

RESEARCH ARTICLE

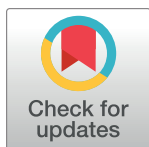
Single and multi-trait genomic prediction for agronomic traits in *Euterpe edulis*

Guilherme Bravim Canal¹, Cynthia Aparecida Valiati Barreto², Francine Alves Nogueira de Almeida¹, Iasmine Ramos Zaidan¹, Diego Pereira do Couto¹, Camila Ferreira Azevedo², Moysés Nascimento², Marcia Flores da Silva Ferreira^{1*}, Adésio Ferreira¹

¹ Department of Agronomy, Federal University of Espírito Santo, Alegre, Espírito Santo, Brazil,

² Department of Statistics, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil

* marcia.ferreira@ufes.br



OPEN ACCESS

Citation: Canal GB, Barreto CAV, de Almeida FAN, Zaidan IR, do Couto DP, Azevedo CF, et al. (2023) Single and multi-trait genomic prediction for agronomic traits in *Euterpe edulis*. PLoS ONE 18(4): e0275407. <https://doi.org/10.1371/journal.pone.0275407>

Editor: Paulo Eduardo Teodoro, Federal University of Mato Grosso do Sul, BRAZIL

Received: September 16, 2022

Accepted: February 16, 2023

Published: April 7, 2023

Copyright: © 2023 Canal et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting Information](#) files.

Funding: We thank the Conselho Nacional de Pesquisa (CNPq, Brazil) (Researcher productivity fellowship AF and MFSF), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil) – Finance Code 001, and the Fundação de Amparo à Pesquisa do Espírito Santo (FAPES, Vitória – ES, Brazil) in partnership with VALE, for the financial support to this study. The

Abstract

Popularly known as juçaizeiro, *Euterpe edulis* has been gaining prominence in the fruit growing sector and has demanded the development of superior genetic materials. Since it is a native species and still little studied, the application of more sophisticated techniques can result in higher gains with less time. Until now, there are no studies that apply genomic prediction for this crop, especially in multi-trait analysis. In this sense, this study aimed to apply new methods and breeding techniques for the juçaizeiro, to optimize this breeding program through the application of genomic prediction. This data consisted of 275 juçaizeiro genotypes from a population of Rio Novo do Sul-ES, Brazil. The genomic prediction was performed using the multi-trait (G-BLUP MT) and single-trait (G-BLUP ST) models and the selection of superior genotypes was based on a selection index. Similar results for predictive ability were observed for both models. However, the G-BLUP ST model provided greater selection gains when compared to the G-BLUP MT. For this reason, the genomic estimated breeding values (GEBVs) from the G-BLUP ST, were used to select the six superior genotypes (UFES.A.RN.390, UFES.A.RN.386, UFES.A.RN.080, UFES.A.RN.383, UFES.S.RN.098, and UFES.S.RN.093). This was intended to provide superior genetic materials for the development of seedlings and implantation of productive orchards, which will meet the demands of the productive, industrial and consumer market.

Introduction

Euterpe edulis, a palm tree native to the Atlantic Forest, is distributed along the entire length of this biome and in the gallery forests of the Brazilian cerrado [1–3]. It belongs to the genus *Euterpe*, which has become popular in recent years through açai, processed pulp from the fruits of species of this genus [4]. The most exploited species is *Euterpe oleracea*, açai, which was recognized as a fruit tree of economic expression in 2008 [5].

As a crop, *Euterpe edulis* has become popular under the name of juçaizeiro. It is a perennial species that can exceed 15 meters in height, with a single stem [6] that does not regrow [7]—with extremely rare cases of tillers. Its inflorescences are infrafoliate, located in the transition

fundings had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

region between the heart of the palm and the stipe [8]. It is a monoecious species with flowering in clusters and its flowers distributed in a triad system, with two male flowers and one female with strong protandry [8, 9].

Jucaizeiro has shown great economic potential for use in the pulp processing industry [10–12]. The production of fruits of the species begins approximately six years after planting [4], and subsequently occurs recurrently once a year. Given its productive potential and wild state, there is a demand for superior genetic materials to use the species as a new crop. However, due to a series of intrinsic characteristics, such as the wild state, slow development, and difficulty in the practice of controlled pollination and propagation exclusively via seminal; the classic breeding practices may not be enough to lead to satisfactory genetic gains in the current short time.

