarding single-trait, bi-trait and tri-trait analyses. Contemporary groups and additive genetic effects were included in the general mixed model. Maternal genetic and maternal permanent environmental effects were also included for W210. The direct heritability estimates ranged from 0.16 (W210) to 0.32 (MY305). The maternal heritability estimate for W210 was 0.03. Genetic correlation estimates among milk production traits and among growth traits ranged from 0.92 to 0.99 and from 0.92 to 0.99, respectively. For milk production and growth traits, the genetic correlations ranged from 0.33 to 0.56. The genetic correlations among AFC and all other traits were negative (-0.43 to -0.27). Scrotal perimeter traits and body weights showed genetic correlations ranging from 0.41 to 0.46. Scrotal perimeter and milk production traits showed genetic correlations ranging from 0.11 to 0.30. The phenotypic correlations were similar in direction (same sign) and lower than the corresponding genetic correlations. These results suggest the viability and potential of joint selection for dairy and growth traits in Guzerá cattle, taking into account reproductive traits.

Keywords: Genetic parameters; Dairy cattle; Beef cattle; Sexual precocity; *Bos indicus*

2.2. Introduction

The Guzerá cattle breed is one of the most productive among the Zebu breeds present in Brazil and are widely used in crossbreeding with the aim of producing animals adapted to tropical climatic conditions. In the last 30 years, in addition to genetic selection for beef production, some breeders have included selection for milk yield in Guzerá cattle to obtain economic benefits from the sale of both meat and milk in dual-purpose production systems (Bruneli et al., 2020; Peixoto et al., 2021).

Dual-purpose cattle systems allow a wide variety of production models because they depend on the preferences of the farmer, local market, household consumption, and local environment, resulting in different proportions of incomes generated from sales of meat and milk (Galué et al., 2008; Albarrán-Portillo et al., 2015). Furthermore, systems with both beef and milk production constitute a subsistence system and an important activity for the economic development of small holders in Latin America (Rangel et al., 2020).

In cattle breeds, in general, selection has allowed large genetic gain that has resulted in beef animals with higher growth rates and dairy cows with higher milk production capacity. In dual-purpose systems, it is important to understand the relationship between growth and milk production traits to evaluate possible genetic antagonisms before defining selection objectives.

Furthermore, genetic progress focused on traits of highest economic value in cattle was accompanied by declines in animal fertility, making low fertility the main reason for cow disposal (Wathes et al., 2014). Therefore, it is essential to include reproductive traits in genetic selection processes. In the near future, traits related to feed supply and quality, adaptive robustness, animal welfare, and functionality should also be considered for sustainable selection.

In Brazilian beef cattle, the most commonly used trait as an indicator of fertility and sexual precocity is the scrotal perimeter (SP), which is easy to measure, and it shows a positive genetic correlation with testicular biometry traits (e.g., length, width, and testicle volume), sperm motility and mass activity and a negative genetic correlation with major and total sperm defects (Carvalho Filho et al., 2020). In turn, age at first calving (AFC) is highly related to fertility and reproductive efficiency in both beef and dairy cattle, thus affecting heifer precocity (Berry and Evans, 2014; Claus et al., 2017). The AFC presents heritability of moderate magnitude, approximately 0.20, and favorable (i.e., negative) genetic correlation with milk yield and growth traits in Guzerá cattle (Brito et al., 2020).

Although the genetic improvement programs of the Guzerá breed for milk and beef are developed independently, joint genetic evaluation is feasible due to the high genetic connection between pedigree datasets of the breeding programs (Peixoto et al., 2021). In addition, favorable genetic correlations among dairy (milk yield), growth (weaning weight, yearling weight, and weight at 550 days), and reproductive (age at first calving) traits were previously reported in Guzerá cattle (Brito et al., 2020).

Studies on the dual-purposes in Guzerá cattle are recent. It is expected that joint genetic evaluations considering milk and beef traits will contribute to the genetic improvement of the dual-purpose in this breed. In this context, this study aimed to estimate genetic parameters for 305-day cumulative milk yield and components, growth, and reproductive traits in Guzerá cattle.

2.3. Material and methods

2.3.1. Animals and data

The phenotypic and pedigree data used in this study came from the Zebu Breeds Genealogical Registry Service (SRGRZ), databases of the Brazilian Association of Zebu Cattle (ABCZ), and the National Program for the Improvement of Guzerá for Dairy Purpose

(PNMGuL), jointly coordinated by Embrapa Dairy Cattle and the Brazilian Center for the Genetic Improvement of Guzerá (CBMG^2).

Records from 197,283 Guzerá males and females born between 1954 and 2018 were used in this study. The pedigree file comprised 545,310 animals, including 148,231 dams and 8,859 sires from 18 generations. The first generation was assumed to be the base population.

The traits evaluated were 305-day first-lactation cumulative yields (kg) of milk (MY305), fat (FY305), protein (PY305), lactose (LY305), and total solids (SY305); age at first calving (AFC) in days; adjusted scrotal perimeter (cm) at the ages of 365 (SP365) and 450 (SP450) days; and adjusted body weight (kg) at the ages of 210 days (W210), 365 (W365) and 450 (W450) days.

The body weights were adjusted to account for the age of the animal at measurement and previous weight, according to the BIF Guidelines (BIF Guidelines, 2018). The scrotal perimeter was adjusted to 365 and 450 days by nonlinear logistic regression, a method used in the official genetic evaluations of the breed. A total of 64,050 SP measurements were available from 29,604 animals (one to ten measurements per animal).

To estimate the nonlinear model parameters, the Gauss-Newton iterative method in nonlinear regression (NLIN) from SAS Software (SAS Institute, 2018) was used. The logistic model used is described as follows:

$$SP_t = \frac{A}{1+B(\exp(-kt))} + e,$$

where SP_t is the scrotal perimeter at t days of age; A is the estimated SP at maturity; B indicates the proportion of asymptotic mature testis size to be obtained after birth (established by the initial values of SP and t); k is maturing index, establishing the rate with which SP approaches A; and e random error associated with each measurement. The final estimates of the parameters A, B, and k, were 40.1638, 2.7315, and 0.0035, respectively.

Contemporary groups (CGs) for MY305, FY305, PY305, LY305, and SY305 were formed by fitting herd, year, and season of calving. For AFC, SP365, SP450, W210, W365, and W450, the contemporary groups were a combination of herd, year, and season of birth. Additionally, for growth traits, sex was also included in the CG. Both calving and birth seasons were defined as dry (April to September) or rainy (October to March). Data that belonged to the CG with fewer than three records were excluded. The number of CGs per trait and the number of animals by CG are shown in Table 1, as well as descriptive statistics for each studied trait.

Table 1. Number of observations (N), mean, standard deviation (SD), coefficient of variation (CV, %), minimum (MIN), maximum (MAX), number of contemporary groups (CG), and range of number of animals by CG (N by CG) included in the analyses of each trait

Traits[a]	**N**	**Mean**	**SD**	**CV**	**MIN**	**MAX**	**CG**	**N by CG**
MY305 (kg)	5,229	1991.17	980.90	49.26	105.00	6487.00	516	3 to 67
FY305 (kg)	1,853	82.82	37.06	44.74	6.00	281.17	181	3 to 51
PY305 (kg)	1,543	62.08	27.16	43.74	4.00	232.00	113	3 to 51
LY305 (kg)	1,457	77.50	34.15	44.07	5.00	226.00	108	3 to 51
SY305 (kg)	1,228	227.02	96.30	42.42	13.00	722.00	93	3 to 51
AFC (days)	83,244	1251.50	208.84	16.69	671.00	1680.00	9,592	3 to 197
W210 (kg)	122,684	173.70	39.92	22.98	50.00	300.00	9,974	3 to 188
W365 (kg)	88,065	227.41	53.36	23.46	67.00	414.00	7,683	3 to 169
W450 (kg)	88,456	275.75	58.58	21.25	102.27	499.00	7,710	3 to 169
SP365 (cm)	26,988	21.55	3.45	15.99	9.17	33.98	1,511	3 to 229
SP450 (cm)	27,047	23.85	3.75	15.74	10.25	37.98	1,519	3 to 230

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solid yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight; SP365 = 365-day scrotal perimeter; and SP450 = 450-day scrotal perimeter.

2.3.2. Genetic analyses

The (co)variance components were obtained by the restricted maximum likelihood method (REML) using AIREMLF90 software (Misztal et al., 2014). The following general model was assumed:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1a} + \mathbf{Z_2m} + \mathbf{Z_3mp} + \mathbf{e},$$

where $\mathbf{y}$ is the vector of phenotypes; $\mathbf{b}$ is the vector of fixed effects of a contemporary group and covariate; $\mathbf{a}$ is the vector of random additive direct genetic effects; $\mathbf{m}$ is the vector of random maternal genetic effects; $\mathbf{mp}$ is the vector of random maternal permanent environmental effects; $\mathbf{Z_1}$, $\mathbf{Z_2}$ and $\mathbf{Z_3}$ are incidence matrices related to the $\mathbf{a}$, $\mathbf{m}$ and $\mathbf{mp}$ to $\mathbf{y}$, respectively; and $\mathbf{e}$ is the residual vector. The following (co)variance structures were also assumed:

$$\mathrm{Var}\begin{bmatrix}\mathbf{a}\\ \mathbf{m}\\ \mathbf{mp}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{G_{0\mathrm{a}}} \otimes \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G_{0\mathrm{m}}} \otimes \mathbf{A} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{MP_0} \otimes \mathbf{I_{mp}} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R_0} \otimes \mathbf{I}\end{bmatrix},$$

in which $\mathbf{G_{0\mathrm{a}}}$ is the additive direct genetic (co)variance matrix; $\mathbf{G_{0\mathrm{m}}}$ is the maternal genetic (co)variance matrix; $\mathbf{A}$ is the numerator relationship matrix; $\mathbf{MP_0}$ is the maternal permanent environment (co)variance matrix; $\mathbf{I}_{\mathrm{mp}}$ is an identity matrix of order equal to the number of dams with progeny; $\mathbf{R_0}$ is the residual (co)variance matrix; $\mathbf{I}$ is an identity matrix of order equal to the number of animals; and $\otimes$ is the Kronecker product.

Single-trait (for direct and maternal heritabilities) and bi-trait (for genetic, phenotypic, and residual correlations) analyses were carried out. Specifically, for the correlations between milk production traits (i.e., MY305, FY305, PY305, LY305, and SY305) and scrotal perimeter at different ages (i.e., SP365 and SP450), tri-trait analyses were performed using W210 as an "anchor" trait. This procedure was performed due to the difference in the number of animals phenotyped in these two groups of traits. Therefore, W210 was used to promote better connections between measured traits since it had a higher number of observations and covers both sexes. The genetic covariances and correlations presented for W210 were from bi-trait analyses.

The maternal genetic and maternal permanent environmental effects were adjusted only for W210. The direct maternal genetic covariances were assumed to be zero. The age at calving was considered a linear covariate for MY305, FY305, PY305, LY305, and SY305, and the age of dams at calving was considered a linear covariate for W210, W365, and W450.

2.4. Results

The variance components obtained by single-trait analysis for evaluated traits are given in Table 2. Direct heritability and genetic and phenotypic correlation estimates among the studied traits are presented in Table 3.

The direct heritability estimates ranged from 0.16 (W210) to 0.32 (MY305). The maternal heritability estimate for W210 was 0.03 (0.004). The genetic correlations among all traits evaluated were favorable. Genetic correlations among milk production traits (milk yield and milk components) were positive and high, ranging from 0.92 (FY305 x LY305) to 0.99 (MY305 x LY305). For W210, W365, and W450, the genetic correlations were also positive and high, ranging from 0.92 (W210 x W365 and W210 x W450) to 0.99 (W365 x W450). When comparing milk production traits and growth traits, the estimated genetic correlations were moderate, ranging from 0.33 (SY305 x W210) to 0.56 (FY305 x W365 and PY305 x W365).

Table 2. Estimates of variance components and their standard errors (within parentheses) for 305-day milk yield, 305-day milk components, growth traits, and reproductive traits studied in Guzerá cattle obtained by single-trait analysis.

Traits[a]	σ^2_a	σ^2_m	σ^2_{mp}	σ^2_e
MY305	158,150.00 (20,124.00)	-	-	340,320.00 (15,534.00)
FY305	202.53 (51.61)	-	-	714.74 (45.89)
PY305	144.11 (33.11)	-	-	345.67 (26.90)
LY305	251.66 (57.49)	-	-	570.79 (46.45)
SY305	2,037.10 (513.24)	-	-	4,621.40 (421.14)
AFC	5,912.50 (264.82)	-	-	23,596.00 (222.58)
W210	109.05 (5.67)	20.64 (2.64)	70.09 (3.08)	493.48 (4.18)
W365	304.12 (12.46)	-	-	952.60 (9.68)
W450	360.46 (14.98)	-	-	1,194.40 (11.71)
SP365	1.48 (0.13)	-	-	5.83 (0.10)
SP450	1.78 (0.15)	-	-	6.89 (0.12)

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solid yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight; SP365 = 365-day scrotal perimeter; SP450 = 450-day scrotal perimeter; σ^2_a = additive genetic variance; σ^2_m = maternal genetic variance; σ^2_{mp} = maternal permanent environmental variance; and σ^2_e = residual variance.

Table 3. Direct heritability estimates (diagonal), genetic correlations (above diagonal) and phenotypic correlations (below diagonal), and their standard errors (within parentheses) for 305-day milk yield, 305-day milk components, growth traits, and reproductive traits studied in Guzerá cattle obtained by bi-trait and tri-trait analysis**.

Traits[a]	**MY305**	**FY305**	**PY305**	**LY305**	**SY305**	**AFC**	**W210**	**W365**	**W450**	**SP365**	**SP450**
MY305	**0.32** **(0.04)**	0.96 (0.024)	0.97 (0.016)	0.99 (0.001)	0.98 (0.013)	-0.28 (0.079)	0.35 (0.081)	0.44 (0.075)	0.40 (0.078)	0.30** (0.100)	0.26** (0.102)
FY305	0.86 (0.006)	**0.22** **(0.05)**	0.95 (0.025)	0.92 (0.034)	0.98 (0.014)	-0.31 (0.139)	0.43 (0.122)	0.56 (0.105)	0.54 (0.107)	0.22** (0.161)	0.16** (0.165)
PY305	0.90 (0.004)	0.91 (0.005)	**0.29** **(0.06)**	0.96 (0.013)	0.98 (0.009)	-0.29 (0.138)	0.42 (0.124)	0.56 (0.103)	0.53 (0.117)	0.22** (0.155)	0.11** (0.123)
LY305	0.92 (0.003)	0.90 (0.006)	0.98 (0.001)	**0.31** **(0.06)**	0.98 (0.009)	-0.29 (0.138)	0.37 (0.129)	0.53 (0.108)	0.50 (0.111)	0.26** (0.156)	0.17** (0.163)
SY305	0.92 (0.004)	0.96 (0.003)	0.96 (0.003)	0.95 (0.003)	**0.31** **(0.07)**	-0.27 (0.151)	0.33 (0.141)	0.50 (0.119)	0.47 (0.122)	0.22** (0.171)	0.13** (0.178)
AFC	-0.07 (0.021)	-0.12 (0.035)	-0.12 (0.039)	-0.11 (0.040)	-0.17 (0.042)	**0.20** **(0.01)**	-0.39 (0.034)	-0.41 (0.030)	-0.43 (0.031)	-0.36** (0.049)	-0.40** (0.047)
W210	0.06 (0.031)	0.01 (0.046)	0.05 (0.050)	0.04 (0.051)	0.09 (0.057)	-0.27 (0.007)	**0.16** **(0.01)**	0.92 (0.007)	0.92 (0.008)	0.45 (0.046)	0.41 (0.046)
W365	0.09 (0.034)	0.07 (0.050)	0.06 (0.055)	0.08 (0.057)	0.18 (0.062)	-0.36 (0.007)	0.72 (0.002)	**0.24** **(0.01)**	0.99 (0.001)	0.44 (0.042)	0.42 (0.042)

W450	0.08	0.07	0.08	0.10	0.18	-0.35	0.71	0.99	**0.23**	0.46	0.44
	(0.033)	(0.049)	(0.054)	(0.056)	(0.061)	(0.007)	(0.002)	(0.001)	**(0.01)**	(0.042)	(0.042)
SP365	0.08	0.05	0.06	0.07	0.06	-0.07	0.40	0.55	0.54	**0.20**	0.96
	(0.026)	(0.037)	(0.040)	(0.041)	(0.044)	(0.010)	(0.007)	(0.006)	(0.006)	**(0.02)**	(0.007)
SP450	0.07	0.04	0.03	0.05	0.03	-0.08	0.39	0.54	0.53	0.88	**0.21**
	(0.027)	(0.038)	(0.042)	(0.043)	(0.046)	(0.010)	(0.007)	(0.006)	(0.006)	(0.002)	**(0.02)**

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solid yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight; SP365 = 365-day scrotal perimeter; and SP450 = 450-day scrotal perimeter. **Exclusively for the correlation between milk production traits (i.e., MY305, FY305, PY305, LY305, and SY305) and scrotal perimeter at different ages (i.e., SP365 and SP450); tri-trait analyses were performed using W210 as the “anchor” trait, e.g., for the comparison between MY305 and SP365, MY305xW210xSP365 was performed and analogous for the others. For all other traits, bi-trait analyses were performed two by two. The genetic parameters presented for W210 were from bi-trait analyses.

Regarding reproductive traits, the genetic correlations among AFC and all other traits evaluated were negative, presenting values ranging from -0.31 (AFC x FY305) to -0.27 (AFC x SY305) with milk production traits and from -0.43 (AFC x W450) to -0.39 (AFC x W210) with growth traits. Body weights and scrotal perimeter measured at different ages showed positive genetic correlations, ranging from 0.41 (W210 x SP450) to 0.46 (W450 x SP365). The genetic correlation between SP365 and AFC was -0.36, and that between SP450 and AFC was -0.40. Low to moderate genetic correlations were estimated between scrotal perimeter at different ages and milk production traits, ranging from 0.11 (PY305 x SP450) to 0.30 (MY305 x SP365).

Phenotypic correlation estimates among all traits were similar in direction and smaller than their corresponding genetic correlations. Overall, the phenotypic and residual correlations were similar in direction and magnitude, except between milk production traits and growth traits. Among milk production traits, the phenotypic correlations were high and positive, ranging from 0.86 (MY305 x PY305) to 0.98 (PY305 x LY305). All traits presented negative and low to moderate phenotypic correlation estimates with AFC. For correlations with milk production traits, values ranged from -0.17 (AFC x SY305) to -0.07 (AFC x MY305) and with growth traits, values ranged from -0.36 (AFC x W365) to -0.27 (AFC x W210). Between AFC and SP365 and SP450, the phenotypic correlations were equal to -0.07 and -0.08, respectively.

Considering growth traits and scrotal perimeter measured at different ages, the phenotypic correlation estimates were positive, presenting moderate magnitudes with values from 0.39 (W210 x SP450) to 0.55 (W365 x SP365). Growth and scrotal perimeter traits presented positive and low to moderate phenotypic correlation estimates with milk production traits, ranging from 0.01 (FY305 x W210) to 0.18 (SY305 x W365, and SY305 x W450). The phenotypic correlation between SP365 and SP450 was positive and high (0.88).

The residual correlation estimates are shown in Table 4. The residual correlation estimates were high among milk production traits and among growth traits, ranging from 0.84 (MY305 x FY305) to 0.98 (PY305 x LY305) and from 0.66 (W210 x W450) to 0.98 (W365 x W450), respectively. Between growth traits and scrotal perimeter traits, the residual correlation estimates were moderate and ranged from 0.42 (W210 x SP365, and W210 x SP450) to 0.59 (W365 x SP365). AFC had a negative residual correlation with all traits evaluated, with values ranging from -0.34 (AFC x W365) to -0.001 (AFC x MY305). The residual correlation estimates between milk production traits and growth traits were low, ranging from -0.15 (PY305 x W365) to 0.05 (SY305 x W365). When different traits are not observed in the same individual, there is no residual covariance between the traits (Mrode and Thompson, 2014). Thus, residual correlation estimates between milk production traits and scrotal perimeter traits and between AFC and scrotal perimeter traits were not computed.

Table 4. Residual correlations and their standard errors (within parentheses) for 305-day milk yield, 305-day milk components, growth traits, and reproductive traits estimated for Guzerá cattle obtained by bi-trait analyses.

Traits[a]	**MY305**	**FY305**	**PY305**	**LY305**	**SY305**	**AFC**	**W210**	**W365**	**W450**	**SP365**	**SP450**
MY305	-	0.84 (0.011)	0.89 (0.008)	0.89 (0.007)	0.90 (0.007)	-0.001 (0.030)	-0.03 (0.047)	-0.05 (0.050)	-0.05 (0.048)	-	-
FY305	-	-	0.90 (0.009)	0.89 (0.010)	0.93 (0.007)	-0.07 (0.047)	-0.14 (0.067)	-0.11 (0.073)	-0.09 (0.071)	-	-
PY305	-	-	-	0.98 (0.002)	0.95 (0.054)	-0.07 (0.051)	-0.06 (0.076)	-0.15 (0.085)	-0.11 (0.083)	-	-
LY305	-	-	-	-	0.94 (0.006)	-0.05 (0.056)	-0.07 (0.079)	-0.12 (0.089)	-0.07 (0.086)	-	-
SY305	-	-	-	-	-	-0.14 (0.059)	0.02 (0.087)	0.05 (0.096)	0.06 (0.093)	-	-
AFC	-	-	-	-	-	-	-0.27 (0.009)	-0.34 (0.010)	-0.33 (0.010)	-	-
W210	-	-	-	-	-	-	-	0.67 (0.003)	0.66 (0.003)	0.42 (0.011)	0.42 (0.011)
W365	-	-	-	-	-	-	-	-	0.98 (0.0002)	0.59 (0.011)	0.58 (0.011)
W450	-	-	-	-	-	-	-	-	-	0.56 (0.010)	0.56 (0.010)

SP365	-	-	-	-	-	-	-	-	-	-	0.86 (0.003)
SP450	-	-	-	-	-	-	-	-	-	-	-

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solid yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight; SP365 = 365-day scrotal perimeter; and SP450 = 450-day scrotal perimeter.

2.5. Discussion

This study estimated genetic associations among milk yield, milk components, growth traits, and reproductive traits in Guzerá cattle. To the best of our knowledge, the study by Brito et al. (2020) is the only study reporting the genetic correlations among milk, growth, and reproductive traits in this breed. In the present study, however, a larger number of animals and important traits, such as milk components and scrotal perimeter, for the genetic improvement of dual-purpose Guzerá were included in the analyses. The milk components, especially major components, i.e., fat, protein, and lactose, are important because they affect directly milk properties in industrial processing and in the quality of dairy products. In addition, SP is important because it is easy to measure and has a negative-favorable genetic correlation with sexual precocity in both males and females (Terakado et al., 2015; Buzanskas et al., 2017), making it a potential selection criterion for herds seeking sexual precocity. Thus, the present study is conclusive and of great importance for the genetic improvement of the breed.

Overall, heritabilities were moderate (0.16 to 0.32), indicating that all traits studied could respond to selection with lesser or greater intensity and could achieve a satisfactory genetic progress rate in the breeding program of Guzerá cattle. The heritability estimate for MY305 (0.32) was similar to that reported by Brito et al. (2020) (0.29), and it was higher than those reported by Santos et al. (2013) (0.24) and Gama et al. (2016) (0.24) in studies also conducted on Guzerá cattle. It must be highlighted that heritability is a property of a population and its mating system in a specific environment and time. Different heritability estimates among studies on the same breed can be explained by differences in the sample size, type of records used in each study (e.g., first or multiple lactations), and estimation methods.

Heritability estimates for milk components (0.22 to 0.31) were greater than those obtained by Silva et al. (2020) in Guzerá cattle. Although reproductive traits are strongly influenced by environmental components, the heritability estimate for AFC was moderate

(0.20). Our results reflect, however, that this trait in Guzerá cattle has a larger genetic effect; thus, selection for improvements in reproductive performance is possible. Breeding value for AFC was recently included in the Guzerá sire summary to increase precocity; therefore, we expect soon to be able to evaluate the genetic response to selection. This is an important trait to be included as a selection goal since the Guzerá breed is still reproductively late.

Regarding SP365 and SP450, the heritability estimates were moderate (0.20 and 0.21, respectively) and were lower than those found in the literature for beef cattle, including Guzerá (Buzanskas et al., 2017; Tramonte et al., 2019). For growth traits, the heritability estimates were moderate (0.16 to 0.24). They corroborated the estimates reported by Gama et al. (2016), but they were lower than those in most of the consulted literature for Guzerá cattle (Tramonte et al., 2019; Brito et al., 2020). Although low, the maternal heritability for W210 (0.03) was similar to that reported by other studies on Guzerá cattle (Tramonte et al., 2019; Brito et al., 2020).

Animal breeding strategies are determined by the relative importance of traits and the genetic correlation between traits. In regard to the dual-purpose focus, there have been few studies that have addressed genetic relationships between dairy and beef traits. The genetic correlations between MY305 and 305-day cumulative milk components were high (0.92 to 0.99). No studies have reported genetic correlations between MY305 and 305-day cumulative milk components in Guzerá cattle. A high genetic correlation indicates a strong genetic association among these traits, i.e., selection for higher MY305 promotes increased 305-day cumulative milk components.

Among growth traits, the genetic correlations were high (0.92 to 0.99), emphasizing that most genes controlling weight at a given age are the same at other ages, making it possible to select animals at younger ages. The genetic correlations among growth traits found in the present study corroborate the literature on cattle breeds (Kluska et al., 2018; Tramonte et al., 2019; Brito et al., 2020).

The genetic correlation estimates between AFC and all evaluated traits were low to moderate and favorable. Between AFC and MY305, the genetic correlation was equal to -0.28. Although of moderate magnitude, this result suggests that females selected for MY305 would improve their sexual precocity. To the best of our knowledge, no other studies on other Zebu cattle breeds have reported genetic correlations between AFC and MY305. Genetic selection involving milk yield and AFC should be performed with caution to avoid an increasing risk of dystocia in younger cows, mainly those with high genetic merit for milk yield (Stefani et al., 2021). The genetic correlations among AFC and milk components were low (0.31 to 0.27) but favorable, suggesting that the selection process for AFC would result in a milk yield with higher 305-day cumulative solids content or vice versa. Thus, considering the high estimates obtained for heritabilities of milk constituents in this study, the selection for these traits could reach higher genetic gains for these traits.

Genetic correlations between AFC and growth traits were low but favorable (0.43 and 0.39). In turn, genetic correlations between AFC and scrotal perimeter at different ages (i.e., SP365 and SP450) were 0.36 and 0.40, respectively, which are values consistent with those reported in the literature for beef cattle (Buzanskas et al., 2017; Kluska et al., 2018). The endocrine axis regulating puberty in bulls and heifers is similar, which may explain the existence of genetic correlations between male and female reproductive traits, which can be exploited in breeding programs (Land, 1973; Hull and Harvey, 2014). Thus, considering all traits of interest for the improvement of Guzerá cattle, it is possible to reach the highest genetic progress for female precocity using a scrotal perimeter as a selection criterion without impairing genetic progress regarding body weight.

Regarding scrotal perimeter, the genetic correlations with MY305 and 305-day cumulative milk components (0.11 to 0.30) were low to moderate, suggesting little genetic association among these traits. Thus, selection for MY305 and milk components would not increase scrotal perimeter, but will not cause harm either. Some studies reported genetic correlations among SP and growth traits, with estimates ranging from 0.30 to 0.78 (Kluska et al., 2018; Tramonte et al., 2019). In the present study, the genetic correlations between SP traits and growth traits were moderate and ranged from 0.41 to 0.46, indicating that the selection to increase body weights at different ages also increases SP365 and SP450 and vice versa. The positive and favorable genetic correlation between SP365 and SP450 found in the present study (0.96) and in other studies on beef cattle (0.90 to 0.94) (Buzanskas et al., 2017; Tramonte et al., 2019) indicated that one-year-old bulls with a higher SP also had a higher scrotal perimeter at a yearling age. The genetic correlations involving SP365 suggest that its use as a selection criterion in breeding programs is advantageous because in addition to allowing the early selection of animals, it does not promote losses to the other traits, with the possibility of a genetic response in female precocity.

Between milk production traits and growth traits, the genetic correlations were moderate and favorable (0.33 to 0.56) and were highest among W365 and milk production traits (i.e., MY305, FY305, PY305, LY305, and SY305), ranging from 0.44 to 0.56. These results suggest that selection for MY305 and milk components would also result in animals with higher weight at one year of age. Few studies have evaluated genetic relationships between milk and growth traits in cattle, but our results suggest the existence of pleiotropic effects among these traits. The estimates found in the current study are slightly higher than those reported by Gama et al. (2016) and Brito et al. (2020), both with Guzerá cattle.

Specifically for milk components genetically correlated with AFC, growth traits, and scrotal perimeter traits, the standard errors of the estimates were high. This can be explained by data differences among the traits, with milk components having a much lower number of observations and fewer animals with phenotypes for all traits evaluated. The standard errors were slightly higher for genetic correlations between milk components and scrotal perimeter traits, possibly due to the difference in the number of animals recorded in the two groups of traits. Thus, the genetic correlation estimates among these traits should be interpreted with caution.

The phenotypic correlation estimates behaved the same way as the corresponding genetic correlation estimates, but in general, they had lower magnitudes than the genetic correlations. Phenotypic correlation estimates among milk, growth, and reproductive traits were smaller in magnitude than genetic correlation estimates in other studies involving genetic parameter estimates in dual-purpose breeds (Mazza et al., 2016; Brito et al., 2020). A phenotypic correlation lower than its corresponding genetic correlation with a small positive residual correlation can be explained by the fact that the genes underlying the two traits are similar, but the environments in which they express and that influence these traits have a low correlation (Searle, 1961). In the current study, this condition occurred only among SY305 and growth traits.

Comparing the milk production traits (i.e., MY305, FY305, PY305, LY305, and SY305), the residual correlations were positive and high (0.84 to 0.98), suggesting that residual effects (i.e., nonadditive and environmental) similarly influence these traits. For these traits, the genetic and their corresponding phenotypic correlations were similar in sign and magnitude. The same occurred among the beef traits (i.e., W210, W365, W450, SP365, and SP450), where the residual correlations ranged from 0.42 to 0.98.

The confidence of breeders in the genetic evaluation results of the National Program for the Improvement of Guzerá for Dairy Purpose has increased the number of herds and animals participating in this program, as well as the use of this breed in dual-purpose production systems

(Peixoto et al., 2014). The non-antagonism among milk, growth and reproductive traits in Guzerá cattle makes simultaneous selection for beef and dairy traits feasible. Each breeder has the option to specialize the herd for milk, beef, or both purposes. The genetic gains for each trait will not be the same as that of direct selection for only one of them, but joint selection is possible.

2.6. Conclusion

All evaluated traits presented moderate heritability estimates and presented the possibility of responding to selection in the Guzerá breed.

Furthermore, the favorable genetic correlations among all evaluated traits suggest the possibility of joint genetic selection for milk and beef production in Guzerá cattle, with a favorable correlated response in reproductive traits.

2.7. Acknowledgments

The authors acknowledge the Centro Brasileiro de Melhoramento Genético do Guzerá (CBMG^2) and Associação Brasileira de Criadores de Zebu (ABCZ) for providing datasets. The authors also acknowledge the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and INCT (Instituto Nacional de Ciência e Tecnologia) for financial support.

2.8. References

Albarrán-Portillo B, Rebollar-Rebollar S, García-Martínez A, Rojo-Rubio R, Avilés-Nova F and Arriaga-Jordán CM 2015. Socioeconomic and productive characterization of dual-purpose farms oriented to milk production in a subtropical region of Mexico. Tropical Animal Health and Production 47, 519–523.

Berry DP and Evans RD 2014. Genetics of reproductive performance in seasonal calving beef cows and its association with performance traits. Journal of Animal Science 92, 1412–1422.

BIF Guidelines 2018. Guidelines for uniform beef improvement programs (Ninth Edition - Revised March, 2018).

Brito LC, Peixoto MGCD, Carrara ER, Fonseca e Silva F, Ventura HT, Bruneli FAT and Lopes PS 2020. Genetic parameters for milk, growth, and reproductive traits in Guzerá cattle under tropical conditions. Tropical Animal Health and Production 52, 2251–2257.

Bruneli FAT, Peixoto MGCD, Júnior MLS, Pereira RJ, Arbex WA, Penna VM, Zadra LEF, Verneque R da S, Lôbo RB and Carvalho MRS 2020. National Breeding Program of Guzerá Cattle for Milk: Progeny Testing, National Zootechnical Archive and MOET Nucleus Results. Brazilian Agricultural Research Corporation Documentos, 251, Juiz de Fora: Embrapa Dairy Cattle.

Buzanskas ME, Pires PS, Chud TCS, Bernardes PA, Rola LD, Savegnago RP, Lôbo RB and Munari DP 2017. Parameter estimates for reproductive and carcass traits in Nelore beef cattle. Theriogenology 92, 204–209.

Carvalho Filho I, Marques DBD, Campos CF, Guimarães JD, Guimarães SEF, Lopes PS, Silva FF e and Veroneze R 2020. Genetic parameters for fertility traits in Nellore bulls. Reproduction in Domestic Animals 55, 38–43.

Claus LAM, Koetz Júnior C, Roso VM, Borges MHF, Barcellos JOJ and Ribeiro EL de A 2017. Genetic parameters of age at first calving, weight gain, and visual scores in Nelore heifers. Revista Brasileira de Zootecnia 46, 303–308.

Galué FU, Peña ME, Rincón R, Romero J and Rendón-Ortín M 2008. Gestión y tecnología en sistemas ganaderos de doble propósito (Taurus-Indicus). Revista Cientifica de la Facultad de Ciencias Veterinarias de la Universidad del Zulia 18, 715–724.

Gama MPM, Ventura HT, Alencar Pereira M, El Faro L and Paz CCP 2016. 0369 Genetic associations between milk production and growth traits in Guzerat breed. Journal of Animal Science 94, 178–178.

Hull KL and Harvey S 2014. Growth Hormone and Reproduction: A Review of Endocrine and Autocrine/Paracrine Interactions. International Journal of Endocrinology 2014, 1–24.

Kluska S, Olivieri BF, Bonamy M, Chiaia HLJ, Feitosa FLB, Berton MP, Peripolli E, Lemos MVA, Tonussi RL, Lôbo RB, Magnabosco C de U, Di Croce F, Osterstock J, Pereira ASC, Munari DP, Bezerra LA, Lopes FB and Baldi F 2018. Estimates of genetic parameters for growth, reproductive, and carcass traits in Nelore cattle using the single step genomic BLUP procedure. Livestock Science 216, 203–209.

Land RB 1973. The Expression of Female Sex-limited Characters in the Male. Nature 241, 208–209.

Mazza S, Guzzo N, Sartori C and Mantovani R 2016. Genetic correlations between type and test-day milk yield in small dual-purpose cattle populations: The Aosta Red Pied breed as a case study. Journal of Dairy Science 99, 8127–8136.

Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A and Vitezica Z 2014. Manual for BLUPF90 family of programs. Athens, USA. Retrieved on 14 February 2022, from http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.

Mrode RA and Thompson R 2014. Linear models for the prediction of animal breeding values. CAB International, Wallingford, Oxon, UK.

Peixoto MGCD, Carvalho MRS, Egito AA, Steinberg RS, Bruneli FÂT, Machado MA, Santos FC, Rosse IC and Fonseca PAS 2021. Genetic Diversity and Population Genetic Structure of a Guzerá (*Bos indicus*) Meta-Population. Animals 11, 1125.

Peixoto MGCD, Santos DJ de A, Borquis RRA, Bruneli FÂT, Panetto JC do C and Tonhati H 2014. Random regression models to estimate genetic parameters for milk production of Guzerat cows using orthogonal legendre polynomials. Pesquisa Agropecuaria Brasileira 49, 372–383.

Rangel J, Perea J, De-Pablos-Heredero C, Espinosa-García JA, Mujica PT, Feijoo M, Barba C and García A 2020. Structural and Technological Characterization of Tropical Smallholder Farms of Dual-Purpose Cattle in Mexico. Animals 10, 86.

Santos DJA, Peixoto MGCD, Borquis RRA, Verneque RS, Panetto JCC and Tonhati H 2013. Genetic parameters for test-day milk yield, 305-day milk yield, and lactation length in Guzerat cows. Livestock Science 152, 114–119.

SAS Institute 2018. SAS/STAT® 15.1 User’s Guide.

Searle SR 1961. Phenotypic, Genetic and Environmental Correlations. Biometrics 17, 474.

Silva RPA da, Lôbo RNB, El Faro L, dos Santos GG, Bruneli FÂT and Peixoto MGCD 2020. Genetic parameters for somatic cell count (SCC) and milk production traits of Guzerá cows using data normalized by different procedures. Tropical Animal Health and Production 52, 2513–2522.

Stefani G, Brandão Aquaroli D, Batista Gonçalves Costa Júnior J, Santana Júnior ML, Tonhati H, Sesana RC and El Faro L 2021. Genetic parameters for dystocia, milk yield and age at first calving in Brazilian Holstein cows. Journal of Applied Animal Research 49, 1–5.

Terakado APN, Boligon AA, Baldi F, Silva JAI V. and Albuquerque LG 2015. Genetic associations between scrotal circumference and female reproductive traits in Nelore cattle1. Journal of Animal Science 93, 2706–2713.

Tramonte NC, Grupioni NV, Stafuzza NB, Guidolin DGF, Savegnago RP, Bezerra LAF, Lôbo RB and Munari DP 2019. Genetic parameters, genetic trends, and principal component analysis for productive and reproductive traits of Guzera beef cattle. Revista Brasileira de Zootecnia 48.

Wathes DC, Pollott GE, Johnson KF, Richardson H and Cooke JS 2014. Heifer fertility and carry over consequences for life time production in dairy and beef cattle. animal 8, 91–104.

CHAPTER 3

Genomic prediction in Brazilian Guzerá cattle: application of a single-step approach to productive and reproductive traits

Carrara, E.R.[1], Peixoto, M.G.C.D.[2], Veroneze, R.[1], Silva, A.A.[3], Bruneli, F.A.T.[2], Zadra, L.E.F.[4], Ventura, H.T.[5], Josahkian, L.A.[5], and Lopes, P.S.[1]

[1] Department of Animal Sciences, Universidade Federal de Viçosa, Viçosa, MG, Brazil

[2] Embrapa Dairy Cattle, Juiz de Fora, MG, Brazil.

[3] School of Agricultural and Veterinarian Sciences, São Paulo State University, Jaboticabal, SP, Brazil

[4] Brazilian Center for the Genetic Improvement of Guzerá, Belo Horizonte, MG, Brazil.

[5] Brazilian Association of Zebu Cattle, Uberaba, MG, Brazil.

3.1. Abstract

The Guzerá breed represents the third largest indicine purebred and one of the most productive among the Zebu breeds in Brazil. At present, Guzerá herds are genetically selected by conventional breeding methods based on progeny testing and Multiple Ovulation and Embryo Transfer nucleus scheme. Efforts are being made to genotype Guzerá animals in Brazil, however, there are no studies on the inclusion of genomic information in the genetic evaluations of Guzerá cattle. Thus, this study aimed to investigate the feasibility of genomic prediction using the single-step approach for productive and reproductive traits in Guzerá cattle. Evaluations included the 305-day first-lactation cumulative yields (kg) of milk (MY305), fat (FY305), protein (PY305), lactose (LY305), and total solids (SY305); age at first calving in days (AFC); and adjusted body weight (kg) at the ages of 210 days (W210), 365 (W365) and 450 (W450) days, from a database containing 197,283 measurements from Guzerá males and females born between 1954 and 2018. The pedigree included 433,823 animals spanning up to 14 overlapping generations. A total of 1,618 genotyped animals for 43,943 single nucleotide polymorphisms (SNP) were used in analyses. The (co)variance components and the breeding values were obtained using pedigree information (pedigree-based) and genomic information (genomic-based). The genomic method used was single-step genomic best linear unbiased predictor (ssGBLUP). Contemporary groups and additive genetic effects were included in the general mixed model. Maternal genetic and permanent environmental effects were also included for W210. Prediction ability and bias were accessed using cross-validation. Individual accuracies of breeding values, rank correlations, and genetic trends were obtained for both models. Overall, the predictive abilities were similar between the methods and ranged from 0.30 to 0.45 for the pedigree-based model, and from 0.27 to 0.47 for the genomic-based model. The bias also was similar between the methods, ranging from 0.96 to 1.41 in pedigree-based model, and from 0.88 to 1.35 in genomic-based model. The mean individual accuracy had an evident increase in genomic evaluation, with values ranging from 0.26 to 0.54 in the pedigree-based model, and from 0.41 to 0.56 in the genomic-based model. No relevant changes were observed for both sexes in the rank of animals, however, considering the youngest animals, the rank correlations were lower than 0.91, reaching 0.67. Even if based on a small number of genotyped animals and a small database for some traits, provided interesting insights about the future application of single-step genomic BLUP in the Brazilian Guzerá cattle. The results suggest that the single-step genomic BLUP is feasible and may be applied to national genetic evaluations of the breed to increase the accuracy of breeding values, but still without great impact on predictive ability and bias.

Keywords: Genetic parameters; Genetic trends; Genomic prediction; ssGBLUP; Zebu

3.2. Introduction

The Guzerá breed, with a population of approximately 500,000 registered individuals (ABCZ, 2020), represents the third largest indicine purebred and one of the most productive among the Zebu breeds in Brazil. Guzerá animals have been selected for meat, milk, or both production traits in the country, where they are considered a dual-purpose breed (Peixoto et al., 2010). The breed has been widely used as a purebred and in crossbreeding strategies for producing animals adapted to tropical climatic conditions, focusing on dual-purpose production systems (Stafuzza et al., 2017). Although the genetic improvement programs of the Guzerá breed for milk and beef are developed independently, several bulls are "double proven", i.e., they are genetically evaluated for dairy and beef traits (Bruneli et al., 2020).

At present, Guzerá herds are genetically selected by conventional breeding methods developed by the Brazilian Center for the Genetic Improvement of Guzerá (CBMG[2]), based on progeny testing and MOET (Multiple Ovulation and Embryo Transfer) nucleus scheme

(Bruneli et al., 2020). From the genomic perspective, efforts are being made to genotype Guzerá animals in Brazil, and the genetic markers have been used in studies involving genetic diversity and population genetic structure (Fonseca et al., 2016; Peixoto et al., 2021); identification of genetic variations (Rosse et al., 2017; Stafuzza et al., 2017); and genome-wide association studies (dos Santos et al., 2017; Paiva et al., 2020). However, there are no studies on the genomic information inclusion in the genetic evaluations of Guzerá cattle.