Within breeding programs, the genome-wide selection (GWS) proposed by Meuwissen et al. (2001) [13], uses techniques such as G-BLUP (Genomic Best Linear Unbiased Prediction), and the use of knowledge of genomic relationship—to increase selective accuracy [14, 15]. The application of the GWS in a natural population with open-pollination, as the base population of improvement of the jucaizeiro species, in which the degree of relationship between the individuals is unknown, is scarce in the literature. However, several studies were conducted on forest species from open-pollinated populations—with knowledge of families [16–20] and support the assumption of increased accuracy with GWS application in a natural population without knowledge of relationship structure. This is due to the fact that, in these works, the use of the genomic relationship matrix (G) corrects the unrealistic priori that, in a given family, all individuals share the same genetic similarity with each other. In addition to the above, the accuracy of genomic prediction is generally higher, especially for breeding populations with superficial genealogy and disconnected families [21].

Genomic prediction, based on single traits, has become a widely applied procedure among breeders after reducing genotyping costs [22–24]—mainly using the G-BLUP method. However, this model may not satisfactorily reflect the interactive complexities between the analyzed traits, because it does not capitalize on the flow of information between traits, through available information on genetic co-variances [25]. This co-variances are generated through shared genetic influence (pleiotropy) and/or by the non-random association between the alleles (linkage disequilibrium), which, consequently, are responsible for generating complex relationships between quantitative characters [26]. Therefore, the multi-trait approach (G-BLUP MT) applied to GWS has been highlighted for being able to combine information and capture the effect of association between traits in order to predict genetic values more accurately [25, 27–29].

Currently, for the jucaizeiro, the techniques and approaches applied to its scarce existing breeding programs are simple. In general, the selection of the best genotypes is carried out in natural populations, based on data from a single year, and with a mass phenotypic selection technique. This is due to the absence of experimental fields with adult representatives of the species, without pre-established delineation and lack of knowledge of relationship between the individuals. In this sense, the selective bias can compromise the success of the species as a crop, by indicating low productivity genetic materials. Therefore, the present work aims to apply new methods and breeding techniques for jucaizeiro, to optimize the breeding programs through the application of genomic prediction. This study was carried out in a natural open-pollinated population containing 275 genotypes, with no knowledge of relationship between individuals; in order to select the genotypes with the greatest potential to serve the productive and industrial sector; and to compare the efficiency between multi-trait (G-BLUP MT) and single-trait (G-BLUP ST) models.

Materials and methods

Field experimental conduction

Experimental evaluations were carried out from 2018 to 2021 in the municipality of Rio Novo do Sul in the state of Espírito Santo, Brazil (Fig 1). The 275 genotypes evaluated were selected for good phytosanitary and physiological conditions and full reproductive age. The genotypes are located in a commercial plantation belonging to two companies that process juçaizeiro fruit pulp, Açai juçara® and Bonalotti®. The plantation was formed by a mixture of individuals that emerged by spontaneous development, through the action of dispersers, and the