The genotyping of animals for genetic markers allow estimation of realized genetic relationships between individuals through identity-by-state. Thereby, the genomic prediction can increase the accuracy of breeding values mainly for young animals, reducing generation interval with consequent increase in genetic gain (Lopes et al., 2018).

Single-step genomic best linear unbiased prediction (ssGBLUP) is a unified approach that simultaneously combines phenotypic, pedigree, and genomic information to generate more accurate predictions of genomic-estimated breeding value (GEBV; Misztal et al., 2009). This method involves replacing the pedigree-based relationship matrix (**A**; Henderson, 1976; Quaas, 1976) in the mixed model equations (traditional BLUP) by a relationship matrix which combines pedigree and genomic relationships (**H**; Legarra et al., 2009; Misztal et al., 2009). It has been successfully applied in many livestock species, including dairy cattle (Matilainen et al., 2018; Oliveira et al., 2019; Silva et al., 2019) and beef cattle (Lourenco et al., 2015; Mehrban et al., 2019; Tonussi et al., 2021). Similarly, the single-step genomic restricted maximum likelihood (ssGREML) can be used to estimate variance components considering simultaneously phenotyped, genotyped and non-genotyped animals (Veerkamp et al., 2011; Forni et al., 2011; Cesarani et al., 2018), in other words using the **H** matrix.

Studies investigating the feasibility of genomic prediction in Guzerá remain scarce, particularly due to the limited number of genotyped animals. It is expected that the advantages of genomic selection observed in others breeds dairy or beef cattle will also be observed in

Guzerá. Therefore, the aim of this study was to investigate the performance of genomic prediction using the single-step approach for productive and reproductive traits in Guzerá cattle.

3.3. Material and methods

3.3.1. Phenotypes and pedigree

The data used in this study were provided by Zebu Breeds Genealogical Registry Service (SRGRZ) of the Brazilian Association of Zebu Cattle (ABCZ), and by the National Program for the Improvement of Guzerá for Dairy Purpose (PNMGuL), coordinated jointly by Embrapa Dairy Cattle and the Brazilian Center for the Genetic Improvement of Guzerá (CBMG^2).

The phenotypic data comprised 197,283 measurements of productive and reproductive traits from Guzerá males and females born between 1954 and 2018. Evaluations included the 305-day first-lactation cumulative yields (kg) of milk (MY305), fat (FY305), protein (PY305), lactose (LY305), and total solids (SY305); age at first calving in days (AFC); and adjusted body weight (kg) at the ages of 210 days (W210), 365 (W365) and 450 (W450) days. Data from animals without valid measurements or with measurements outside the acceptable range (mean ± three standard deviations) were excluded.

Contemporary groups (CGs) for MY305, FY305, PY305, LY305, and SY305 were formed by combining herd, year, and season of calving. For AFC, W210, W365, and W450, the contemporary groups were a combination of herd, year, and season of birth. Additionally, for growth traits, sex was also included in the CG. Both calving and birth seasons were defined as dry (April to September) or rainy (October to March). Data from CG with fewer than three records and with only one sire as a parent were excluded. The number of records, the descriptive statistics of the datasets after editing, and the number of CG by trait are shown in Table 1.

Table 1. Number of animals with record (and genotyped), mean and standard deviation (SD), minimum (MIN) and maximum (MAX), and the number of contemporary groups (CG) included in the analyses of each trait

Traits[a]	**Number of animals with record (and genotyped)**	**Mean ± SD**	**MIN**	**MAX**	**CG**
MY305 (kg)	5,229 (976)	1991.17 ± 980.90	105.00	6487.00	516
FY305 (kg)	1,853 (621)	82.82 ± 37.06	6.00	281.17	181
PY305 (kg)	1,543 (629)	62.08 ± 27.16	4.00	232.00	113
LY305 (kg)	1,457 (599)	77.50 ± 34.15	5.00	226.00	108
SY305 (kg)	1,228 (513)	227.02 ± 96.30	13.00	722.00	93
AFC (days)	83,244 (1,029)	1251.50 ± 208.84	671.00	1680.00	9,592
W210 (kg)	122,684 (609)	173.70 ± 39.92	50.00	300.00	9,974
W365 (kg)	88,065 (574)	227.41 ± 53.36	67.00	414.00	7,683
W450 (kg)	88,456 (574)	275.75 ± 58.58	102.27	499.00	7,710

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids224 yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight.

The pedigree included 433,823 animals (121,497 dams and 7,200 sires) spanning up to 14 overlapping generations. Animals with both known sire and dam made up 77.47% of the pedigree, 4.10% had only sire known, 15.68% had only dam known, and 2.75% had both sire and dam unknown. The average pedigree-based inbreeding (F_{PED}) was 0.007±0.02 considering the complete pedigree, and F_{PED} was 0.030±0.04 considering only the inbreeding animals (disregarding F_{PED}=0).

Although dairy traits present a smaller number of records, and, consequently, fewer animals related with the phenotyped ones in the pedigree, non-pruning of the pedigree was carried for any of the analyzed traits.

3.3.2. Genotypes

A total of 1,734 animals (54 bulls and 1,680 cows) were genotyped using different single nucleotide polymorphisms (SNP) chips. Nine hundred seventy-three samples were genotyped using the Illumina BovineSNP50 v2 BeadChip, including 54,609 SNPs, 50 sample genotyped using the Illumina BovineHD BeadChip, including 777,962 SNPs; (Illumina, Inc. San Diego, California), and the remaining 711 samples were genotyped using the GeneSeek Genomic Profiler *indicus* chip (GGP *indicus*; Neogen Corp., Lansing, MI), including 54,791 SNPs.

A genotype quality control (QC) was performed separately for each chip to keep only autosomal SNPs with call rate higher than 0.95, minor allele frequency (MAF) higher than 0.02. SNPs with extreme departure from the Hardy-Weinberg Equilibrium (estimated as the difference between expected and observed frequency of heterozygous > 0.15) were also discarded. Samples were required to have a call rate higher than 0.95, parent-progeny pairs were tested for Mendelian conflicts, and samples with inconsistent identification were excluded.

After filtering, 1,618 genotyped animals (37 bulls and 1,581 cows) remained for further analyses, being 879 cows with autosomal 29,705 SNPs from BovineSNP50; 34 bulls with autosomal 514,505 SNPs from BovineHD; and 705 animals (702 cows and 3 bulls) with 44,149 SNPs from GGP *indicus*.

The GGP *indicus* is a panel built specifically for *Bos indicus* cattle and for this reason, it was prioritized as the reference panel in an imputation analysis to use all genotypic information available. The common SNPs between the BovineSNP50 and the GGP *indicus* chip were 7,173, and between BovineHD and the GGP *indicus* chip were 41,206. The missing genotypes were imputed to GGP *indicus* panel using the FImpute version 3 (Sargolzaei et al., 2014).

To evaluate the imputation accuracy, the animals from GGP *indicus* panel (N = 705 animals) were divided into reference and validation sets based on their birth year. The reference set comprised 508 animals that were born before 2015. The validation set comprised 197 animals that were born after 2015. The validation set had their genotypes masked according to each imputed panel.

Imputation accuracy was estimated based on Spearman's correlation coefficient between imputed and true SNP markers, and the concordance rate (CR) was assessed as a proportion of correctly imputed SNPs in relation to all imputed SNPs. In general, imputation accuracy was higher than 0.91 and CR higher than 92.9%. After imputation analyses, a new QC was performed using imputed genotype data and following the same parameters as above. Finally, a total of 43,943 autosomal SNPs and 1,618 animals (37 bulls and 1,581 cows) were retained for further analyses. All animals with both record and genotype were female.

3.3.3. Statistical modeling

In genomic-based analysis, the (co)variance components were obtained by ssGREML using single-trait (for direct and maternal heritabilities) and bi-trait (for genetic and phenotypic correlations) models. The variance components from single-trait genomic-based model were used in prediction of genomic breeding values (GEBV) by ssGBLUP method. The variance components used in prediction of pedigree-based breeding values (EBV) were obtained from single-trait model by REML in a previous study and more details can be found in the previous chapter, also published in Carrara et al. (2022).

The following mixed linear model was assumed:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1a} + \mathbf{Z_2m} + \mathbf{Z_3mp} + \mathbf{e},$$

where $\mathbf{y}$ is the vector of phenotypes; $\mathbf{b}$ is the vector of fixed effects of a contemporary group and covariate; $\mathbf{a}$ is the vector of random additive direct genetic effects; $\mathbf{m}$ is the vector of random maternal genetic effects; $\mathbf{mp}$ is the vector of random maternal permanent environmental effects, with $\mathbf{mp} \sim \mathrm{N}(0, \mathbf{MP_0} \otimes \mathbf{I_{mp}})$, where $\mathbf{MP_0}$ is the maternal permanent environment (co)variance matrix; $\mathbf{I}_{\mathrm{mp}}$ is an identity matrix of order equal to the number of dams with progeny; $\mathbf{Z_1}$, $\mathbf{Z_2}$ and $\mathbf{Z_3}$ are incidence matrices related to the $\mathbf{a}$, $\mathbf{m}$ and $\mathbf{mp}$ to $\mathbf{y}$, respectively; and $\mathbf{e}$ is the residual vector, with $\mathbf{e} \sim \mathrm{N}(0, \mathbf{R_0} \otimes \mathbf{I})$, where $\mathbf{R_0}$ is the residual (co)variance matrix; $\mathbf{I}$ is an identity matrix of order equal to the number of animals. In the pedigree-based model, it was assumed $\mathbf{a} \sim \mathrm{N}(0, \mathbf{G_{0}}_{\mathrm{a}} \otimes \mathbf{A})$ and $\mathbf{m} \sim \mathrm{N}(0, \mathbf{G_{0}}_{\mathrm{m}} \otimes \mathbf{A})$, whereas in the genomic-based model it was assumed $\mathbf{a} \sim \mathrm{N}(0, \mathbf{G_{0}}_{\mathrm{a}} \otimes \mathbf{H})$ and $\mathbf{m} \sim \mathrm{N}(0, \mathbf{G_{0}}_{\mathrm{m}} \otimes \mathbf{H})$, in which $\mathbf{G_{0}}_{\mathrm{a}}$ is the additive direct genetic (co)variance matrix; $\mathbf{G_{0}}_{\mathrm{m}}$ is the maternal genetic (co)variance matrix; $\mathbf{A}$ is the numerator relationship matrix; $\mathbf{H}$ is a matrix that combines pedigree and genomic information; and $\otimes$ is the Kronecker product.

The maternal effects (genetic and permanent environment) were assumed only for W210. For MY305, FY305, PY305, LY305, and SY305, the age at calving was considered a linear covariate. For W210, W365, W450, the age of dams at calving was considered linear covariate.

The inverse of the matrix $\mathbf{H}$ was created as (Aguilar et al., 2010; Christensen and Lund, 2010):

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

where $\mathbf{A^{-1}}$ is the inverse of the pedigree-based numerator relationship matrix, $\mathbf{G}$ is the genomic relationship matrix; and $\mathbf{A_{22}}$ is the proportion of $\mathbf{A}$ for the genotyped animals. The matrix $\mathbf{G}$ was obtained based on the first method proposed by VanRaden (2008), where the allele frequencies were calculated based on the genotyped animals information. The default scaling ($\tau = 1$ and $\omega = 1$) and weighting ($\alpha = 0.95$ and $\beta = 0.05$) factors were used.

3.3.4. Prediction ability and bias

Five-fold cross-validation was performed to assess the predictive ability of breeding values using BLUP and by ssGBLUP. The animals with phenotypes and genotypes were randomly allocated to five mutually exclusive subsets (fold). Then, four of them were used as a training group and one used as validation group (whose phenotypes were not included in the analysis). The analysis was performed five times for each trait; thus, each subset was considered as validation group once. The subsets contained approximately the same number of animals.

The corrected phenotype (y^*) for each trait and animal was calculated as the sum of the breeding value and the residual ($\mathrm{y}^* = \hat{\mathrm{a}} + \hat{\mathrm{e}}$) for both BLUP ($\mathrm{y}^*_{\mathrm{BLUP}}$) and ssGBLUP ($\mathrm{y}^*_{\mathrm{ssGBLUP}}$) analyses. The Pearson's correlation between EBV or GEBV and the respective corrected phenotypes of the animals in the validation population was defined as the predictive ability. The mean of predictive abilities of the five subsets from cross-validation by trait were computed.

To assess the prediction bias, the corrected phenotypes were regressed on EBV or GEBV ($\mathrm{y}^*_{\mathrm{BLUP}} = \mathrm{b}_0 + \mathrm{b}_1\mathrm{EBV}$, and $\mathrm{y}^*_{\mathrm{ssGBLUP}} = \mathrm{b}_0 + \mathrm{b}_1\mathrm{GEBV}$). The b_0 and b_1 are the intercept and the slope regression coefficients, respectively. The b_1 was used as an indicator of bias with an expected value of one.

3.3.5. Individual accuracy

The individual accuracies of EBV or GEBV, were obtained considering all animals and considering the youngest animals (birth year ≥ 2015) with genotype and without phenotype in both BLUP and ssGBLUP analyses following the equation (Van Vleck, 1993; Mrode and Thompson, 2014):

$$\mathrm{ACC}_{\mathrm{j}} = \sqrt{1 - \frac{\mathrm{SEP}_{\mathrm{j}}^2}{(1+\mathrm{F}_{\mathrm{j}})\sigma_{\mathrm{a}}^2}},$$

where $\mathrm{ACC_j}$ is the accuracy of EBV or GEBV for an individual *j*; $\mathrm{SEP_j}$ is the standard error of prediction of EBV or GEBV for animal *j*, obtained from the square root of diagonal elements of the inverse of the left-hand side of mixed model equations (Misztal et al., 2014); $\mathrm{F_j}$ is the inbreeding coefficient of animal *j*; and σ_a^2 is the additive genetic variance from pedigree-based or genomic-based analyses. The inbreeding coefficients from the genomic-based model were obtained using the Method 2 of Legarra et al. (2020).

3.3.6. Rank correlation and genetic trend

Spearman's correlations between EBVs and GEBVs considering all males, all females, the top 5%, 25%, and 50% animals within each sex, and the youngest animals were calculated for each trait to assess possible reranking of animals. The youngest animals were born in and after 2015 and, they have genotype and do not have phenotype.

Additionally, the genetic trends were estimated by regressing the animals GEBV or EBV on birth year considering separately males and females. Segmented linear regressions were performed based on the date of the first sire summary (2000s) (Bruneli et al., 2020; ANCP, 2022).

3.3.7. Software

The statistical analyses were performed using the BLUPF90 family programs (Misztal et al., 2014): the files containing the allele calls of samples were converted using Illumina2preGS; the quality control was performed using PreGSf90; the (co)variance components were estimated using AIREMLF90; the inbreeding coefficients from the ssGBLUP were obtained using PreGSf90; the breeding values were predicted using BLUPF90; and the corrected phenotypes were obtained using PREDICTF90.

3.4. Results

3.4.1. Genetic parameter estimates

Direct heritability and genetic and phenotypic correlation estimates among the studied traits obtained by genomic-based analyses are presented in Table 2. The variance components obtained by genomic-based analyses can be viewed in Supplementary Table 1.

The direct heritability estimates ranged from 0.16 (W210) to 0.39 (SY305). The maternal heritability estimate for W210 was 0.03 (0.004). Heritability estimates were highest in the milk production traits, ranging from 0.24 (FY305) to 0.39 (SY305), while for AFC it was equal to 0.20, and for growth traits they ranged from 0.16 (W210) to 0.24 (W365).

The genetic correlations among all traits evaluated were favorable. Genetic correlations among milk production traits were positive and high, ranging from 0.92 (FY305 x LY305) to 0.99 (MY305 x LY305). The growth traits also presented positive and high genetic correlations, ranging to 0.92 (W210 x W365 and W210 x W450) to 0.99 (W365 x W450). The genetic correlations between milk production traits and growth traits were positive, moderate, and presented values from 0.33 (SY305 x W210) to 0.56 (FY305 x W365 and PY305 x W365). The genetic correlations among AFC and all other evaluated traits were negative and favorable, presenting values ranging from -0.31 (AFC x FY305) to -0.27 (AFC x SY305) for milk production traits and from -0.43 (AFC x W450) to -0.39 (AFC x W210) for growth traits.

Phenotypic correlation estimates among all traits were similar in direction and smaller than their corresponding genetic correlations. Among milk production traits, phenotypic correlations ranged from 0.86 (MY305 x FY305) to 0.97 (PY305 x LY305), and among growth traits ranged from 0.71 (W210 x W450) to 0.99 (W365 x W450). The AFC presented phenotypic correlations ranging from -0.36 (AFC x W365) to -0.06 (AFC x MY305) with all other traits. Between milk production traits and growth traits, the phenotypic correlations ranged from 0.01 (FY305 x W210) to 0.16 (SY305 x W365 and SY305 x W450).

Table 2. Direct heritability estimates (diagonal), genetic correlations (above diagonal) and phenotypic correlations (below diagonal), as well as their standard errors (within parentheses), for 305-day milk yield, 305-day milk components, growth traits, and reproductive traits studied in Guzerá cattle obtained by ssGREML

Traits[a]	**MY305**	**FY305**	**PY305**	**LY305**	**SY305**	**AFC**	**W210**	**W365**	**W450**
MY305	**0.36** **(0.036)**	0.96 (0.024)	0.97 (0.016)	0.99 (0.001)	0.98 (0.013)	-0.28 (0.079)	0.35 (0.081)	0.44 (0.075)	0.40 (0.078)
FY305	0.86 (0.006)	**0.24** **(0.053)**	0.95 (0.025)	0.92 (0.034)	0.98 (0.014)	-0.31 (0.139)	0.43 (0.122)	0.56 (0.105)	0.54 (0.107)
PY305	0.90 (0.004)	0.91 (0.005)	**0.32** **(0.059)**	0.96 (0.013)	0.98 (0.009)	-0.29 (0.138)	0.42 (0.124)	0.56 (0.103)	0.53 (0.117)
LY305	0.92 (0.003)	0.90 (0.010)	0.97 (0.002)	**0.34** **(0.061)**	0.98 (0.009)	-0.29 (0.138)	0.37 (0.129)	0.53 (0.108)	0.50 (0.111)
SY305	0.92 (0.004)	0.94 (0.003)	0.96 (0.003)	0.95 (0.003)	**0.39** **(0.070)**	-0.27 (0.151)	0.33 (0.141)	0.50 (0.119)	0.47 (0.122)
AFC	-0.06 (0.021)	-0.11 (0.035)	-0.10 (0.040)	-0.08 (0.040)	-0.15 (0.042)	**0.20** **(0.008)**	-0.39 (0.034)	-0.41 (0.030)	-0.43 (0.031)
W210	0.07 0.030)	0.01 0.045)	0.05 (0.049)	0.04 (0.051)	0.09 (0.056)	-0.27 (0.007)	**0.16** **(0.008)**	0.92 (0.007)	0.92 (0.008)
W365	0.10 0.033)	0.07 (0.025)	0.06 (0.054)	0.07 (0.057)	0.16 (0.062)	-0.36 (0.007)	0.72 (0.002)	**0.24** **(0.009)**	0.99 (0.001)
W450	0.08 (0.032)	0.08 (0.049)	0.08 (0.053)	0.08 (0.056)	0.16 (0.061)	-0.35 (0.007)	0.71 (0.002)	0.99 (0.001)	**0.23** **(0.009)**

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight.

Additionally, the averages relationship based on pedigree (A_{22}) and on genomic (G) in studied population were 0.0407 and 0.0445, respectively.

3.4.2. Prediction ability and bias

The average predictive ability and bias of the EBVs (pedigree-based) and GEBVs (genomic-based) for milk production, reproductive, and growth traits are shown in Table 3. Overall, the predictive abilities were similar between the methods and ranged from 0.30 to 0.45 for pedigree-based model, and from 0.27 to 0.47 for genomic-based model.

A slight increase was observed for dairy traits, which have a higher proportion of genotyped animals. The bias values also were similar between the methods. For pedigree-based model bias ranged from 0.96 to 1.41, and for genomic-based model the values ranged from 0.88 to 1.35.

The difference between pedigree-based and genomic-based approaches were more evident when evaluating the individual accuracy (Table 4). The mean individual accuracy for all animals ranged from 0.26 to 0.54 in the pedigree-based model, and from 0.41 to 0.56 in the genomic-based model. Considering youngest animals, the individual accuracy ranged from 0.28 to 0.50 and from 0.43 to 0.55 in the pedigree-based model and genomic-based model, respectively.

The gain in accuracy was greater for milk production traits considering all and the youngest animals. For milk production traits, the gain in accuracy ranged from 28% (MY305) to 116% (SY305) considering all animals, and from 33% (MY305) to 96% (SY305) for youngest animals. For AFC and growth traits, the gain in accuracy ranged from 0% (AFC) to 4% (W450) for all animals, and from 4% (AFC) to 7% (W365 and W450) for youngest animals. As expected, the gain in accuracy was higher for the young animals for most traits.

Table 3. Average of predictive abilities and of bias (with standard deviation; SD) of EBVs (pedigree-based) and GEBVs (genomic-based) estimated for milk production, reproductive, and growth traits in Guzerá cattle using five-fold cross validation.

	Predictive abilities		**Bias**	
Traits[a]	**Pedigree-based**	**Genomic-based**	**Pedigree-based**	**Genomic-based**
MY305	0.45 (0.08)	0.47 (0.08)	1.28 (0.22)	1.25 (0.19)
FY305	0.30 (0.11)	0.31 (0.14)	1.25 (0.56)	1.22 (0.59)
PY305	0.33 (0.10)	0.33 (0.07)	1.10 (0.35)	1.02 (0.21)
LY305	0.36 (0.08)	0.39 (0.07)	1.16 (0.21)	1.19 (0.15)
SY305	0.33 (0.04)	0.40 (0.07)	1.14 (0.21)	1.31 (0.37)
AFC	0.30 (0.04)	0.27 (0.03)	0.96 (0.16)	0.88 (0.16)
W210	0.45 (0.04)	0.43 (0.04)	1.41 (0.33)	1.35 (0.27)
W365	0.39 (0.15)	0.41 (0.14)	1.06 (0.41)	1.15 (0.41)
W450	0.36 (0.13)	0.38 (0.11)	1.02 (0.36)	1.13 (0.32)

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight.

Table 4. Mean (with standard deviation; SD) individual accuracy of breeding values predicted through pedigree-based (BLUP) and genomic-based analyses (ssGBLUP) for milk production, reproductive, and growth traits in Guzerá cattle, for all and for youngest animals

	All animals		**Youngest animals**	
Traits[a]	**BLUP**	**ssGBLUP**	**BLUP**	**ssGBLUP**
MY305	0.44 (0.23)	0.56 (0.14)	0.42 (0.10)	0.56 (0.06)
FY305	0.28 (0.19)	0.44 (0.10)	0.28 (0.12)	0.45 (0.05)
PY305	0.27 (0.20)	0.41 (0.12)	0.29 (0.13)	0.43 (0.06)
LY305	0.27 (0.20)	0.43 (0.12)	0.29 (0.13)	0.45 (0.06)
SY305	0.26 (0.20)	0.54 (0.08)	0.28 (0.13)	0.55 (0.04)
AFC	0.53 (0.13)	0.53 (0.14)	0.50 (0.07)	0.52 (0.05)
W210	0.52 (0.14)	0.53 (0.12)	0.44 (0.08)	0.46 (0.06)
W365	0.54 (0.16)	0.54 (0.15)	0.46 (0.08)	0.49 (0.06)
W450	0.54 (0.16)	0.55 (0.14)	0.46 (0.08)	0.49 (0.06)

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight.

3.4.3. Rank correlation and genetic trend

The rank correlations between EBVs and GEBVs considering all males and all females, for all the traits, were higher than 0.75 (Table 5). For AFC and growth traits, the rank correlations were higher than 0.94 considering all animals or top 5%, 25%, and 50%, both in males and females. For MY305, the rank correlations ranged from 0.80 to 0.91 for males, and from 0.82 to 0.92 for females.

For milk components, there is a possible reranking of animals when using genomic information in both sexes. The rank correlations for top 50% males ranged from 0.74 to 0.77, reaching 0.48 to 0.59 for top 5% males. For the 50% top females, the correlations ranged from 0.79 to 0.86, reaching 0.68 to 0.73 for the top 5%. Rank correlations increased as the proportion of animals increased, i.e., from the top 5% males (0.48 to 0.99) to all males (0.76 to 0.99), and from the top 5% females (0.68 to 0.99) to all females (0.79 to 0.99). Considering the youngest animals, the rank correlations ranged from 0.67 (W365) to 0.91 (MY305).

Table 5. Rank correlations between breeding values from pedigree-based and genomic-based models for all males, all females, the top 5%, 25%, and 50% animals within each sex, and youngest animals in Guzerá cattle

	Males				**Females**				**Youngest animals**
Traits[a]	**5%**	**25%**	**50%**	**All**	**5%**	**25%**	**50%**	**All**	
MY305	0.80	0.84	0.89	0.91	0.82	0.89	0.89	0.92	0.91
FY305	0.48	0.74	0.77	0.86	0.70	0.74	0.76	0.86	0.81
PY305	0.51	0.64	0.75	0.76	0.68	0.73	0.75	0.80	0.77
LY305	0.58	0.74	0.75	0.79	0.73	0.73	0.79	0.80	0.78
SY305	0.59	0.63	0.74	0.76	0.73	0.73	0.76	0.79	0.79
AFC	0.94	0.96	0.98	0.99	0.96	0.96	0.98	0.99	0.84
W210	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.76
W365	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.67
W450	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.69

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight.

The genetic trends obtained based on both BLUP and ssGBLUP are shown in Figures 1 and 2 and a similar pattern was observed between the approaches, for both sexes. The genetic gains before (≤2000) and after (>2000) the publication of the first sire summaries were similar in pedigree-based and genomic-based approaches (Table 6).

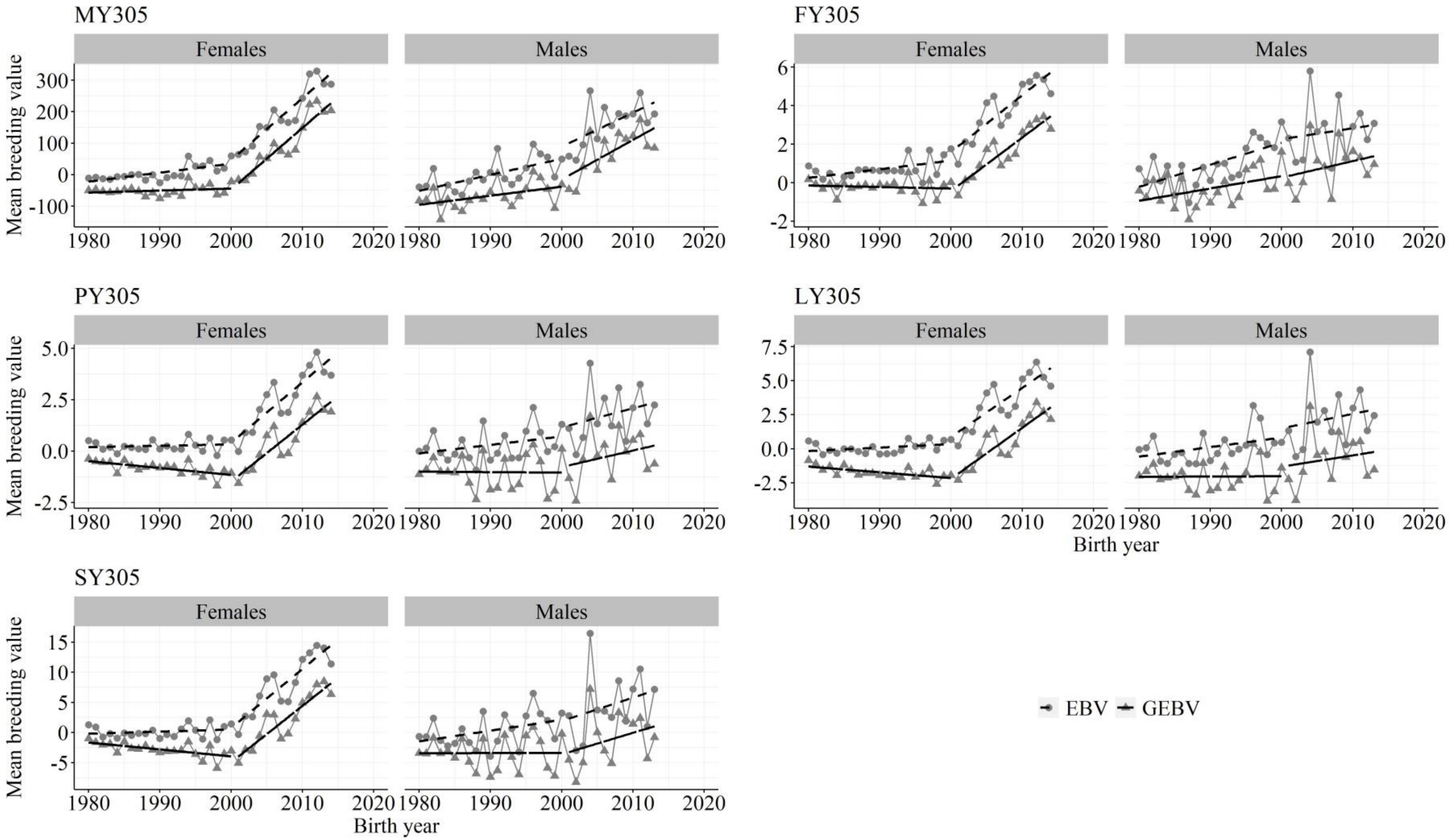


Figure 1. Genetic trends based on EBV and GEBV by birth year, considering males and females for 305-day first-lactation cumulative yields of milk (MY305), fat (FY305), protein (PY305), lactose (LY305), and total solids (SY305) for the Guzerá cattle population studied.

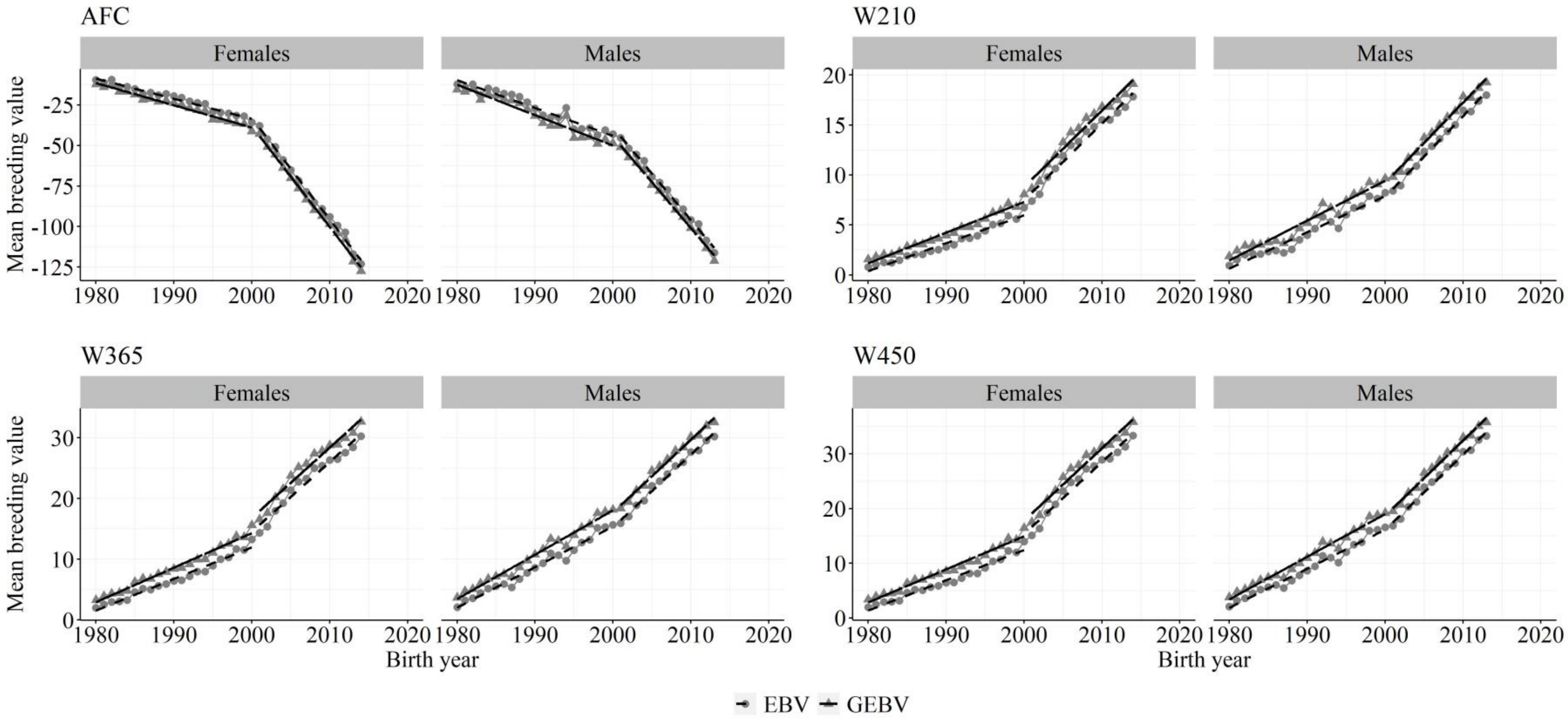


Figure 2. Genetic trends based on EBV and GEBV by birth year, considering males and females for age at first calving (AFC), and adjusted body weight at the ages of 210 days (W210), 365 (W365) and 450 (W450) days for the Guzerá cattle population studied.

Table 6. Genetic gain segmented based on the first sire summary (2000s) with respective standard errors and p-values for evaluated traits under pedigree-based and genomic-based models by sex in Guzerá cattle.

		Males		**Females**	
Traits[a]	**Segment**	**Pedigree-based**	**Genomic-based**	**Pedigree-based**	**Genomic-based**
MY305	≤ 2000	5.08 ± 1.39 (<0.002)*	2.88 ± 1.28 (<0.001)*	2.85 ± 0.67 (<0.001)*	0.64 ± 0.56 (<0.270)^{NS}
	> 2000	10.81 ± 4.27 (<0.03)*	12.57 ± 4.00 (<0.009)*	19.64 ± 2.26 (<0.001)*	19.56 ± 2.13 (<0.001)*
FY305	≤ 2000	0.11 ± 0.03 (<0.001)*	0.06 ± 0.03 (<0.045)*	0.05 ± 0.02 (<0.020)*	-0.01 ± 0.01 (<0.580)^{NS}
	> 2000	0.06 ± 0.11 (<0.622)^{NS}	0.09 ± 0.09 (<0.339)^{NS}	0.30 ± 0.05 (<0.001)*	0.28 ± 0.04 (<0.001)*
PY305	≤ 2000	0.04 ± 0.03 (<0.140)^{NS}	-0.002 ± 0.03 (<0.941)^{NS}	0.01 ± 0.001 (<0.486)*	-0.03 ± 0.01 (<0.002)*
	> 2000	0.09 ± 0.09 (<0.338)^{NS}	0.08 ± 0.09 (<0.367)^{NS}	0.29 ± 0.05 (<0.001)*	0.28 ± 0.04 (<0.001)*
LY305	≤ 2000	0.07 ± 0.04 (<0.074)^{NS}	0.002 ± 0.04 (<0.958)^{NS}	0.03 ± 0.01 (<0.088)^{NS}	-0.04 ± 0.01 (<0.067)^{NS}
	> 2000	0.11 ± 0.16 (<0.513)^{NS}	0.08 ± 0.14 (<0.557)^{NS}	0.36 ± 0.07 (<0.001)*	0.37 ± 0.06 (<0.001)*
SY305	≤ 2000	0.18 ± 0.09 (<0.070)^{NS}	0.001 ± 0.10 (<0.989)^{NS}	0.03 ± 0.04 (<0.351)^{NS}	-0.12 ± 0.03 (<0.002)*
	> 2000	0.39 ± 0.39 (<0.346)^{NS}	0.36 ± 0.32 (<0.286)^{NS}	0.98 ± 0.15 (<0.001)*	0.94 ± 0.13 (<0.001)*
AFC	≤ 2000	-1.71 ± 0.11 (<0.001)*	-1.86 ± 0.11 (<0.001)*	-1.24 ± 0.05 (<0.001)*	-1.37 ± 0.05 (<0.001)*
	> 2000	-5.72 ± 0.14 (<0.001)*	-5.65 ± 0.14 (<0.001)*	-6.27 ± 0.14 (<0.001)*	-6.22 ± 0.15 (<0.001)*
W210	≤ 2000	0.36 ± 0.02 (<0.001)*	0.40 ± 0.02 (<0.001)*	0.28 ± 0.01 (<0.001)*	0.31 ± 0.01 (<0.001)*
	> 2000	0.81 ± 0.02 (<0.001)*	0.80 ± 0.03 (<0.001)*	0.77 ± 0.04 (<0.001)*	0.77 ± 0.04 (<0.001)*
W365	≤ 2000	0.67 ± 0.02 (<0.001)*	0.73 ± 0.02 (<0.001)*	0.52 ± 0.02 (<0.001)*	0.57 ± 0.02 (<0.001)*
	> 2000	1.20 ± 0.04 (<0.001)*	1.20 ± 0.04 (<0.001)*	1.15 ± 0.07 (<0.001)*	1.16 ± 0.07 (<0.001)*
W450	≤ 2000	0.71 ± 0.03 (<0.001)*	0.78 ± 0.03 (<0.001)*	0.55 ± 0.02 (<0.001)*	0.60 ± 0.02 (<0.001)*

	> 2000	1.38 ± 0.04 (<0.001)*	1.37 ± 0.05 (<0.001)*	1.32 ± 0.07 (<0.001)*	1.32 ± 0.08 (<0.001)*

---

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight.

* Significant to 0.05. [NS] Not significant to 0.05.

3.5. Discussion

The results obtained in this study illustrate the feasibility of using the ssGBLUP approach for the genetic evaluations of Guzerá cattle, despite the still small number of genotyped animals available. Comparisons between BLUP and ssGBLUP were based on predictive abilities, bias, individual accuracies, rank correlations, and genetic trends. The most important economic traits for the improvement of Guzerá as dual-purpose breed were evaluated in this study. Therefore, our study represents an important step towards the implementation of genomic prediction for Guzerá cattle in Brazil.

Although considered a dual-purpose breed in Brazil, the genetic evaluation of Guzerá cattle is carried out independently for milk and beef production. The existence of "double proven" sires, i.e., sires genetically evaluated for dairy and beef traits (Bruneli et al., 2020), allows the farmers to choose concomitantly for meat and milk production improvement. Moreover, there is no genetic antagonism among dairy, growth and reproductive traits in Guzerá cattle as showed in previous studies (Brito et al., 2020; Carrara et al., 2022) and also on this paper. Thus, it is possible to select for one purpose without loss in the others.

The inclusion of genomic information for variance components estimation and genetic parameters for the studied population, did not cause large differences in comparison to estimates obtained through pedigree-based model (Carrara et al., 2022). Considering the standard errors, the heritability estimates do not differ between REML and ssGREML, however, comparing only the absolute values, we found a small increase in the heritabilities estimates for the dairy traits when using ssGREML.

The mean difference between relationship based on pedigree and on genomic in studied population was low, which may explain the similarity between the estimates of heritability and genetic correlations obtained from pedigree-based and genomic-based analyses. Misztal et al. (2013) point out that the main reasons for the occurrence of differences between pedigree-based and genomic-based relationships are incomplete pedigree and pedigree mistakes. More than 90% of the animals included in the present study have both sire and dam known, and probably few mistakes in pedigree of the genotyped animals, reflecting in small differences between relationship based on pedigree and on genomic.

Veerkamp et al. (2011) and Kluska et al. (2018) reported that the use of SNP-based relationships improved the precision of genetic parameters estimates (in terms of decrease of standard error). In our study, we did not observe differences between the methodologies, probably because of few genotyped animals were used in the analysis and we also used H matrix which combines pedigree and genomic information.

Regarding the predictive ability, the highest difference between BLUP and ssGBLUP was observed for SY305, wherein an increase of 0.07 was observed when using ssGBLUP approach. It is important to mention that for SY305 higher proportion of the phenotyped animals was also genotyped (40%) than in the other evaluated traits. Thus, it may be one possible reason for the predictive ability increase. For remaining traits, similar predictive abilities were observed between the approaches. Most probably the small number of genotyped animals affected the predictive abilities in our study. Silva et al. (2016) reported that ssGBLUP provided higher predictive abilities than traditional BLUP for residual feed intake and feed conversion ratio in Nelore cattle, with difference of 96% and 3.5% respectively. However, these authors used a small database with a higher proportion of genotyped animals, containing 9,551 animals in the pedigree, of which 896 animals were phenotyped and 788 animals were genotyped. We have not been able to find any studies showing predictive ability of traditional BLUP and ssGBLUP for dairy traits in *Bos indicus* cattle. In *Bos taurus* cattle, several studies reported superior predictive ability of ssGBLUP for dairy cattle (Matilainen et al., 2018), beef cattle (Lourenco et al., 2015; Mehrban et al., 2019), and dual-purpose cattle (Cesarani et al., 2021).

According to Goddard and Hayes (2009), the size of the training population is one of the main factors that affect the accuracy of genomic prediction. Song et al. (2019) using simulated pig data evaluated the effect of reference population sizes in the predictive abilities, and the authors reported that when a smaller number of genotyped animals was included in the analysis, the difference between traditional BLUP and ssGBLUP methods was minimal. In our study, the number of genotyped animals is still low because of genomic prediction implementation in Guzerá cattle in Brazil is at an early stage, and the training population consisted mainly on females.

Only females were used in the reference population, which have fewer offspring than the sires. In this way, increase in predictive abilities are expected as the size of the Guzerá cattle reference population continues to increase, especially if bulls are included. In addition, as pointed for genetic parameters, the similarity between mean of relationship based on pedigree and genomic relationship may also account for the lack of differences between the predictive abilities of BLUP and ssGBLUP.

Differences between pedigree-based and genomic-based models were more evident when individual accuracies were analyzed. In this case, an increase of up to 116% was observed when using ssGBLUP. The milk production traits showed higher gains in accuracy than the other traits evaluated, possibly because of the higher proportion of genotyped animals with phenotypes for these traits.