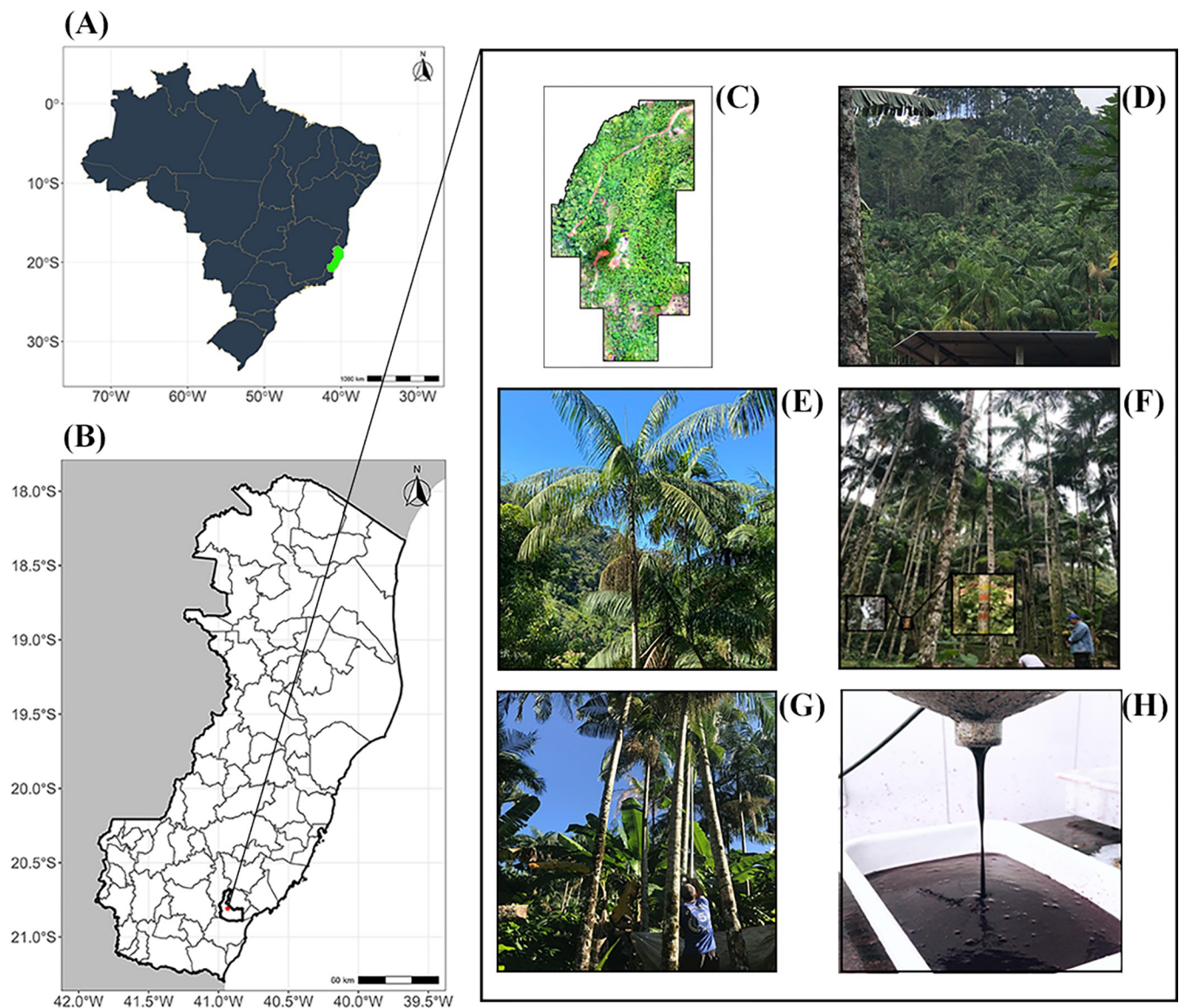


Fig 1. Geographical location of the experimental field. Geographical location of the experimental field where the *Euterpe edulis* genotypes are located. Map generated with R free environment software. (A) Delimitation of Brazil, in green, the geographic location of the state of Espírito Santo; (B) Graphic representation of the state of Espírito Santo, in red, the location of the municipality of Rio Novo do Sul; (C) Orthomosaic of the experimental area; (D) Image referring to a fragment of the experimental planting; (E) Juçaizeiro (*Euterpe edulis*) from experimental planting in productive phase; (F) Registration of genotypes identification in field; (G) Juçaizeiro (*Euterpe edulis*) fruits harvest; (H) Processed pulp of juçaizeiro (*Euterpe edulis*) fruits.

<https://doi.org/10.1371/journal.pone.0275407.g001>

enrichment by the owners by the sowing of seeds in the area. Commercial planting does not have pre-defined spacing and basically no productive management treatment, mowing is only done at harvest time. Consequently, the field evaluations followed a blank design, as the descent relationship between the genotypes was initially unknown.

Phenotyping of plants. Between 2018 and 2021, the number of bunches per plant (NB) was evaluated by visually counting the number of bunches with fruits of each matrix plant. In the years 2018, 2019 and 2021, the mass of fruits per bunch (MFB) (kg), rachis length (RL) (cm), equatorial diameter of fruits (EDF) (mm) and pulp yield (PY) (%) were evaluated. For the evaluation of MFB, in all evaluations, the same trained professional performed the determination of the harvest point of the bunches. The harvest of each genotype was carried out when the fruit maturation point reached the stage used for fruit processing in the industry. Each year, one bunch per plant was harvested, the fruits were separated from the rachillas, and they were weighed on a scale with an accuracy of 0.1 g. For RL, after separating the fruits, the length of the rachis of the inflorescence was measured with a tape measure.

A sample of fruits was taken from each genotype, packed in properly identified plastic bags and transported to the Plant Biometry laboratory at the Federal University of Esp rito Santo, where morphometric evaluations of fruits and seeds were carried out in a completely randomized design. For EDF, data were measured in millimeters (mm) obtained using a 6" digital caliper (Zaasprecision ), performed on five fruits individually, as recommended by Mar al et al. (2015) [30], who observed high repeatability values for these characteristics. The authors concluded that five measurements are necessary to carry out the measurement to reach coefficients of determination of 95%. Measurements above this amount increase costs and evaluative time, bringing little additional information to the works.

The PY was estimated by the following relationship:

$$PY = \frac{FFM - SFM}{FFM} * 100$$

In the equation, FFM is fruit fresh mass and SFM is seed fresh mass, measured by weighing four replicates of 25 fruits and seeds, using an analytical balance (0.0001g).

Genomic DNA extraction. Genomic DNA was obtained from leaf samples of the genotypes under study. Extraction was performed using the cetyltrimethylammonium bromide or CTAB method by Doyle (1990) [31] with modifications [32].

DNA concentrations and integrity were estimated using a NanodropTM 2000 spectrophotometer (Thermo Scientific). DNA quality was verified on 0.8% agarose gel. DNA genotypes prepared for genotyping using the DArTseqTM methodology were sent to the Service of Genetic Analysis for Agriculture (SAGA) in Mexico for high-throughput genotyping using the DArTseqTM technology

The genome representation of the 275 genotypes was obtained from the reduction of DNA complexity using two restriction enzymes, HpaII and MseI. The ends of the cleaved fragments were linked to a code adapter and a common adapter to identify each sample. The fragments were amplified using PCR reaction; subsequently, equimolar amounts of amplification products from each sample of the 96-well microtiter plate were pooled, purified and quantified, then sequenced on the Illumina Novaseq 6000 System platform. The sequences were analyzed using Dartsoft14, an automated genomics data analysis program and DArTdb, a laboratory management system, developed and patented by DArT Pvt. Ltd. (Australia), generating SNP marker data as described by Killian et al. (2012) [33] and Sansaloni et al. (2020) [34].

Quality control of molecular markers. The dataset of codominant markers of the SNP type was submitted to quality control analysis in the R [35]. We removed the markers with

Call Rate (CR) $\leq 90\%$ and Minor Allele Frequency (MAF) $\leq 5\%$. After quality control, the marker dataset reduced by 81.75% from 44,457 markers to 8,112.

Phenotypic data analysis. The phenotypic values were corrected for the effect of years in the R software, according to the correction proposed by Carvalho et al. (2020) [36]. The linear model used was:

$$y_{ik} = \mu + a_k + g_i + e_{ik}$$

where y_{ik} is the phenotypic value for the genotype i and the year k ; μ is the population mean; a_k is the fixed effect of the k^{th} year; g_i is the fixed effect of the i^{th} genotype; and e_{ik} is the random effect of residual, with $e \sim N(0, I\sigma_e^2)$. The empirical best linear unbiased estimates (eBLUES) was calculated for each trait individually and their values were obtained by $\hat{y}_i = \mu + \hat{a}_k + g_i$.