The gain in accuracy was higher for young animals for most traits, as observed for other authors (Lopes et al., 2018). Even with lower proportion of genotyped animals, the AFC and growth traits presented a small gain in accuracy using ssGBLUP, when considering only younger animals. This result further confirms the benefits of including genomic information into the Guzerá genetic evaluation, especially for young animals. Increases in breeding values accuracy resulting from the adoption of genomic evaluation would lead to greater genetic gains for the evaluated traits. Cesarani et al. (2021) evaluated 131,308 milkability phenotypes from dual-purpose Italian Simmental cattle, including 9,526 genotyped animals (2,455 animals with phenotype and genotype) under traditional BLUP and ssGBLUP, the authors reported 37% higher individual accuracy in ssGBLUP. According to the authors, the main reason for the greater accuracies with ssGBLUP than with traditional BLUP are smaller prediction error due to larger amounts of information and more accurate relationships between individuals.

There are some possible reasons for the superiority of ssGBLUP in dairy traits compared to AFC and growth traits in our study. Firstly, the dairy traits showed higher heritabilities (Table 2) and the accuracy is dependent on the additive variance and, consequently, on the heritability of the trait (Hayes et al., 2009). There are more animals with genotype and phenotype for milk production traits than for reproductive and growth traits. The milk production traits presented from 19 to 42% of the animals with both phenotype and genotype. The AFC showed just over 1% of animals with both phenotype and genotype, and the growth traits (W210, W365, and W450) showed less than 1% of animals with both phenotype and genotype. In addition, the genotyped animals are predominantly from dairy and/or dual-purpose herds. The AFC and growth traits were mostly measured in herds selected for meat production, and although there is a genetic connection among the Guzerá herds (Peixoto et al., 2021), the genotyped animals may not have caused substantial differences in the genetic evaluations by BLUP and ssGBLUP.

The bias was higher than 1.00 for almost all traits, with the exception of AFC, in both BLUP and ssGBLUP approaches. Assuming b_1 values within 1.00 ± 0.15 are acceptable (Tsuruta et al., 2011) the traits MY305, FY305, LY305, SY305, and W210, were outside of the acceptable range. The bias in genomic-based analyses may be affected by compatibility between a pedigree-based relationship matrix (**A**) and a genomic relationship matrix (**G**). Moreover, the way the bias is calculated can influence its interpretation. For example, Costa et al. (2019) evaluated prediction bias for female reproductive traits in Nellore cattle and reported

b_1 lower than 1.00 when GEBVs were regressed on adjusted phenotypes. In turn, they reported b_1 greater than 1.00 when the same GEBVs were regressed on EBVs. Therefore, the bias in obtained in our study may have originated from the data and modeling themselves, since bias was also observed in pedigree-based model. This argument is also supported by the fact that the b_1 from pedigree-based and from genomic-based models was similar.

No relevant changes were observed for both sexes in the rank of animals for MY305, AFC, W210, W365, and W450 and the rank correlations were higher than 0.80, indicating that the use of BLUP or ssGBLUP did not influence the animals' ranking. Therefore, similar selection decisions for these traits would be taken based on the results of both approaches. On the other hand, considering youngest animals, the rank correlations were lower than when considering all animals, reaching 0.67. Even for growth traits, which have a smaller number of genotyped animals, the correlations were lower than 0.76, indicating reranking of youngest animals.

In breeding programs, it is expected that younger generations will excel if selection is really taking place. However, young animals have less information, leading to less accurate breeding values in pedigree-based evaluations. The genomic-based evaluation provides more accurate breeding values than pedigree-based analysis, and may also cause differences in young animals breeding values because genomic relationships create stronger ties than pedigree relationships (Misztal et al., 2020). Consequently, the chance of reranking of younger animals between pedigree-based and genomic-based models increases. For milk components (i.e., FY305, PY305, LY305, and SY305) the rank correlations were lower than the other evaluated traits, mainly under lower proportion of selected animals, suggesting that reranking may occur using genomic-based prediction. The dairy traits had fewer measurements and may present greater differences between pedigree-based and genomic-based analyses.

Genetic trends for dairy traits showed slightly lower average GEBVs than average EBVs, but the pattern of genetic values across birth years was the same for both methods. For the other traits, genetic trends were the same between methods. There is genetic improvement for all traits and it is intensified in the years following 2000 for all traits, which is the year wherein the first sire summaries was published (Bruneli et al., 2020; ANCP, 2022). For this reason, we used segmented regression in 2000 to actually show genetic progress after selection began. A negative trend for AFC was observed, as expected. It is noteworthy that AFC was recently included in the summary of PNMGuL, and its negative genetic may be a correlated response, since AFC shows a favorable negative genetic correlation with 305-day cumulative milk yield (Brito et al., 2020; Carrara et al., 2022) which is the main target trait for selection.

Other studies for dairy traits and female reproductive traits report similar pattern for EBVs and GEBVs by birth year (Masuda et al., 2018; Matilainen et al., 2018). Genomic selection is not yet applied in the studied population, which may have caused the similarity of genetic trends between pedigree-based and genomic-based methods.

In the present study, the feasibility of using ssGBLUP approach for productive and reproductive traits of Guzerá cattle were evaluated considering a small number of genotyped animals and a small database. The ssGBLUP provided higher individual accuracies (individual EBV or GEBV accuracy) for all traits evaluated, mainly for dairy traits. However, this gain was not as large for AFC and growth traits. Predictive abilities were similar between approaches and bias also was similar and was observed in both approaches, suggesting it did not originate from genomic-based evaluations. Genetic trends were not affected by the inclusion of genomic information and showed the same behavior for all traits.

3.6. Conclusion

This study, even based on a small number of genotyped animals and a small database for some traits, provided interesting insights about the use and future application of single-step genomic BLUP in the Brazilian Guzerá cattle. The results suggest that the single-step genomic BLUP is feasible and may be applied to national genetic evaluations for reproductive, dairy and growth traits in Guzerá cattle to increase the accuracy of breeding values. However, the number of genotyped animals has not yet been sufficient to influence genetic evaluation in terms of predictive ability and bias.

3.7. Acknowledgments

The authors acknowledge the Centro Brasileiro de Melhoramento Genético do Guzerá (CBMG^2) and Associação Brasileira de Criadores de Zebu (ABCZ) for providing datasets. The authors also acknowledge the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and INCT (Instituto Nacional de Ciência e Tecnologia) for financial support.

3.8. References

ABCZ 2020. Zebu breeds. Brazilian Association of Zebu Breeders (ABCZ). Retrieved on 2 November 2021, from *https://www.abcz.org.br/a-abcz/racas-Zebuinas*.

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S and Lawlor TJ 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score1. Journal of Dairy Science 93, 743–752.

ANCP 2022. Guzerá program. Associação Nacional de Criadores e Pesquisadores. Retrieved on 23 June 2022, from *https://www.ancp.org.br/programas/guzera/*.

Brito LC, Peixoto MGCD, Carrara ER, Fonseca e Silva F, Ventura HT, Bruneli FAT and Lopes PS 2020. Genetic parameters for milk, growth, and reproductive traits in Guzerá cattle under tropical conditions. Tropical Animal Health and Production 52, 2251–2257.

Bruneli FAT, Peixoto MGCD, Júnior MLS, Pereira RJ, Arbex WA, Penna VM, Zadra LEF, Verneque R da S, Lôbo RB and Carvalho MRS 2020. National Breeding Program of Guzerá

Cattle for Milk: Progeny Testing, National Zootechnical Archive and MOET Nucleus Results. Brazilian Agricultural Research Corporation Documentos, 251, Juiz de Fora: Embrapa Dairy Cattle.

Carrara ER, Peixoto MGCD, Veroneze R, Silva FF e, Ramos PVB, Bruneli FAT, Zadra LEF, Ventura HT, Josahkian LA and Lopes PS 2022. Genetic study of quantitative traits supports the use of Guzerá as dual-purpose cattle. Animal Bioscience.

Cesarani A, Garcia A, Hidalgo J, Degano L, Vicario D, Macciotta NPP and Lourenco D 2021. Genomic information allows for more accurate breeding values for milkability in dual-purpose Italian Simmental cattle. Journal of Dairy Science 104, 5719–5727.

Cesarani A, Pocrnic I, Macciotta NPP, Fragomeni BO, Misztal I and Lourenco DAL 2018. Bias in heritability estimates from genomic restricted maximum likelihood methods under different genotyping strategies. Journal of Animal Breeding and Genetics 136, 40–50.

Christensen OF and Lund MS 2010. Genomic prediction when some animals are not genotyped. Genetics Selection Evolution 42, 1–8.

Costa RB, Irano N, Diaz IDPS, Takada L, Hermisdorff I da C, Carvalheiro R, Baldi F, de Oliveira HN, Tonhati H and de Albuquerque LG 2019. Prediction of genomic breeding values for reproductive traits in Nellore heifers. Theriogenology 125, 12–17.

Fonseca PA de S, Santos FC, Rosse IC, Ventura RV, Brunelli FÂT, Penna VM, da Silva Verneque R, Machado MA, da Silva MVGB, Carvalho MRS and Peixoto MGCD 2016. Retelling the recent evolution of genetic diversity for Guzerá: Inferences from LD decay, runs of homozygosity and Ne over the generations. Livestock Science 193, 110–117.

Forni S, Aguilar I and Misztal I 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. Genetics Selection Evolution 43, 1.

Goddard ME and Hayes BJ 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics 10, 381–391.

Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. Journal of Dairy Science 92, 433–443.

Henderson CR 1976. A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. Biometrics 32, 69.

Kluska S, Olivieri BF, Bonamy M, Chiaia HLJ, Feitosa FLB, Berton MP, Peripolli E, Lemos MVA, Tonussi RL, Lôbo RB, Magnabosco C de U, Di Croce F, Osterstock J, Pereira ASC, Munari DP, Bezerra LA, Lopes FB and Baldi F 2018. Estimates of genetic parameters for growth, reproductive, and carcass traits in Nelore cattle using the single step genomic BLUP procedure. Livestock Science 216, 203–209.

Legarra A, Aguilar I and Colleau JJ 2020. Short communication: Methods to compute genomic inbreeding for ungenotyped individuals. Journal of Dairy Science 103, 3363–3367.

Legarra A, Aguilar I and Misztal I 2009. A relationship matrix including full pedigree and genomic information. Journal of Dairy Science 92, 4656–4663.

Lopes FB, Wu X-L, Li H, Xu J, Perkins T, Genho J, Ferretti R, Tait RG, Bauck S and Rosa GJM 2018. Improving accuracy of genomic prediction in Brangus cattle by adding animals with imputed low-density SNP genotypes. Journal of Animal Breeding and Genetics 135, 14–27.

Lourenco DAL, Tsuruta S, Fragomeni BO, Masuda Y, Aguilar I, Legarra A, Bertrand JK, Amen TS, Wang L, Moser DW and Misztal I 2015. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus1. Journal of Animal Science 93, 2653–2662.

Masuda Y, VanRaden PM, Misztal I and Lawlor TJ 2018. Differing genetic trend estimates from traditional and genomic evaluations of genotyped animals as evidence of preselection bias in US Holsteins. Journal of Dairy Science 101, 5194–5206.

Matilainen K, Strandén I, Aamand GP and Mäntysaari EA 2018. Single step genomic evaluation for female fertility in Nordic Red dairy cattle. Journal of Animal Breeding and Genetics 135, 337–348.

Mehrban H, Lee DH, Naserkheil M, Moradi MH and Ibáñez-Escriche N 2019. Comparison of conventional BLUP and single-step genomic BLUP evaluations for yearling weight and carcass traits in Hanwoo beef cattle using single-trait and multi-trait models. PLOS ONE 14, e0223352.

Misztal I, Aggrey SE and Muir WM 2013. Experiences with a single-step genome evaluation. Poultry Science 92, 2530–2534.

Misztal I, Legarra A and Aguilar I 2009. Computing procedures for genetic evaluation including phenotypic , full pedigree , and genomic information. Journal of Dairy Science 92, 4648–4655.

Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A and Vitezica Z 2014. Manual for BLUPF90 family of programs. Athens, USA. Retrieved on 14 February 2022, from *http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf*.

Misztal I, Tsuruta S, Pocrnic I and Lourenco D 2020. Core-dependent changes in genomic predictions using the Algorithm for Proven and Young in single-step genomic best linear unbiased prediction. Journal of Animal Science 98, 1–8.

Mrode RA and Thompson R 2014. Linear models for the prediction of animal breeding values. CAB International, Wallingford, Oxon, UK.

Oliveira HR, Lourenco DAL, Masuda Y, Misztal I, Tsuruta S, Jamrozik J, Brito LF, Silva FF and Schenkel FS 2019. Application of single-step genomic evaluation using multiple-trait random regression test-day models in dairy cattle. Journal of Dairy Science 102, 2365–2377.

Paiva JT, Peixoto MGCD, Bruneli FAT, Alvarenga AB, Oliveira HR, Silva AA, Silva DA, Veroneze R, Silva FF and Lopes PS 2020. Genetic parameters, genome-wide association and gene networks for milk and reproductive traits in Guzerá cattle. Livestock Science 242, 104273.

Peixoto MGCD, Carvalho MRS, Egito AA, Steinberg RS, Bruneli FÂT, Machado MA, Santos FC, Rosse IC and Fonseca PAS 2021. Genetic Diversity and Population Genetic Structure of a Guzerá (*Bos indicus*) Meta-Population. Animals 11, 1125.

Peixoto MGCD, Poggian CF, Verneque RS, Egito AA, Carvalho MRS, Penna VM, Bergmann JAG, Viccini LF and Machado MA 2010. Genetic basis and inbreeding in the Brazilian Guzerat (*Bos indicus*) subpopulation selected for milk production. Livestock Science 131, 168–174.

Quaas RL 1976. Computing the Diagonal Elements and Inverse of a Large Numerator Relationship Matrix. Biometrics 32, 949.

Rosse IC, Assis JG, Oliveira FS, Leite LR, Araujo F, Zerlotini A, Volpini A, Dominitini AJ, Lopes BC, Arbex WA, Machado MA, Peixoto MGCD, Verneque RS, Martins MF, Coimbra RS, Silva MVGB, Oliveira G and Carvalho MRS 2017. Whole genome sequencing of Guzerá cattle reveals genetic variants in candidate genes for production, disease resistance, and heat tolerance. Mammalian Genome 28, 66–80.

dos Santos FC, Peixoto MGCD, Fonseca PA de S, Pires M de FÁ, Ventura RV, Rosse I da C, Bruneli FAT, Machado MA and Carvalho MRS 2017. Identification of Candidate Genes for Reactivity in Guzerat (*Bos indicus*) Cattle: A Genome-Wide Association Study. PLOS ONE 12, e0169163.

Sargolzaei M, Chesnais JP and Schenkel FS 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15, 478.

Silva RMO, Fragomeni BO, Lourenco DAL, Magalhães AFB, Irano N and Carvalheiro R 2016. Accuracies of genomic prediction of feed efficiency traits using different prediction and validation methods in an experimental Nelore cattle population 1. 3613–3623.

Silva AA, Silva DA, Silva FF, Costa CN, Lopes PS, Caetano AR, Thompson G and Carvalheira J 2019. Autoregressive single-step test-day model for genomic evaluations of Portuguese Holstein cattle. Journal of Dairy Science 102, 6330–6339.

Song H, Zhang J, Zhang Q and Ding X 2019. Using Different Single-Step Strategies to Improve the Efficiency of Genomic Prediction on Body Measurement Traits in Pig. Frontiers in Genetics 9, 1–10.

Stafuzza NB, Zerlotini A, Lobo FP, Yamagishi MEB, Chud TCS, Caetano AR, Munari DP, Garrick DJ, Machado MA, Martins MF, Carvalho MR, Cole JB and Barbosa da Silva MVG 2017. Single nucleotide variants and InDels identified from whole-genome re-sequencing of Guzerat, Gyr, Girolando and Holstein cattle breeds. PLOS ONE 12, e0173954.

Tonussi RL, Londoño-Gil M, de Oliveira Silva RM, Magalhães AFB, Amorim ST, Kluska S, Espigolan R, Peripolli E, Pereira ASC, Lôbo RB, Aguilar I, Lourenço DAL and Baldi F 2021. Accuracy of genomic breeding values and predictive ability for postweaning liveweight and age at first calving in a Nellore cattle population with missing sire information. Tropical Animal Health and Production 53, 432.

Tsuruta S, Misztal I, Aguilar I and Lawlor TJ 2011. Multiple-trait genomic evaluation of linear type traits using genomic and phenotypic data in US Holsteins. Journal of Dairy Science 94, 4198–4204.

VanRaden PM 2008. Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science 91, 4414–4423.

Veerkamp RF, Mulder HA, Thompson R and Calus MPL 2011. Genomic and pedigree-based genetic parameters for scarcely recorded traits when some animals are genotyped. Journal of Dairy Science 94, 4189–4197.

Van Vleck LD 1993. Variance of prediction error with mixed model equations when relationships are ignored. Theoretical and Applied Genetics 85, 545–549.

3.9. Supplementary material

3.9.1. Supplementary Tables

Supplementary Table 1. Estimates of variance components and their standard errors (within parentheses) for 305-day milk yield, 305-day milk components, growth traits, and reproductive traits studied in Guzerá cattle obtained by genomic-based single-trait analyses

Traits[a]	$\boldsymbol{\sigma^2_a}$	$\boldsymbol{\sigma^2_m}$	$\boldsymbol{\sigma^2_{mp}}$	$\boldsymbol{\sigma^2_e}$
MY305	179,720.00 (20,505.00)	-	-	325,780.00 (15,225.00)
FY305	226.05 (52.80)	-	-	697.26 (45.49)
PY305	156.57 (32.30)	-	-	335.12 (25.58)
LY305	278.25 (56.97)	-	-	551.62 (44.62)
SY305	2,621.30 (537.30)	-	-	4,136.50 (408.47)
AFC	5,882.20 (264.39)	-	-	23,620.00 (222.36)
W210	110.15 (5.69)	20.59 (2.63)	69.90 (3.08)	492.91 (4.18)
W365	307.76 (12.50)	-	-	949.95 (9.63)
W450	365.12 (15.03)	-	-	1,191.10 (11.73)

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight; σ^2_a = additive genetic variance; σ^2_m = maternal genetic variance; σ^2_{mp} = maternal permanent environmental variance; σ^2_e = residual variance.

CHAPTER 4

Inferences about inbreeding using runs of homozygosity of Guzerá dual-purpose cattle

Carrara, E.R.[1], Peixoto, M.G.C.D.[2], Veroneze, R.[1], Bruneli, F.A.T.[2], Zadra, L.E.F.[3], Ventura, H.T.[4], Josahkian, L.A.[4], and Lopes, P.S.[1]

[1] Department of Animal Sciences, Universidade Federal de Viçosa, Viçosa, MG, Brazil

[2] Embrapa Dairy Cattle, Juiz de Fora, MG, Brazil.

[3] Brazilian Center for the Genetic Improvement of Guzerá, Belo Horizonte, MG, Brazil.

[4] Brazilian Association of Zebu Cattle, Uberaba, MG, Brazil.

4.1. Abstract

Inbreeding coefficients have traditionally been determined using pedigree information. With the availability of genetic marker information, inbreeding coefficients can be more accurately determined by runs of homozygosity (ROH). Our aim was to obtain inbreeding coefficients for a population of Guzerá cattle by runs of homozygosity and to compare these coefficients with pedigree-based inbreeding coefficients. In addition, we aimed to evaluate the presence of inbreeding depression for reproductive, dairy and growth traits in the population. Genotype data of 44,284 SNPs were available for 1,733 Guzerá cattle that were part of a much larger data set of pedigree entries (433,823 animals). Inbreeding was measured using pedigree data F_{PED} and runs of homozygosity F_{ROH}. In addition, inbreeding coefficients were calculated based on the genomic relationship matrix and based on the relation between the observed and expected number of homozygous genotypes. Inbreeding depression was assessed by phenotypes: 305-day first-lactation cumulative yields of milk, fat, protein, lactose, and total solids; age at first calving in days; and adjusted body weight at the ages of 210 days, 365 and 450 days. ROH were identified in all animals, totaling 16,379 homozygous segments, with an average number of 9.45 ± 4.21 segments and a mean length of 6.76 ± 7.12 Mb. BTA6 had the highest number of ROH and the highest numbers of animals with ROH (1,563 and 1,063, respectively). The average F_{PED} was 0.007 ± 0.022 and the average F_{ROH} was 0.023 ± 0.0003, low values compared to other cattle breeds. Low correlation was observed between F_{PED} and F_{ROH} (0.31). F_{PED} also showed low correlations with alternatives genomic-based methods (0.31 to 0.36). F_{ROH} presented low-to-moderate correlations with alternative genomic-based methods (0.43 to 0.78). No inbreeding depression was observed in the reproductive, dairy and growth traits evaluated. The Guzerá cattle population studied has maintained a lower level of inbreeding than other cattle breeds, and still no significant effects of inbreeding level on dairy and beef cattle traits.

Keywords: Beef cattle; Dairy cattle; Inbreeding depression; Zebu

4.2. Introduction

The mating of individuals related by descent leads to inbreeding. Two individuals that share a common ancestor can have the essential consequence of carrying replicates of the same allele came from that ancestor, increasing the chance to pass those replicates to their offspring if they mate. Thus, inbred animals can carry two alleles at a locus that are replicates of the same allele from an earlier generation (Falconer and Mackay, 1996). High inbreeding rates can increase the occurrence of homozygosis for deleterious recessive alleles and reduce genetic variability in the population (Charlesworth and Willis, 2009). It is highlighted that genetic variability is key to the success of any animal breeding program, and in this context, controlling inbreeding levels is important to avoid unfavorable effects on traits as a result of inbreeding depression.