Genomic prediction (GP). Genomic prediction was performed using the multi-trait (G-BLUP MT) and single-trait (G-BLUP ST) models. For this the SOMMER package version 3.4 [37] was implemented in the R software [35]. The G-BLUP MT model for the prediction of the genomic estimated breeding values (GEBVs) of the individuals used was:

$$y = Xb + Zu + e$$

Where y is the vector of previously estimated eBLUES and structured as $y = [y_1 \ y_2 \ \dots \ y_n]'$; $y_1, y_2, \dots, y_n$ is the vector of observations for each characteristic; b is the vector of means of each characteristic structured as $b = [b_1 \ b_2 \ \dots \ b_n]'$ and with X incidence matrix. u is vector of individual additive genomic genetic values of each trait structured as $u = [u_1 \ u_2 \ \dots \ u_n]'$. Z is the incidence matrix, with the variance structure given by $u \sim N(0, G \otimes \Sigma_u)$, where G is the genomic relationship matrix between individuals for additive effects, Σ_u is the additive genetic covariance matrix and $\otimes$ denotes the Kronecker product; e is the random error vector with $e \sim N(0, I \otimes \Sigma_e)$ where Σ_e is the residual covariance matrix.

Covariance matrices can be written as:

$$\Sigma_u = \begin{bmatrix} \sigma_{u_1}^2 & \sigma_{u_{12}}^2 & \dots & \sigma_{u_{1n}}^2 \\ \sigma_{u_{21}}^2 & \sigma_{u_2}^2 & \dots & \sigma_{u_{2n}}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{u_{n1}}^2 & \sigma_{u_{n2}}^2 & \dots & \sigma_{u_n}^2 \end{bmatrix} \text{ and } \Sigma_e = \begin{bmatrix} \sigma_{e_1}^2 & \sigma_{e_{12}}^2 & \dots & \sigma_{e_{1n}}^2 \\ \sigma_{e_{21}}^2 & \sigma_{e_2}^2 & \dots & \sigma_{e_{2n}}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{e_{n1}}^2 & \sigma_{e_{n2}}^2 & \dots & \sigma_{e_n}^2 \end{bmatrix}$$

where $\sigma_{u_i}^2$ and $\sigma_{e_i}^2$ are, respectively, the additive and residual genetic variance associated with the i^{th} trait with $i = 1, \dots, n = 5$; $\sigma_{u_{ij}}^2$ and $\sigma_{e_{ij}}^2$ are, respectively, the additive and residual genetic covariance associated with the i^{th} and j^{th} traits with $j = 1, \dots, n = 5$ and $i \neq j$. The variance and covariance components were obtained via the restricted maximum likelihood method (REML). The additive genomic relationship matrix (G) was obtained as described by VanRaden (2008) [38] by the centralization of the matrix of markers:

$$G = \frac{W'W}{\sum_{i=1}^n 2p_i(1 - p_i)}$$

The parameterization for the incidence matrix W is presented below and is in accordance with the classical theory of quantitative genetics [39]:

$$W = \begin{cases} \text{If } MM, \text{ then } 2 - 2p_i \rightarrow 2q_i \\ \text{If } Mm, \text{ then } 1 - 2p_i \rightarrow q_i - p_i \\ \text{If } mm, \text{ then } 0 - 2p_i \rightarrow -2p_i \end{cases}$$

where, p_i and q_i are allele frequencies of M and m , respectively.

With the genetic values of the individuals for the NB and MFB traits, the fruit production per plant (FPP) for each genotype was estimated.

The narrow-sense heritability of the i^{th} trait was estimated following the equation below:

$$h_u^2 = \frac{\hat{\sigma}_{u_i}^2}{\hat{\sigma}_{u_i}^2 + \hat{\sigma}_{e_i}^2}$$

where $\hat{\sigma}_{u_i}^2$ and $\hat{\sigma}_{e_i}^2$ are the additive genetic variance and the residual variance of the i^{th} trait, respectively. The genetic correlation between the i^{th} and j^{th} traits was obtained through the following equation:

$$r_{u_{ij}} = \frac{\hat{\sigma}_{u_{ij}}^2}{\sqrt{\hat{\sigma}_{u_i}^2 \times \hat{\sigma}_{u_j}^2}}$$

where $\hat{\sigma}_{u_i}^2$ and $\hat{\sigma}_{u_j}^2$ are the additive genetic variance of the i^{th} and j^{th} trait, respectively. $\hat{\sigma}_{u_{ij}}^2$ is the additive genetic covariance between the i^{th} and j^{th} traits.

Predictive ability and cross-validation. The predictive ability ($r_{y,\hat{u}}$) and the standard error of $r_{y,\hat{u}}$, were estimated through the cross-validation procedure, randomly subdividing the population into 5 folds. Thus, 220 genotypes were used for training set and 55 genotypes were used for validation set. For each fold, the $r_{y,\hat{u}}$ was obtained by the correlation between the predicted GEBV's ($\hat{u}$) and the corrected phenotypic values (y).

Genetic selection based on selection index. With the GEBVs predicted by the G-BLUP ST method, we initially estimated the production of fruit production per plant (PFP) by multiplying the GEBVs between NB and MFB. The selection of the best genotypes was based on multiple variables. For this, the method developed by Mulamba and Mock (1978) [40] was used, which is based on the sum of the individual ranks of each characteristic, creating a global rank. For the selection, the parameters of NB, MFB, PFP, RL and PY in the positive direction and EDF in the negative direction were considered. Two processes were performed before rank summing. The first was to transform the EDF values into classes, in order to create four classes. Class I for small fruits (values below the first quartile of the distribution of the genetic values); class II for small/medium size fruits (genotypes with values between the first and second quartile of the distribution of the genetic values); class III for medium fruits (genotypes with values between the second and third quartile of the genetic values), and class IV for large fruits (genotypes with values above the third quartile of the genetic values). The second change was the normalization of the number of ranks for each trait, to avoid traits with fewer ranks having a greater influence on the selection process. The standardization followed the expression below:

$$r_{ij} = \frac{nr_j - p_{ij}}{nr_j - 1}$$

where r_{ij} is the standardized rank value for genotype i and characteristic j ; nr_j is the number of ranks of trait j and p_{ij} and is the rank of genotype i for the trait j .

In order to compare the selective efficiency between the G-BLUP ST and G-BLUP MT models, it was estimated the expected selection gain for the population of the next selection cycle and for the commercial seed donor population. For both methods and all characteristics, with the exception of PFP, estimates were obtained by:

$$SG_{ij} = h_{ij}^2 * SD_i$$

where, SG is the expected selection gain for characteristic i by model j ; h_{ij}^2 is the heritability of characteristic i estimated by model j and SD_i is the selection differential for characteristic i in model j estimated based on corrected phenotypic values.

The SG for PFP was estimated based on the product of the SG estimates of NB and MFB.

Comparison between methodologies. Cohen's Kappa coefficient [41] was used to analyze the agreement between the best selected individuals between the G-BLUP ST e G-BLUP MT models, for the commercial seed donor population (formed by the six best ranked individuals), and the next selective cycle population (formed by the 50 best ranked individuals). Cohen's Kappa coefficient is given by:

$$kappa = \frac{NAO - NAEC}{NOA - NAEC}$$

where NAO are the number of observed agreements, $NAEC$ is the number of expected agreements by chance, and NOA is the number of analyzed observations [42].

Results

Data description

A summary of the descriptive statistics including mean, standard deviation, and maximum and minimum values for the five agronomic traits evaluated in this work is shown in Table 1.

Table 1. Summary of the descriptive statistics. Amplitudes, mean, standard deviation (SD) and coefficient of variation (CV) for rachis length (RL), fruit mass per bunch (MFB), number of bunches (NB), equatorial fruit diameter (EDF) and pulp yield (PY) for the years 2018, 2019, 2020 and 2021.

Var.	Year	Minimum	Maximum	Average	SD	CV (%)
RL (cm)	2018	32	99	60.6	14.35	23.68
	2019	33	104	62.72	14.15	22.56
	2021	34.3	90.3	59.31	10.39	17.52
MFB (kg)	2018	0.4	8.3	3.05	1.57	51.47
	2019	0.5	8.7	3.45	1.49	43.19
	2021	0.3	8.7	3.68	1.87	50.81
NB (unit)	2018	1	7	4.24	1.35	31.84
	2019	0	8	3.28	1.36	41.46
	2020	0	7	3.69	1.03	27.91
	2021	0	7	2.18	1.36	62.38
EDF (mm)	2018	11.15	18.9	14.03	1.15	8.20
	2019	10.36	21.3	13.24	1.09	8.23
	2021	10	19.94	13.31	1.25	9.39
PY (%)	2018	4.57	50.81	30.62	6.5	21.23
	2019	16.75	51.52	35.96	5.37	14.93
	2021	12.55	46.63	29.76	6.06	20.36

<https://doi.org/10.1371/journal.pone.0275407.t001>

It is noted that in general, the population presented a small variation in the average phenotypic response between the years evaluated.