Defined as a reduction in the mean of a population for a quantitative trait, inbreeding depression has long been known in the plant and animal populations (Charlesworth, 2006; Leroy, 2014). For decades, inbreeding in cattle has been and continues to be evaluated using pedigree information (Miglior et al., 1992; Peixoto et al., 2010; Pereira et al., 2016; Carrara et al., 2020), based on statistical probability expectations that two alleles are identical by descent (IBD).

With the availability of genome sequencing technologies, there are further opportunities to understand the genetic basis of inbreeding depression (Ferenčaković et al., 2017). For example, genetic markers of type single-nucleotide polymorphism (SNP) can be used to calculate a genomic relationship matrix (G) among individuals, where the expected value of a diagonal element of G is 1 + F, where F is equal to the inbreeding coefficient of the animal (VanRaden, 2008; Hayes et al., 2009; Yang et al., 2010). This genomic estimate measures realized inbreeding, which can vary between animals with the same pedigree, i.e., full siblings (Pryce et al., 2014). In addition to inbreeding coefficients based on the G matrix, the relation of the observed and expected number of homozygous genotypes is another example of genomic measure that can be used to estimate the degree of inbreeding across the genome (Purcell et al., 2007).

One limitation of the measures cited above is that they do not distinguish between IBD and identity by state (IBS). In this sense, an alternative are the runs of homozygosity (ROH), defined as continuous homozygous segments of DNA sequence (Gibson et al., 2006), which have been widely used in studies of inbreeding and inbreeding depression in cattle (Fonseca et al., 2016; Ferenčaković et al., 2017; Peripolli et al., 2018; Mulim et al., 2022). When parents have a common ancestor and pass shared chromosomal segments on to their progeny, autozygosity occurs. Given high correlation of ROHs with autozygosity (0.70) (McQuillan et al., 2008), they have been used to quantify autozygosity with high accuracy (Keller et al., 2011; Marras et al., 2015), i.e., ROH can be used to identify the location of specific regions of the genome that are IBD (Peripolli et al., 2017). Bjelland et al. (2013) evaluated inbreeding in Holstein cattle using ROH and other genomic inbreeding coefficient estimators and concluded that only ROH can distinguish between IBD and IBS markers.

The Brazilian population of Guzerá cattle evolved from breeds originated from a small number of animals imported from India at the end of the 19th century and spread to Brazil, where they were able to adapt to different environmental conditions, especially harsh ones. The breed remained predominant in cattle breeding until the 1930s, but the population declined until the middle of the last century due to its extensive use in crossbreeding programs (Fonseca et al., 2016; Santana et al., 2020; Peixoto et al., 2021). Currently, this breed represents the third largest purebred indicine breed and one of the most productive among Zebu breeds in Brazil.

The Brazilian population of Guzerá animals has genetic variability that allows them to be used in both beef and dairy herds, and also in both purposes (Peixoto et al., 2021), since animals are naturally dual-purpose and there is no genetic antagonism between dairy and beef traits (Brito et al., 2020; Carrara et al., 2022). Thus, they represent an important genetic resource for Brazilian livestock production and controlling the level of inbreeding in herds is therefore crucial. Studies evaluating inbreeding in the Guzerá animal population have previously been conducted using pedigree-based inference (Peixoto et al., 2010; Pereira et al., 2016) and more recently using SNPs markers (Fonseca et al., 2016). However, there are no reports of studies in the breed evaluating inbreeding depression using genomic information.

Our aims with this study were to use pedigree and phenotypic data from a Guzerá cattle population, which a subpopulation is genotyped for SNP markers, to: i) Assess autozygosity using runs of homozygosity; ii) estimate pedigree-based and genomic-based inbreeding coefficients for population; iii) estimate the inbreeding depression for dairy and beef traits using inbreeding coefficients based on runs of homozygosity and based on pedigree data; iii) compare these estimates.

4.3. Material and methods

4.3.1. Data source and statement of animal rights

The data used in this study were provided by Zebu Breeds Genealogical Registry Service (SRGRZ) of the Brazilian Association of Zebu Cattle (ABCZ), and by the National Program for the Improvement of Guzerá for Dairy Purpose (PNMGuL), coordinated jointly by Embrapa Dairy Cattle and the Brazilian Center for the Genetic Improvement of Guzerá (CBMG[2]). Thus, approval of the animal care and use committee was not needed because this research used existing datasets historically collected by the animal breeding program.

4.3.2. Genotypes

A total of 973 cows were genotyped with the Illumina BovineSNP50 v2 BeadChip (Illumina, Inc. San Diego, California), containing 54,609 SNPs; 50 bulls with the Illumina BovineHD BeadChip (Illumina, Inc. San Diego, California), containing 777,962 SNPs; and 820 cows and five bulls with the GeneSeek Genomic Profiler *indicus* chip (GGP *indicus*; Neogen Corp., Lansing, MI), containing 54,791 SNPs, totaling 1,848 animals genotyped.

The common SNPs between the BovineSNP50 and the GGP *indicus* chips were 9,636, and between BovineHD and the GGP *indicus* chips were 48,903. The missing genotypes were imputed to GGP *indicus* panel using the FImpute version 3 (Sargolzaei et al., 2014). The GGP *indicus* is a panel built specifically for *Bos indicus* cattle, with the selection of SNPs with higher MAF in Zebu cattle populations and designed to optimize equidistant spacing of markers (Aliloo et al., 2018). For this reason, it was prioritized as the reference panel in an imputation analysis to use all genotypic information available. The imputation accuracy, taken as Spearman correlation coefficient between imputed and true SNPs, was higher than 0.94, in accordance with the results presented by Carrara et al. in previous chapter (unpublished; > 0.91) using the same population and fewer animals genotyped.

Prior imputation, only autosomal SNPs with call rate higher than 0.95 have been kept. In addition, samples also were edited for call rate (<0.95), parent-progeny pairs were tested for Mendelian conflicts, and samples with inconsistent identification were excluded. After imputation, a total of 1,733 animals and 50,641 SNPs were retained for the analyses. These genotype data were used only for ROH inferences. For the remaining analyses, the minor frequency allele (MAF < 0.02) and Hardy-Weinberg equilibrium (estimated as the difference between expected and observed frequency of heterozygous > 0.15) parameters were also used to filter out SNPs, remaining 1,733 animals (1,694 cows and 39 bulls) and 44,284 SNPs.

4.3.3. Runs of homozygosity

The PLINK version 1.07 software (Purcell et al., 2007) was used for the ROH identification based on the following criteria: (i) up to one heterozygous genotype and a maximum of five SNPs with missing genotypes were allowed in a ROH; (ii) the proportion of homozygous overlapping windows was 0.05; (iii) the maximum gap between consecutive homozygous SNPs was 500 kb; (iv) the minimum length of a ROH was set to 500 kb; (v) the minimum number of consecutive SNPs included in a ROH was 30; (vi) a density of one SNP per 50 kb; and (vii) a sliding window of 50 SNPs across the genome. There is no consensus on the definition of the parameters for measuring ROH. Therefore, in this study, ROH parameters were defined according to recent studies in cattle, mainly Zebu (Fonseca et al., 2016; Peripolli et al., 2018; Meyermans et al., 2020; Mulim et al., 2022). ROH were also classified into five length classes: <2, 2–4, 4–8, 8–16, and >16 Mb, identified as $ROH_{<2Mb}$, $ROH_{2–4Mb}$, $ROH_{4–8Mb}$, $ROH_{8–16Mb}$, and $ROH_{>16Mb}$, respectively.

Inbreeding coefficients based on ROH (F_{ROH}) were estimated for each animal as the proportion of genome in ROH over the overall length of the autosomal genome covered by the involved SNPs (McQuillan et al., 2008):

$$F_{ROH} = \frac{\sum_{i=1}^{n} L_{ROH_i}}{L_{total}}$$

where L_{ROH_i} is the ROH length of individual i, and L_{total} is the total length of the autosomal chromosomes (2,487,768,066 bp). Still, for each class of ROH ($F_{ROH<2Mb}$, $F_{ROH2-4Mb}$, $F_{ROH4-8Mb}$, $F_{ROH8-16Mb}$, $F_{ROH>16Mb}$), inbreeding coefficient estimates were made dividing the total sum of ROH segments by the total length of the cattle autosomal genome. All F_{ROH} coefficients were calculated using the functions *Froh_inbreeding* and *Froh_inbreedingClass* from the R package detectRUNS version 0.9.6 (https://r-project.org; Biscarini et al., 2019).

4.3.4. Pedigree

The pedigree included 433,823 animals (121,497 dams and 7,200 sires) spanning up to 14 overlapping generations. Animals with both known sire and dam made up 77.47% of the pedigree, 4.10% had only sire known, 15.68% only dam known, and 2.75% had both sire and dam unknown.

The classical inbreeding coefficient based on all information in the pedigree (F_{PED_FULL}) were estimated for each animal using the function *pedInbreeding* from the R package optiSel version 2.0.5 (https://r-project.org; Wellmann, 2021). The F_{PED_FULL} was defined as the pedigree-based probability that two alleles at a random locus in an individual were IBD (Falconer and Mackay, 1996). F_{PED_FULL} was presented in the descriptive statistics, but only the pedigree-based inbreeding coefficients of the genotyped animals (F_{PED}) was used for comparison with F_{ROH}.

4.3.5. Alternative inbreeding coefficients estimation methods

Three alternative estimates of genomic inbreeding coefficients were also estimated:

i) F_{G1}: inbreeding coefficient based on genotype additive variance, computed from the diagonal elements minus 1 of the genomic relationship matrix obtained according to VanRaden's method 1 (VanRaden, 2008), as follows:

$$F_{G1} = \frac{\sum_{i=1}^{n}(x_i - 2p_i)^2}{2\sum_{i=1}^{n} p_i(1 - p_i)} - 1$$

where x_i is the number of reference allele copies for SNP i; p_i is the reference allele frequency in the population, and n is the number of SNP.

ii) F_{G2}: inbreeding coefficient based on the correlation between uniting gametes, computed from the diagonal elements minus 1 of the genomic relationship matrix obtained according to Yang (Yang et al., 2010), as follows:

$$F_{G2} = \frac{1}{n}\sum_{1}^{n} \frac{x_i^2 - (1 + 2p_i)x_i + 2p_i^2}{2p_i(1 - p_i)}$$

where x_i, p_i, and n are as for G_1. The relationship matrix G_2 is similar to G_1, but G_2 gives more weight to homozygotes for the minor allele than to homozygotes for the major allele and it has a lower sampling variance than the G_1 because it accounts for the sampling error associated with each SNP (Yang et al., 2010).

iii) F_{HOM}: inbreeding coefficient based on the homozygous genotypes observed and expected (Purcell et al., 2007), as follows:

$$F_{HOM} = \frac{H_{exp} - H_{obs}}{L - H_{exp}}$$

where H_{exp} is the expected value for homozygous genotypes, H_{obs} is the observed value for the homozygous genotypes and L is the number of non-missing autosomal SNPs.

4.3.6. Correlation among inbreeding coefficients estimation methods

After evaluation of data distributions (Supplementary Figure 1) with the Shapiro-Wilk test, Spearman rank correlation was used to assess relationships among inbreeding coefficients. Correlations may be overestimated by using Pearson’s correlation, less adequate for non-normal distributions (Gurgul et al., 2016). Since ROH of different length can have different capability to be autozygous, the correlations were also calculated considering ROH classes. A heatmap was created for better visualization of the results.

4.3.7. Phenotypes

The phenotypic data comprised 197,283 measurements of productive and reproductive traits from Guzerá males and females born between 1954 and 2018. The available phenotypes were: 305-day first-lactation cumulative yields (kg) of milk (MY305), fat (FY305), protein

(PY305), lactose (LY305), and total solids (SY305); age at first calving in days (AFC); and adjusted body weight (kg) at the ages of 210 days (W210), 365 (W365) and 450 (W450) days.

The descriptive statistics of the traits are shown in Table 1.

Table 1. Number of animals with record (and genotyped) (N), mean and standard deviation (SD), minimum (Min) and maximum (Max) included in the analyses of each trait

Traits[a]	**N**	**Mean ± SD**	**Min**	**Max**
MY305 (kg)	5,229 (987)	1991.17 ± 980.90	105.00	6487.00
FY305 (kg)	1,853 (627)	82.82 ± 37.06	6.00	281.17
PY305 (kg)	1,543 (636)	62.08 ± 27.16	4.00	232.00
LY305 (kg)	1,457 (603)	77.50 ± 34.15	5.00	226.00
SY305 (kg)	1,228 (518)	227.02 ± 96.30	13.00	722.00
AFC (days)	83,244 (1,044)	1251.50 ± 208.84	671.00	1680.00
W210 (kg)	122,684 (630)	173.70 ± 39.92	50.00	300.00
W365 (kg)	88,065 (591)	227.41 ± 53.36	67.00	414.00
W450 (kg)	88,456 (591)	275.75 ± 58.58	102.27	499.00

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight.

The single-step genomic best linear unbiased prediction (ssGBLUP) method was applied, using the following single-trait mixed linear model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1a} + \mathbf{Z_2m} + \mathbf{Z_3mp} + \mathbf{e}$$

where $\mathbf{y}$ is the vector of phenotypes; $\mathbf{b}$ is the vector of fixed effects of a contemporary group and covariate; $\mathbf{a}$ is the vector of random additive direct genetic effects, with $\mathbf{a} \sim \mathrm{N}(0, \mathbf{H}\sigma_a^2)$; $\mathbf{m}$ is the vector of random maternal genetic effects, with $\mathbf{m} \sim \mathrm{N}(0, \mathbf{H}\sigma_m^2)$; $\mathbf{mp}$ is the vector of random maternal permanent environmental effects, with $\mathbf{mp} \sim \mathrm{N}(0, \mathbf{I}\sigma_{mp}^2)$; $\mathbf{Z_1}$, $\mathbf{Z_2}$ and $\mathbf{Z_3}$ are incidence matrices related to the $\mathbf{a}$, $\mathbf{m}$ and $\mathbf{mp}$ to $\mathbf{y}$, respectively; and $\mathbf{e}$ is the residual vector, with $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$. On the assumptions, σ_a^2, σ_m^2, σ_{mp}^2 and σ_e^2 are the additive direct genetic, maternal genetic, maternal permanent environmental and residual variance components, respectively; **H** is a matrix that combines pedigree and genomic information; and **I** is an identity matrix. The maternal effects (genetic and permanent environment) were assumed only for W210.

Contemporary groups (CGs) for MY305, FY305, PY305, LY305, and SY305 were formed by combining herd, year, and season of calving, and age at calving was considered as a linear covariate. For AFC, W210, W365, and W450, the CGs were a combination of herd, year, and season of birth. In addition, for growth traits, sex was also included in the CG, and the age of dams at calving was considered as a linear covariate. Both calving and birth seasons were defined as dry (April to September) or rainy (October to March). Data from CG with fewer than three records and with only one sire as parent were excluded.

4.3.8. Inbreeding depression

The F_{PED} and F_{ROH} were used to access inbreeding depression in the evaluated traits. Due to the small number of animals with both phenotype and genotype, it was decided not to use inbreeding coefficients as a covariate in the animal model. Thus, inbreeding depression was estimated by linear regression of corrected phenotypes on inbreeding coefficients, for each inbreeding estimation method. The corrected phenotypes for each trait and animal were calculated as the sum of the breeding value and the residual.

The regression equation was:

$$y_i = \mu + b(x_i - \bar{x}) + e_i$$

where y_i is the corrected phenotype of the animal i; μ is the intercept; b is the slope associated with the inbreeding coefficient (inbreeding depression); x_i is the inbreeding coefficient of the animal i; $\bar{x}$ is the average inbreeding coefficient; and e_i is the residual, with $e_i \sim N(0, \sigma^2)$.

4.4. Results

4.4.1. Runs of homozygosity

ROH were identified in all 1,733 individuals, totaling 16,379 homozygous segments distributed on 29 autosomal chromosomes. A summary of the ROH measurements is shown in Table 2.

Table 2. Summary of the runs of homozygosity (ROH) measurements

ROH measurement	**Mean ± SD**[a]	**Min**[a]	**Max**[a]
Number of ROH by individual	9.45 ± 4.21	1.00	31.00
Number of SNP in a ROH	152.00 ± 148.00	30.00	2,608.00
Number of ROH by chromosome	565.00 ± 293.00	203.00	1,563.00
Average general length of ROH (Mb)	6.76 ± 7.12	0.50	128.97
Average length of ROH among chromosomes (Mb)	6.01 ± 0.82	4.30	7.04

[a] SD=standard deviation; Min=minimum; Max=maximum.

BTA29 had the lowest number of ROH (203) and BTA5 and BTA6 had the highest number of ROH (1,387 and 1,563, respectively). BTA5 and BTA6 also showed higher numbers of animals with ROH: 991 animals with ROH on BTA5, and 1,063 animals with ROH on BTA6.

Regarding ROH classes, class $ROH_{<2Mb}$ presented 2,828 ROHs, with an average length of 0.75 Mb. Class $ROH_{2–4Mb}$ showed the highest number of ROHs, being 5,441 ROHs with an average length of 2.92 Mb. Classes $ROH_{4–8Mb}$ and $ROH_{8–16Mb}$ showed 4,391 and 2,586 ROHs respectively, with average lengths equal to 5.64 and 11.11 Mb respectively. The last class, $ROH_{>16Mb}$, showed 1,133 ROHs with average length of 25.80 Mb. The number of ROH by chromosome and class, as well as the average ROH length by chromosome and class can better visualized in Figure 1.

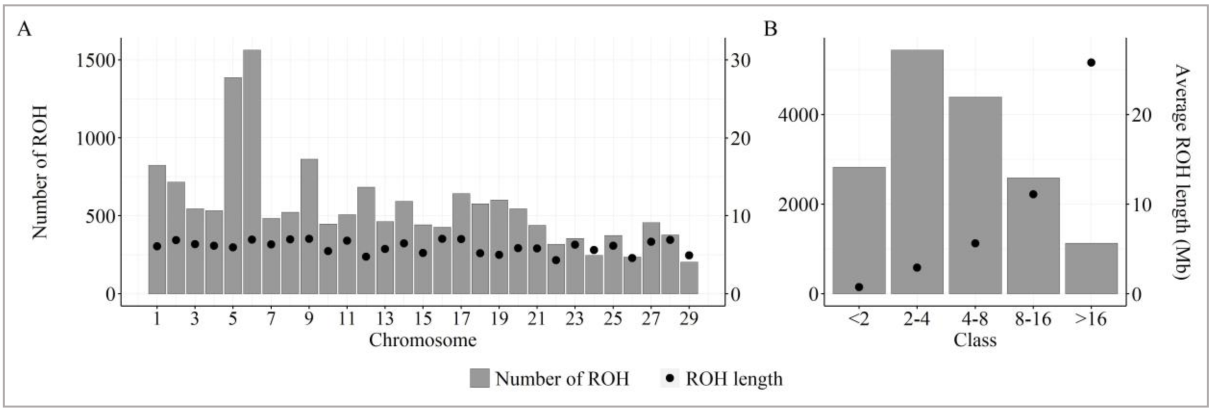


Figure 1. Number of ROH (principal axis; bar) and average ROH length (secondary axis; point) by (A) chromosome, and (B) ROH classes.

4.4.2. Pedigree and ROH inbreeding coefficients

Descriptive statistics for F_{PED_FULL}, F_{PED}, F_{ROH}, and classes of F_{ROH} coefficients are presented in Table 3.

Table 3. General descriptive statistics of the pedigree-based inbreeding coefficients considering full pedigree (F_{PED_FULL}) and only for genotyped animals (F_{PED}), and inbreeding coefficients based on runs of homozygosity total (F_{ROH}) and of different lengths ($F_{ROH<2Mb}$, $F_{ROH2-4Mb}$, $F_{ROH4-8Mb}$, $F_{ROH8-16Mb}$, and $F_{ROH>16Mb}$) for genotyped animals (N)

Inbreeding coefficients estimation method	**N**	**Mean ± SD**	**Median**	**Min[a]**	**Max[a]**
F_{PED_FULL}	433,823	0.007 ± 0.025	0.000	0.000	0.406
F_{PED}	1,733	0.007 ± 0.022	0.000	0.000	0.289
F_{ROH}	1,733	0.023 ± 0.0003	0.019	0.0003	0.144
$F_{ROH<2Mb}$	1,733	0.023 ± 0.0003	0.019	0.0003	0.144
$F_{ROH2-4\ Mb}$	1,730	0.023 ± 0.001	0.018	0.001	0.143
$F_{ROH4-8\ Mb}$	1,672	0.020 ± 0.017	0.015	0.002	0.135
$F_{ROH8-16\ Mb}$	1,388	0.017 ± 0.016	0.012	0.003	0.124
$F_{ROH>16\ Mb}$	700	0.017 ± 0.014	0.011	0.006	0.095

[a] SD=standard deviation; Min=minimum; Max=maximum.

The average of F_{PED_FULL} was 0.007 ± 0.025. Of the 1,733 genotyped animals, 93% had complete parentage information (sire and dam), in 5.94% only the dam was known, in 0.17% only the sire was known, and in 0.75% both the sire and dam were unknown, and the average F_{PED} was 0.007 ± 0.022. Furthermore, 1,351 animals had an F_{PED} of zero, and of these, 91% had complete parentage information and only 0.96% had no sire and/or dam information. The 1,351 genotyped animals with F_{PED} equal to zero had an average F_{ROH} of 0.021 ± 0.015, ranging from 0.0003 to 0.128. Considering all 1,733 genotyped animals, the average FROH was 0.023 ± 0.0003.

Estimates of inbreeding coefficients by known year of birth are shown in Table 4. Most of the animals were born after 2000. The mean values for F_{PED_FULL}, F_{PED}, and F_{ROH} were higher in the younger animals, i.e., those born after 2010. In addition, there were 330,441 (77.5%) and 1,364 (78.3%) individuals with F_{PED_FULL} and F_{PED} equal to zero, respectively, and there was no F_{ROH} equal to zero (Table 5). The second most frequent class of inbreeding coefficient ranged from 0 to 0.01 for both F_{PED_FULL} and F_{PED} and from 0.01 to 0.02 for F_{ROH}. The number of individuals decreased with increasing inbreeding coefficient class, with only 1.7, 1.4, and 0.5% of individuals with F_{PED_FULL}, F_{PED}, and F_{ROH}, higher than 0.10, respectively.

Table 4. Number of individuals (N), mean, standard deviation (SD), median, minimum (Min) and maximum (Max) inbreeding coefficients according to birth year period (Period) by inbreeding coefficients estimation method (F)

Period	**F**[a]	**N**	**Mean ± SD**	**Median**	**Min**	**Max**
≤ 1979	F_{PED_FULL}	7,622	0.002 ± 0.019	0.000	0.000	0.250
	F_{PED}	0	-	-	-	-
	F_{ROH}	0	-	-	-	-
1980 - 1989	F_{PED_FULL}	16,101	0.002 ± 0.022	0.000	0.000	0.250
	F_{PED}	1	0.000 ± 0.000	0.000	0.000	0.000
	F_{ROH}	1	0.028	-	-	-
1990 - 1999	F_{PED_FULL}	39,245	0.002 ± 0.017	0.000	0.000	0.375
	F_{PED}	84	0.000 ± 0.000	0.000	0.000	0.000
	F_{ROH}	84	0.016 ± 0.013	0.012	0.001	0.059
2000 - 2009	F_{PED_FULL}	174,153	0.004 ± 0.021	0.000	0.000	0.375
	F_{PED}	946	0.003 ± 0.011	0.000	0.000	0.125
	F_{ROH}	946	0.021 ± 0.017	0.017	0.000	0.144
≥ 2010	F_{PED_FULL}	189,443	0.011 ± 0.029	0.000	0.000	0.406
	F_{PED}	712	0.014 ± 0.032	0.000	0.000	0.289
	F_{ROH}	712	0.027 ± 0.019	0.022	0.000	0.132

[a] F_{PED_FULL} = pedigree-based inbreeding coefficients for full pedigree; F_{PED} = pedigree-based inbreeding coefficients for genotyped animals; F_{ROH} = inbreeding coefficients based on runs of homozygosity.

Table 5. Number (N) and percentage (%) of individuals, mean, standard deviation (SD), median, minimum (Min) and maximum (Max) inbreeding coefficients according to inbreeding coefficient class (F class) by inbreeding coefficients estimation method (F)

F class	**F^a**	**N**	**%**	**Mean ± SD**	**Median**	**Min**	**Max**
F=0	F_{PED_FULL}	330,411	77.5	0.000 ± 0.000	0.000	0.000	0.000
	F_{PED}	1,364	78.3	0.000 ± 0.000	0.000	0.000	0.000
	F_{ROH}	0	-	-	-	-	-
$0 < F \leq 0.01$	F_{PED_FULL}	42,977	10.1	0.004 ± 0.003	0.003	0.0000	0.010
	F_{PED}	119	6.8	0.003 ± 0.003	0.002	0.0000	0.010
	F_{ROH}	322	18.5	0.006 ± 0.002	0.007	0.0003	0.010
$0.01 < F \leq 0.02$	F_{PED_FULL}	15,575	3.7	0.016 ± 0.002	0.016	0.010	0.020
	F_{PED}	64	3.7	0.015 ± 0.002	0.016	0.10	0.020
	F_{ROH}	622	35.7	0.015 ± 0.003	0.015	0.10	0.020
$0.02 < F \leq 0.04$	F_{PED_FULL}	17,083	4.0	0.031 ± 0.004	0.031	0.020	0.040
	F_{PED}	107	6.1	0.031 ± 0.003	0.031	0.023	0.040
	F_{ROH}	582	33.4	0.028 ± 0.005	0.027	0.20	0.040
$0.04 < F \leq 0.06$	F_{PED_FULL}	2,811	0.7	0.047 ± 0.004	0.047	0.040	0.060
	F_{PED}	13	0.7	0.049 ± 0.004	0.047	0.040	0.055
	F_{ROH}	134	7.7	0.048 ± 0.006	0.047	0.40	0.059
$0.06 < F \leq 0.10$	F_{PED_FULL}	10,330	2.4	0.067 ± 0.008	0.063	0.060	0.100
	F_{PED}	51	2.9	0.066 ± 0.008	0.063	0.063	0.094
	F_{ROH}	74	4.2	0.074 ± 0.010	0.072	0.60	0.098
$0.10 < F \leq 0.20$	F_{PED_FULL}	5,650	1.3	0.131 ± 0.014	0.125	0.100	0.199
	F_{PED}	23	1.3	0.135 ± 0.019	0.126	0.102	0.192
	F_{ROH}	9	0.5	0.122 ± 0.011	0.116	0.110	0.144
$F \geq 0.20$	F_{PED_FULL}	1,716	0.4	0.254 ± 0.015	0.25	0.202	0.406
	F_{PED}	2	0.1	0.270 ± 0.028	0.270	0.250	0.289
	F_{ROH}	0	-	-	-	-	-

[a] F_{PED_FULL} = pedigree-based inbreeding coefficients for full pedigree; F_{PED} = pedigree-based inbreeding coefficients for genotyped animals; F_{ROH} = inbreeding coefficients based on runs of homozygosity.

4.4.3. Correlation among inbreeding coefficients estimate methods

Low correlation was observed between F_{PED} and F_{ROH} (0.31) and between F_{PED} and F_{ROH} in different classes (0.28 to 0.32) (Figure 2). In addition, F_{PED} also showed low correlations with F_{G1} (0.31), F_{G2} (0.34), and F_{HOM} (0.36). When F_{ROH} was compared with F_{G1} and F_{G2}, the correlations were slightly higher but still low-to-moderate (0.43 and 0.49, respectively). The correlation between F_{ROH} and F_{HOM} was high (0.78) and decreased with ROH length. All correlations were significant ($p<0.001$) and the full Spearman correlation matrix of the inbreeding coefficients can be found in Supplementary Table 1.

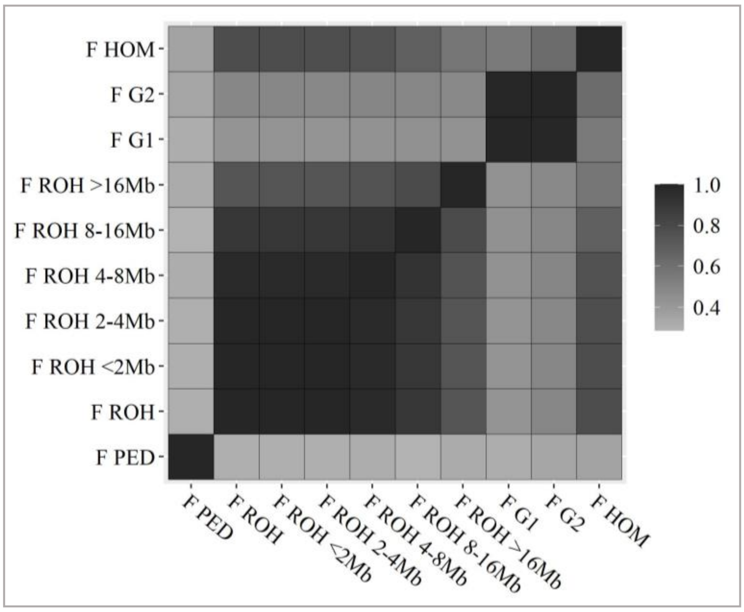


Figure 2. Spearman correlation among inbreeding estimation methods.

4.4.4. Inbreeding depression

Linear regression of corrected phenotypes on F_{PED} and F_{ROH} by trait can be visualized in Figure 3. For better visualization, the regression coefficients associated with inbreeding depression (b) are described in Table 6. All inbreeding coefficients showed no significant effect (p-value>0.05) on the traits evaluated.

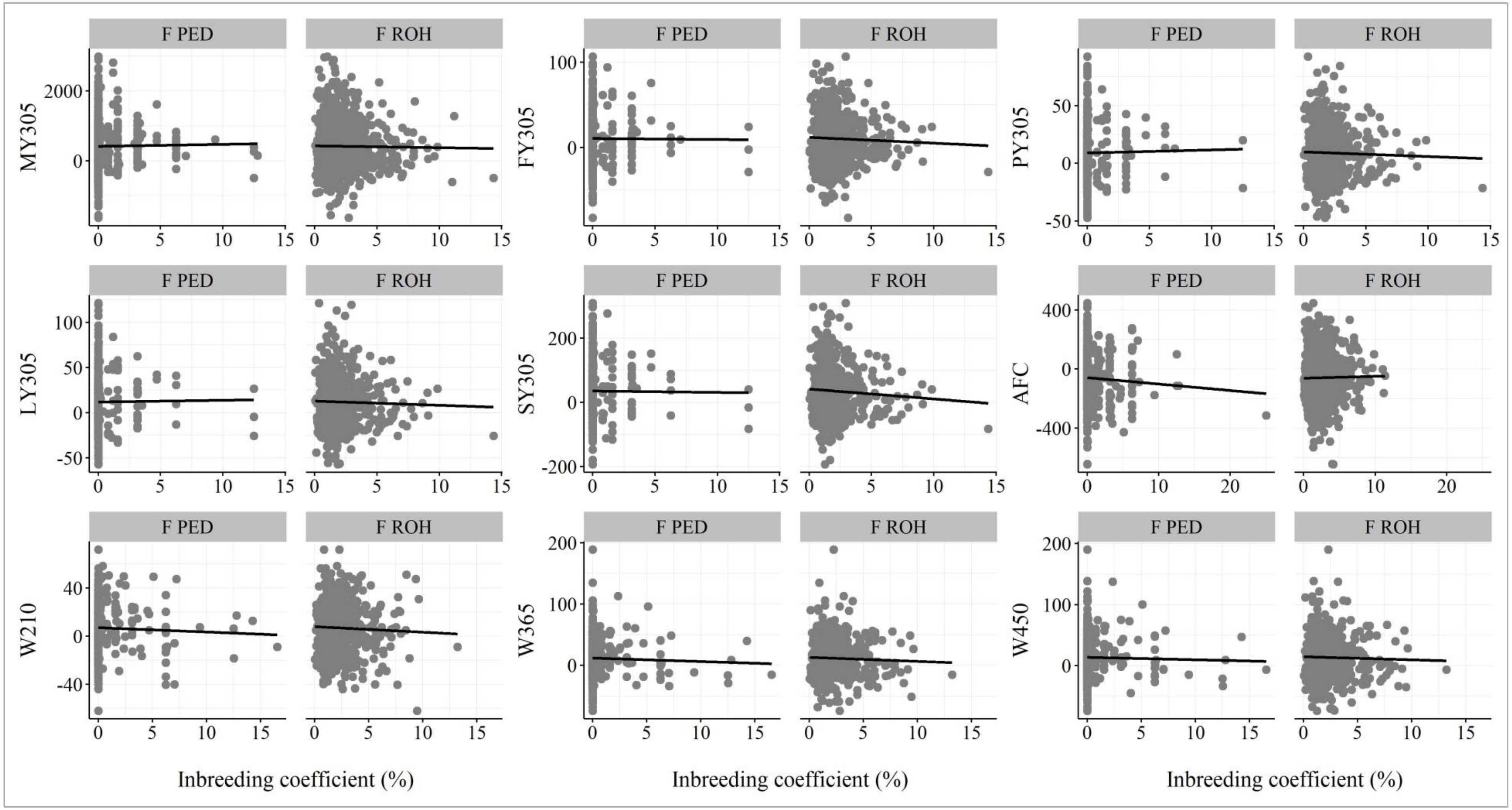


Figure 3. Linear regression of corrected phenotypes on pedigree-based (F_{PED}) and based in runs of homozygosity (F_{ROH}) inbreeding coefficients by trait (MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids224 yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight).

Table 6. Effects of the increase in pedigree-based (F_{PED}) and based on runs of homozygosity (F_{ROH}) inbreeding coefficients on the evaluated traits

Trait[a]	**Inbreeding coefficients estimation method**	**b**[b] **(SE)**	**p-value**
MY305	F_{PED}	5.65 (15.10)	0.724
	F_{ROH}	-5.25 (12.72)	0.680
FY305	F_{PED}	-0.11 (0.90)	0.898
	F_{ROH}	-0.68 (0.70)	0.329
PY305	F_{PED}	0.26 (0.75)	0.727
	F_{ROH}	-0.40 (0.54)	0.455
LY305	F_{PED}	0.19 (0.94)	0.842
	F_{ROH}	-0.46 (0.71)	0.513
SY305	F_{PED}	-0.43 (2.61)	0.868
	F_{ROH}	-3.10 (2.04)	0.130
AFC	F_{PED}	-4.37 (2.91)	0.133
	F_{ROH}	1.30 (2.76)	0.637
W210	F_{PED}	-0.36 (0.45)	0.426
	F_{ROH}	-0.46 (0.46)	0.314
W365	F_{PED}	-0.56 (0.68)	0.411
	F_{ROH}	-0.65 (0.70)	0.351
W450	F_{PED}	-0.40 (0.77)	0.599
	F_{ROH}	-0.51 (0.79)	0.522

[a] MY305 = 305-day milk yield; FY305 = 305-day fat yield; PY305 = 305-day protein yield; LY305 = 305-day lactose yield; SY305 = 305-day total solids224 yield; AFC = age at first calving; W210 = 210-day weight; W365 = 365-day weight; W450 = 450-day weight.

[b] Slope associated with the inbreeding coefficient (inbreeding depression).

* Significant to 0.05%.

4.5. Discussion

In this study, we mainly aimed to analyze the inbreeding coefficient estimates derived from the ROH for the Brazilian population of Guzerá cattle and their impact on economically important traits of this breed. In addition, pedigree-based inbreeding coefficients and three other measures of genomic inbreeding coefficients were obtained to compare the estimates.

We present ROH classified into five length classes ($ROH_{<2Mb}$, $ROH_{2–4Mb}$, $ROH_{4–8Mb}$, $ROH_{8–16Mb}$, and $ROH_{>16Mb}$). Although Purfield et al. (2012) reported that the Bovine SNP50 fails to detect ROH < 1Mb, we chose to present results for all five ROH classes because this is an initial study involving ROH including Guzerá animals genotyped with the GGP *indicus* panel and using imputed data for this panel. In this way, we want evaluate the results under these conditions.

The greatest number of ROH per chromosome was described on BTA5 and BTA6, coinciding with the results observed in other indicine breeds such Gyr (BTA5; Peripolli et al., 2018) and Nellore (BTA5 and BTA6; Mulim et al., 2022). Studies in taurine breeds (Gurgul et al., 2016; Mastrangelo et al., 2016; Caivio-Nasner et al., 2021) have showed the highest number of ROH on BTA1. In our study, there was no relationship between the number of ROHs with chromosome length, as pointed out by Caivio-Nasner et al. (2021) and Mastrangelo et al. (2016) where the number of ROHs tended to decrease with chromosome length.

Several studies in cattle have reported the influence of genes on BTA5 and BTA6 on milk production and its components such as fat, protein and lactose (Freyer et al., 2002; Schopen et al., 2011; Iung et al., 2019; Paiva et al., 2020; Bohlouli et al., 2022). Milk yield is the main trait under selection (Bruneli et al., 2020) in the subpopulation of genotyped animals included in our study and this trait shows a high genetic correlation with milk components such as fat and protein (Carrara et al., 2022). One of the effects of selection is the increased frequency of favorable alleles in important genes and the possibility of homozygosity. Thus, selection for milk production may be one of the reasons for a higher number of ROH in BTA5 and BTA6 in the population analyzed.

The longest ROH was found on BTA1 with 128.97 Mb, however, results in Cinisara taurine cattle (Mastrangelo et al., 2016) and Gyr indicine cattle (Peripolli et al., 2018) have reported the longest on BTA8 (112.65 and 108.97 Mb, respectively). The majority of ROHs were classified as 2-4 Mb, representing 33% of all ROHs found, corroborating what has been reported for the Gyr, Brahman, and Nellore indicine breeds, as well as for other taurine breeds, such as Angus, Hereford, Holstein, and Jersey (Mulim et al., 2022). Probably $ROH_{2\text{-}4}$ were formed in older generations, considering that the recombination events that take place in each generation split the homozygous segments into smaller haploblocks, as also pointed out by Mulim et al. (2022). However, there was not a large discrepancy in the number of ROHs for classes 2-4 compared to classes less than 16Mb in length. Furthermore, only 7% of the ROHs are longer than 16Mb, suggesting that recent recombination events are rarer in this population. It is important to emphasize that there is great variability in the literature regarding the parameters used to identify ROH and that each population has its own characteristics. Therefore, comparison between studies must be interpreted cautiously, as reinforced by Fonseca et al. (2016), Mulim et al. (2022) and Peripolli et al. (2017).

The F_{PED_FULL} and F_{PED} found can be considered to be low when compared to inbreeding coefficients reported in previous studies in Guzerá breed (Pereira et al., 2016), in indicine Gyr cattle (Santana et al., 2014; Peripolli et al., 2018) and taurine breeds (Zhang et al., 2015; Dadousis et al., 2022). An interesting result emerged when comparing our results with the study of Peixoto et al. (2010). These authors used a subpopulation of Guzerá animals selected for milk and born by 2007 and found a mean pedigree-based inbreeding coefficient of 0.009 with the majority of inbred animals with inbreeding coefficient equal to or less than 0.01. More than 10 years later, we found similar results for F_{PED}, including inbreeding coefficients by year of birth. This fact shows that the breeders are careful to avoid mating between related individuals. We used F_{PED} for comparison because although beef, milk and dual-purpose populations are closely connected (Peixoto et al., 2021), the genotyped animals are mostly part of herds that select for milk or dual-purpose.

Although many animals with F_{PED} equal to zero have known parents (93%), the depth of the pedigree of these animals directly affects their inbreeding coefficients. If the pedigree is unable to link the parents to at least one common ancestor, the inbreeding coefficient of the animal may be underestimated or even zero. In this sense, all individuals whose F_{PED} was zero had some level of F_{ROH}, and the average F_{ROH} was slightly higher than the average F_{PED}. The variation between these two estimates can be attributed to the fact that pedigree-based inbreeding coefficient is an expected measure, i.e., a probabilistic measure of the proportion of the genome that is in autozygosity. However, this expectation does not capture recombination events, making ROHs superior in this regard (Keller et al., 2011; Curik et al., 2014).

Moreover, according to (Falconer and Mackay, 1996), the inbreeding coefficient compares the degree of relationship between individuals now, with the degree of relationship between individuals in the base population, i.e., it expresses how much dispersive process has occurred from the base population. Thus, the base population is directly related to the determination of the inbreeding coefficients, which could also explain the differences between the F_{PED} and F_{ROH} estimates.

F_{ROH} values were lower than those reported for Gyr cattle (Peripolli et al., 2018; Mulim et al., 2022) and higher than those reported for Nellore cattle (Mulim et al., 2022) and taurine breeds (Zhang et al., 2015; Dadousis et al., 2022). However, these values may vary even within the same breed (Peripolli et al., 2018; Mulim et al., 2022), as they depend on the number of animals genotyped, the density of markers, and the parameters used to determine ROHs.

Overall, correlations between F_{PED} and genomic-based inbreeding coefficients methods were low. The "old" inbreeding cannot be traced precisely from pedigree data, but can be traced from short ROH (Ferencakovic et al., 2011). Along these lines, studies show a lower correlation between pedigree-based inbreeding coefficients and F_{ROH} for ROH < 4Mb compared to ROH > 4Mb in cattle (Ferencakovic et al., 2011; Peripolli et al., 2018), but this behavior was not observed in our study, where the correlation between F_{PED} and F_{ROH} barely changed with ROH length. In turn, F_{ROH} showed higher correlations with the other genomic methods.

Importantly, the different methods used to estimate inbreeding coefficients have their particularities. Whether by the probability of two alleles being IBD at a locus (Malecót, 1948) or the theoretical correlation between the two gametes in union (Wright, 1965), determining the pedigree-based inbreeding coefficient provides expected estimates for this parameter and does not account for the stochastic nature of Mendelian inheritance. F_{G1} is based on additive genetic variance (VanRaden, 2008) and F_{G2} on the correlation of the union of gametes (Yang et al., 2010). Both G1 and G2 depend on the initial allele frequencies and they weight rare alleles, i.e., rare homozygotes contribute more to inbreeding than frequent homozygotes. In turn, F_{HOM} and F_{ROH} do not require estimation of allele frequencies (McQuillan et al., 2008) and directly reflect homozygosity levels in the genome (Zhang et al., 2015). In this sense, it is plausible that the correlation between F_{ROH} and F_{HOM} in our study has been moderate, as also reported in other studies in cattle (Peripolli et al., 2018; Freitas et al., 2021; Mulim et al., 2022).

It should be noted that many studies use the Pearson correlation instead of the Spearman correlation to compare inbreeding coefficient estimates obtained by different methods. Since the data distribution is different than normal, especially for F_{PED} data, the former may overestimate the correlations in this case, as pointed out by (Gurgul et al., 2016). Indeed, as can be seen in Supplementary Table 1, the Pearson correlations between the inbreeding coefficient estimates obtained in our study were slightly higher than the corresponding Spearman correlations.

In addition to the low level of F_{PED_FULL}, F_{PED} and F_{ROH} in the studied Guzerá cattle population compared to other cattle breeds, no inbreeding depression was observed in the main traits of economic interest for the breed.

Although the inbreeding rate estimated in this study was not high enough to adversely affect the performance of most traits evaluated for this population, inbreeding depression was detected in several livestock and poultry species including cattle (Santana et al., 2010; Pereira et al., 2016; Doekes et al., 2019), pig (Saura et al., 2015), poultry (Sewalem et al., 1999; Xue et al., 2021), and horse (Bussiman et al., 2018), using both pedigree-based and genome-based methods.

Maintaining low levels of inbreeding is critical in cattle breeding populations to avoid inbreeding depression. The Guzerá breed has experienced some historical events in recent decades, such as population bottlenecks and a strong founder effect. Nevertheless, the breed has maintained a good level of genetic diversity (Fonseca et al., 2016; Peixoto et al., 2021), a lower level of inbreeding than other cattle breeds, and still no significant effects of inbreeding level on dairy and beef cattle traits.

4.6. Conclusion

Despite the limited number of animals imported to form the first Guzerá lines in Brazil and the bottleneck effect of using the animals in crossbreeding systems, the autozygotic proportion of the genome in this population was low compared to other cattle populations. Pedigree-based inbreeding coefficients were also low.

Pedigree-based inbreeding coefficients showed low correlation with inbreeding coefficients obtained by genomic-based methods, including those obtained by runs of homozygosity. In turn, the inbreeding coefficients based on runs of homozygosity showed low-to-moderate correlation with the other genomic-based methods.

No inbreeding depression was observed in the reproductive, dairy and beef traits evaluated. Compared to previous studies, there was no increase in inbreeding coefficients in this population. Nevertheless, continuous monitoring of inbreeding levels is essential to prevent loss of genetic variability in the population.

4.7. Acknowledgments

The authors acknowledge the Centro Brasileiro de Melhoramento Genético do Guzerá (CBMG^2) and Associação Brasileira de Criadores de Zebu (ABCZ) for providing datasets. The authors also acknowledge the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and INCT (Instituto Nacional de Ciência e Tecnologia) for financial support.

4.8. References

Aliloo H, Mrode R, Okeyo AM, Ni G, Goddard ME and Gibson JP 2018. The feasibility of using low-density marker panels for genotype imputation and genomic prediction of crossbred dairy cattle of East Africa. Journal of Dairy Science 101, 9108–9127.

Biscarini F, Cozzi P, Gaspa G and Marras G 2019. DetectRUNS: Detect Runs of Homozygosity and Runs of Heterozygosity in Diploid Genomes. R package version 0.9.6.

Bjelland DW, Weigel KA, Vukasinovic N and Nkrumah JD 2013. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. Journal of Dairy Science 96, 4697–4706.

Bohlouli M, Halli K, Yin T, Gengler N and König S 2022. Genome-wide associations for heat stress response suggest potential candidate genes underlying milk fatty acid composition in dairy cattle. Journal of Dairy Science 105, 3323–3340.

Brito LC, Peixoto MGCD, Carrara ER, Fonseca e Silva F, Ventura HT, Bruneli FAT and Lopes PS 2020. Genetic parameters for milk, growth, and reproductive traits in Guzerá cattle under tropical conditions. Tropical Animal Health and Production 52, 2251–2257.

Bruneli FAT, Peixoto MGCD, Júnior MLS, Pereira RJ, Arbex WA, Penna VM, Zadra LEF, Verneque R da S, Lôbo RB and Carvalho MRS 2020. National Breeding Program of Guzerá Cattle for Milk: Progeny Testing, National Zootechnical Archive and MOET Nucleus

Results. Brazilian Agricultural Research Corporation Documentos, 251, Juiz de Fora: Embrapa Dairy Cattle.

Bussiman FO, Perez BC, Ventura R V., Peixoto MGCD, Curi RA and Balieiro JCC 2018. Pedigree analysis and inbreeding effects over morphological traits in Campolina horse population. Animal 12, 2246–2255.

Caivio-Nasner S, López-Herrera A, González-Herrera LG and Rincón JC 2021. Diversity analysis, runs of homozygosity and genomic inbreeding reveal recent selection in Blanco Orejinegro cattle. Journal of Animal Breeding and Genetics 138, 613–627.

Carrara ER, de Genova Gaya L, de Paiva JT, Petrini J, Salvian M, Rovadoscki GA, Chaves Zanetti Reis A, Machado PF and Barreto Mourão G 2020. Impact of inbreeding on milk fatty acids of a Brazilian Holstein cattle. Animal Production Science 60, 1482.

Carrara ER, Peixoto MGCD, Veroneze R, Silva FF e, Ramos PVB, Bruneli FAT, Zadra LEF, Ventura HT, Josahkian LA and Lopes PS 2022. Genetic study of quantitative traits supports the use of Guzerá as dual-purpose cattle. Animal Bioscience.

Charlesworth D 2006. Evolution of Plant Breeding Systems. Current Biology 16, R726–R735.

Charlesworth D and Willis JH 2009. The genetics of inbreeding depression. Nature Reviews Genetics 10, 783–796.

Curik I, Ferenčaković M and Sölkner J 2014. Inbreeding and runs of homozygosity: A possible solution to an old problem. Livestock Science 166, 26–34.

Dadousis C, Ablondi M, Cipolat-Gotet C, van Kaam JT, Marusi M, Cassandro M, Sabbioni A and Summer A 2022. Genomic inbreeding coefficients using imputed genotypes: Assessing different estimators in Holstein-Friesian dairy cows. Journal of Dairy Science.

Doekes HP, Veerkamp RF, Bijma P, De Jong G, Hiemstra SJ and Windig JJ 2019. Inbreeding depression due to recent and ancient inbreeding in Dutch Holstein-Friesian dairy cattle. Genetics Selection Evolution 51, 1–16.

Falconer DS and Mackay TFC 1996. Introduction to Quantitative Genetics. Burnt Mill, England: Longman, New York.

Ferencakovic M, Hamzic E, Gredler B, Curik I and Sölkner J 2011. Runs of homozygosity reveal genomewide autozygosity in the Austrian Fleckvieh cattle. Agriculturae Conspectus Scientificus 76, 325–328.

Ferenčaković M, Sölkner J, Kapš M and Curik I 2017. Genome-wide mapping and estimation of inbreeding depression of semen quality traits in a cattle population. Journal of Dairy Science 100, 4721–4730.

Fonseca PA de S, Santos FC, Rosse IC, Ventura RV, Brunelli FÂT, Penna VM, da Silva Verneque R, Machado MA, da Silva MVGB, Carvalho MRS and Peixoto MGCD 2016. Retelling the recent evolution of genetic diversity for Guzerá: Inferences from LD decay, runs of homozygosity and Ne over the generations. Livestock Science 193, 110–117.

Freitas PHF, Wang Y, Yan P, Oliveira HR, Schenkel FS, Zhang Y, Xu Q and Brito LF 2021. Genetic Diversity and Signatures of Selection for Thermal Stress in Cattle and Other Two Bos Species Adapted to Divergent Climatic Conditions. Frontiers in Genetics 12.

Freyer G, Kühn C, Weikard R, Zhang Q, Mayer M and Hoeschele I 2002. Multiple QTL on chromosome six in dairy cattle affecting yield and content traits. Journal of Animal Breeding and Genetics 119, 200–200.

Gibson J, Morton NE and Collins A 2006. Extended tracts of homozygosity in outbred human populations. Human Molecular Genetics 15, 789–795.

Gurgul A, Szmatoła T, Topolski P, Jasielczuk I, Żukowski K and Bugno-Poniewierska M 2016. The use of runs of homozygosity for estimation of recent inbreeding in Holstein cattle. Journal of Applied Genetics 57, 527–530.

Hayes BJ, Visscher PM and Goddard ME 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genetics Research 91, 47–60.

Iung LHS, Petrini J, Ramírez-Díaz J, Salvian M, Rovadoscki GA, Pilonetto F, Dauria BD, Machado PF, Coutinho LL, Wiggans GR and Mourão GB 2019. Genome-wide association study for milk production traits in a Brazilian Holstein population. Journal of Dairy Science 102, 5305–5314.

Keller MC, Visscher PM and Goddard ME 2011. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics 189, 237–249.

Leroy G 2014. Inbreeding depression in livestock species: Review and meta-analysis. Animal Genetics 45, 618–628.

Malecót G 1948. Les Mathématiques de l'Hérédité. Libraries de l'académie de médecine.

Marras G, Gaspa G, Sorbolini S, Dimauro C, Ajmone-Marsan P, Valentini A, Williams JL and MacCiotta NPP 2015. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Animal Genetics 46, 110–121.

Mastrangelo S, Tolone M, Di Gerlando R, Fontanesi L, Sardina MT and Portolano B 2016. Genomic inbreeding estimation in small populations: Evaluation of runs of homozygosity in three local dairy cattle breeds. Animal 10, 746–754.

McQuillan R, Leutenegger A-L, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A, MacLeod AK, Farrington SM, Rudan P, Hayward C, Vitart V, Rudan I, Wild SH, Dunlop MG, Wright AF, Campbell H and Wilson JF 2008. Runs of Homozygosity in European Populations. The American Journal of Human Genetics 83, 359–372.

Meyermans R, Gorssen W, Buys N and Janssens S 2020. How to study runs of homozygosity using plink? a guide for analyzing medium density snp data in livestock and pet species. BMC Genomics 21, 1–14.

Miglior F, Szkotnicki B and Burnside EB 1992. Analysis of Levels of Inbreeding and Inbreeding Depression in Jersey Cattle. Journal of Dairy Science 75, 1112–1118.

Mulim HA, Brito LF, Pinto LFB, Ferraz JBS, Grigoletto L, Silva MR and Pedrosa VB 2022. Characterization of runs of homozygosity, heterozygosity-enriched regions, and population structure in cattle populations selected for different breeding goals. BMC Genomics 23, 209.

Paiva JT, Peixoto MGCD, Bruneli FAT, Alvarenga AB, Oliveira HR, Silva AA, Silva DA, Veroneze R, Silva FF and Lopes PS 2020. Genetic parameters, genome-wide association and gene networks for milk and reproductive traits in Guzerá cattle. Livestock Science 242, 104273.

Peixoto MGCD, Carvalho MRS, Egito AA, Steinberg RS, Bruneli FÂT, Machado MA, Santos FC, Rosse IC and Fonseca PAS 2021. Genetic Diversity and Population Genetic Structure of a Guzerá (*Bos indicus*) Meta-Population. Animals 11, 1125.

Peixoto MGCD, Poggian CF, Verneque RS, Egito AA, Carvalho MRS, Penna VM, Bergmann JAG, Viccini LF and Machado MA 2010. Genetic basis and inbreeding in the Brazilian Guzerat (*Bos indicus*) subpopulation selected for milk production. Livestock Science 131, 168–174.

Pereira RJ, Santana ML, Ayres DR, Bignardi AB, Menezes GRO, Silva LOC, Machado CHC, Josahkian LA and Albuquerque LG 2016. Inbreeding depression in Zebu cattle traits. Journal of Animal Breeding and Genetics 133, 523–533.

Peripolli E, Munari DP, Silva MVGB, Lima ALF, Irgang R and Baldi F 2017. Runs of homozygosity: current knowledge and applications in livestock. Animal Genetics 48, 255–271.

Peripolli E, Stafuzza NB, Munari DP, Lima ALF, Irgang R, Machado MA, Panetto JC do C, Ventura RV, Baldi F and da Silva MVGB 2018. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. BMC Genomics 19, 1–13.

Pryce JE, Haile-Mariam M, Goddard ME and Hayes BJ 2014. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genetics Selection Evolution 46, 71.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ and Sham PC 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics 81, 559–575.

Purfield DC, Berry DP, McParland S and Bradley DG 2012. Runs of homozygosity and population history in cattle. BMC Genetics 13.

Santana ML, Oliveira PS, Pedrosa VB, Eler JP, Groeneveld E and Ferraz JBS 2010. Effect of inbreeding on growth and reproductive traits of Nellore cattle in Brazil. Livestock Science 131, 212–217.

Santana ML, Pereira RJ, Bignardi AB, El Faro L, Pires MFÁ, Andrade RG, Perez BC, Bruneli FAT and Peixoto MGCD 2020. Dual-purpose Guzerá cattle exhibit high dairy performance under heat stress. Journal of Animal Breeding and Genetics 137, 486–494.

Santana ML, Pereira RJ, Bignardi AB, El Faro L, Tonhati H and Albuquerque LG 2014. History, structure, and genetic diversity of Brazilian Gir cattle. Livestock Science 163, 26–33.

Sargolzaei M, Chesnais JP and Schenkel FS 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15, 478.

Saura M, Fernández A, Varona L, Fernández AI, De Cara MÁR, Barragán C and Villanueva B 2015. Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data. Genetics Selection Evolution 47, 1–9.

Schopen GCB, Visker MHPW, Koks PD, Mullaart E, van Arendonk JAM and Bovenhuis H 2011. Whole-genome association study for milk protein composition in dairy cattle. Journal of Dairy Science 94, 3148–3158.

Sewalem A, Johansson K, Wilhelmson M and Lillpers K 1999. Inbreeding and inbreeding depression on reproduction and production traits of White Leghorn lines selected for egg production traits. British Poultry Science 40, 203–208.

VanRaden PM 2008. Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science 91, 4414–4423.

Wellmann R 2021. Package ‘optiSel’: Pedigree-based evaluations. Retrieved on 12 January 2021, from https://cran.r-project.org/web/packages/optiSel/vignettes/ped-vignette.html.

Wright S 1965. The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. Evolution 19, 395.

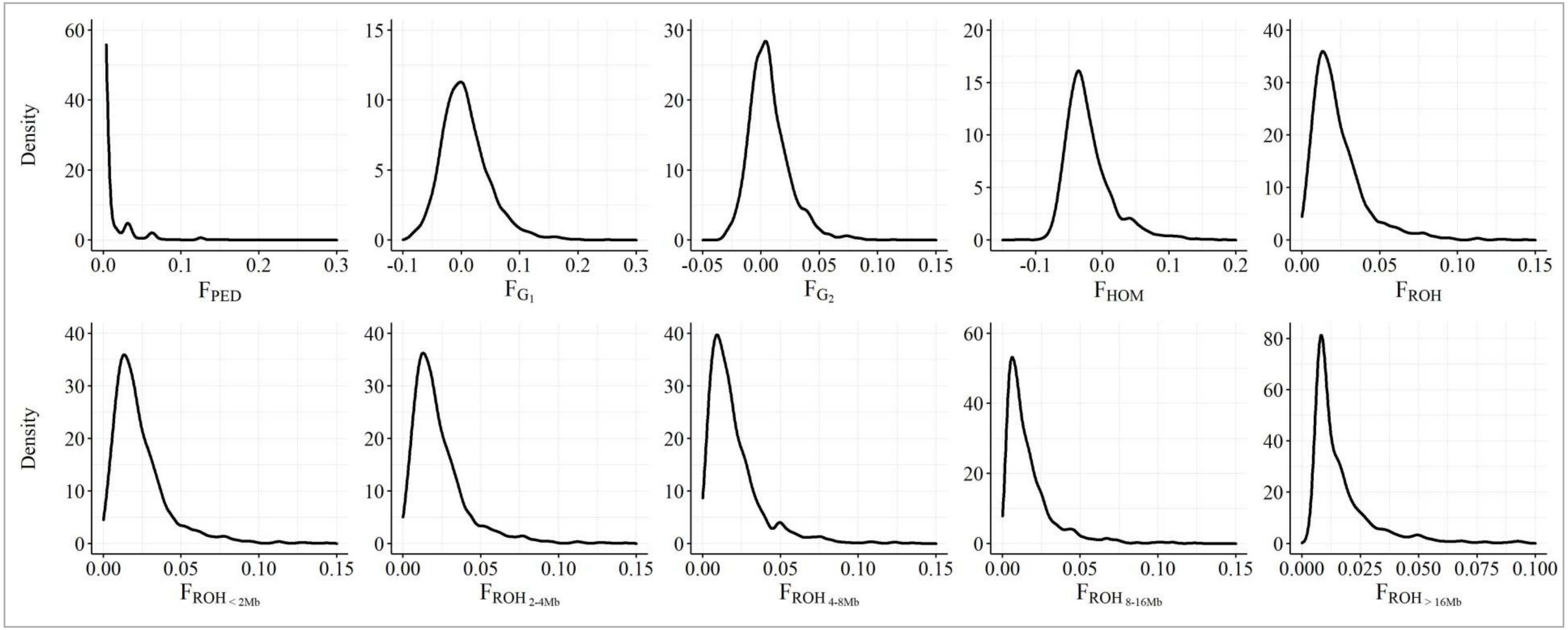
Xue Q, Li G, Cao Y, Yin J, Zhu Y, Zhang H, Zhou C, Shen H, Dou X, Su Y, Wang K, Zou J and Han W 2021. Identification of genes involved in inbreeding depression of reproduction in Langshan chickens. Animal Bioscience 34, 975–984.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME and Visscher PM 2010. Common SNPs explain a large proportion of the heritability for human height. Nature Genetics 42, 565–569.

Zhang Q, Calus MPL, Guldbrandtsen B, Lund MS and Sahana G 2015. Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. BMC Genetics 16, 1–11.

4.9. Supplementary material

4.9.1. Supplementary Figures



Supplementary Figure 1. Density plot of inbreeding coefficients estimated by different methods, significant for Shapiro-Wilk test (p-value < 0.001).

4.9.2. Supplementary Tables

Supplementary Table 1. Spearman correlations (above diagonal) and Pearson correlations (below diagonal) of inbreeding estimation methods.

	F_{PED}	F_{ROH}	$F_{ROH<2Mb}$	$F_{ROH2\text{-}4}$	$F_{ROH4\text{-}8}$	$F_{ROH8\text{-}16}$	$F_{ROH>16}$	F_{G1}	F_{G2}	F_{HOM}
F_{PED}	-	0.31	0.31	0.31	0.31	0.28	0.32	0.31	0.34	0.36
F_{ROH}	0.35	-	1.00	1.00	0.98	0.90	0.74	0.43	0.49	0.78
$F_{ROH<2Mb}$	0.35	1.00	-	1.00	0.98	0.90	0.74	0.43	0.49	0.78
$F_{ROH2\text{-}4\ Mb}$	0.35	1.00	1.00	-	0.98	0.90	0.74	0.43	0.49	0.78
$F_{ROH4\text{-}8\ Mb}$	0.37	0.99	0.99	0.99	-	0.92	0.75	0.44	0.49	0.76
$F_{ROH8\text{-}16\ Mb}$	0.36	0.96	0.96	0.96	0.97	-	0.79	0.44	0.49	0.69
$F_{ROH>16\ Mb}$	0.37	0.87	0.87	0.87	0.88	0.91	-	0.44	0.48	0.57
F_{G1}	0.35	0.55	0.55	0.55	0.56	0.55	0.53	-	0.99	0.55
F_{G2}	0.39	0.62	0.62	0.62	0.62	0.60	0.57	0.99	-	0.62
F_{HOM}	0.37	0.84	0.84	0.84	0.83	0.78	0.67	0.65	0.72	-

CHAPTER 5

5.1. General conclusions

The most important productive and reproductive traits selected in the improvement programs for Brazilian Guzera cattle were included in this thesis.

The genetic parameters (heritability and genetic correlations) showed that the traits can respond to selection and that there is no genetic antagonism between dairy (milk production and its components), beef (weight at 210, 365 and 450 days of age) and reproductive (age at first calving and scrotal perimeter at 365 and 450 days of age) traits in the studied population. This was a conclusive study of the feasibility of joint selection for milk and meat in this population, without loss in reproductive traits.

The inclusion of SNP genetic marker information in genetic evaluations proved beneficial in terms of increased accuracy of breeding values, especially for dairy traits. However, the number of genotyped animals was not yet sufficient to influence genetic evaluation in terms of predictive ability and bias.

Although the Brazilian population of Guzerá animals has experienced historical events such as the bottleneck effect, inbreeding coefficients based on runs of homozygosity and based on pedigree were low compared to other cattle breeds and compared to older studies in the breed, even in younger animals. There was no inbreeding depression in the evaluated traits. Nevertheless, continuous monitoring of inbreeding levels is essential to prevent loss of genetic variability in the population.

The findings from this thesis will contribute to the definition of selection criteria for dual-purpose cattle. In addition, this thesis provides the first results on the impact of incorporating information from SNP markers into genetic evaluations of dairy, reproductive and growth traits in Guzerá cattle (in terms of predictive ability, bias, accuracy of breeding values, rank correlations and genetic trends).