Genetic parameters, genetic correlation and phenotypic correlation

The heritability and the residual and additive genetics variance components were estimated for G-BLUP ST and G-BLUP MT models (Table 2). Considering the standard error, all these estimates were statistically equivalent for all traits in the evaluated models. Heritabilities ranged from 0.29 to 0.80 for NB (GBLUP-MT) and PY (GBLUP-ST), respectively (Table 2).

The genetic and phenotypic correlations between the traits are shown in Fig 2, which were calculated using the (co)variance estimates obtained by the G-BLUP MT model, so that the genetic correlation estimates are presented on the upper diagonal and, on the lower diagonal, estimates of phenotypic correlations are presented.

Comparing the phenotypic and genotypic correlations between the traits in Fig 2, it is observed that only two correlation estimates had changes in their directions (RL and NB; MFB and NB), with the change in correlation between MFB and NB being more pronounced. The results obtained for the phenotypic and genotypic correlations can be classified from low to moderate [43]. For phenotypic correlation, the highest values observed were 0.41 (RL and MFB), in the positive direction, and -0.17 (RL and PY), in the negative direction. While, for the genotypic correlations between traits, the highest values were 0.40 (RL and EDF) in the positive direction, and -0.37 (NB and MFB) in the negative direction (Fig 2). E os erros padrões para as correlações genéticas, foram de magnitudes mais elevadas quando comparado aos estimados para as correlações fenotípicas. Standard errors for genetic correlations ranged from 0.18 (PY and EDF) to 0.37 (MFB and NB) and for phenotypic correlations, standard errors ranged from 0.01 (PY and EDF) to 0.12 (PY and RL) (S1 Table).

Predictive ability

The predictive ability ($r_{y,\hat{u}}$) for the G-BLUP ST and G-BLUP MT models is shown in Table 3. For the G-BLUP ST, the $r_{y,\hat{u}}$ ranged from 0.21 for NB to 0.49 for PY, and for the G-BLUP MT model, the $r_{y,\hat{u}}$ ranged from 0.18 for NB to 0.48 for PY. In general, the values of $r_{y,\hat{u}}$ observed for both methodologies were close, as well as NB and PY maintained their positions with the lowest and highest $r_{y,\hat{u}}$, respectively. The results show that both models are similar to perform the prediction of genomic genetic values.

Matrix selection and expected genetic advancement

The Cohen's Kappa coefficient was calculated for the top 20% of individuals (50) with the highest rank positions based on the Mulamba and Mock index [40], aimed to form the

Table 2. Genetic parameters and variance components. Additive genetic variance (σ_a^2), residual variance (σ_e^2), narrow-sense heritability (h_u^2) and their respective standard errors considering single trait (ST) and multiple trait (MT) models in genomic data.

Model	Par.	RL (cm)	MFB (kg)	NB (unit)	EDF (mm)	PY (%)
Single trait	σ_a^2	59.55(±23.57)	1.04(±0.39)	0.32 (±0.16)	0.81(±0.23)	19.71(±4.54)
	σ_e^2	73.53(±14.12)	1.15(±0.23)	0.72 (±0.11)	0.41(±0.12)	4.96(±2.05)
	h_u^2	0.45 (±0.14)	0.48 (±0.14)	0.31 (±0.13)	0.67(±0.12)	0.80(±0.10)
Multi trait	σ_a^2	61.60(±24.05)	0.90(±0.36)	0.30(±0.16)	0.81(±0.23)	19.48(±4.54)
	σ_e^2	72.26(±14.27)	1.21(±0.23)	0.73(±0.11)	0.41(±0.12)	5.07(±2.06)
	h_u^2	0.46(±0.14)	0.43(±0.14)	0.29(±0.13)	0.66(±0.12)	0.79(±0.10)

Parameters (Par.); Rachis Length (RL); fruit mass per bunch (MFB); number of bunches (NB); equatorial diameter of fruit (EDF) and pulp yield (PY).

<https://doi.org/10.1371/journal.pone.0275407.t002>

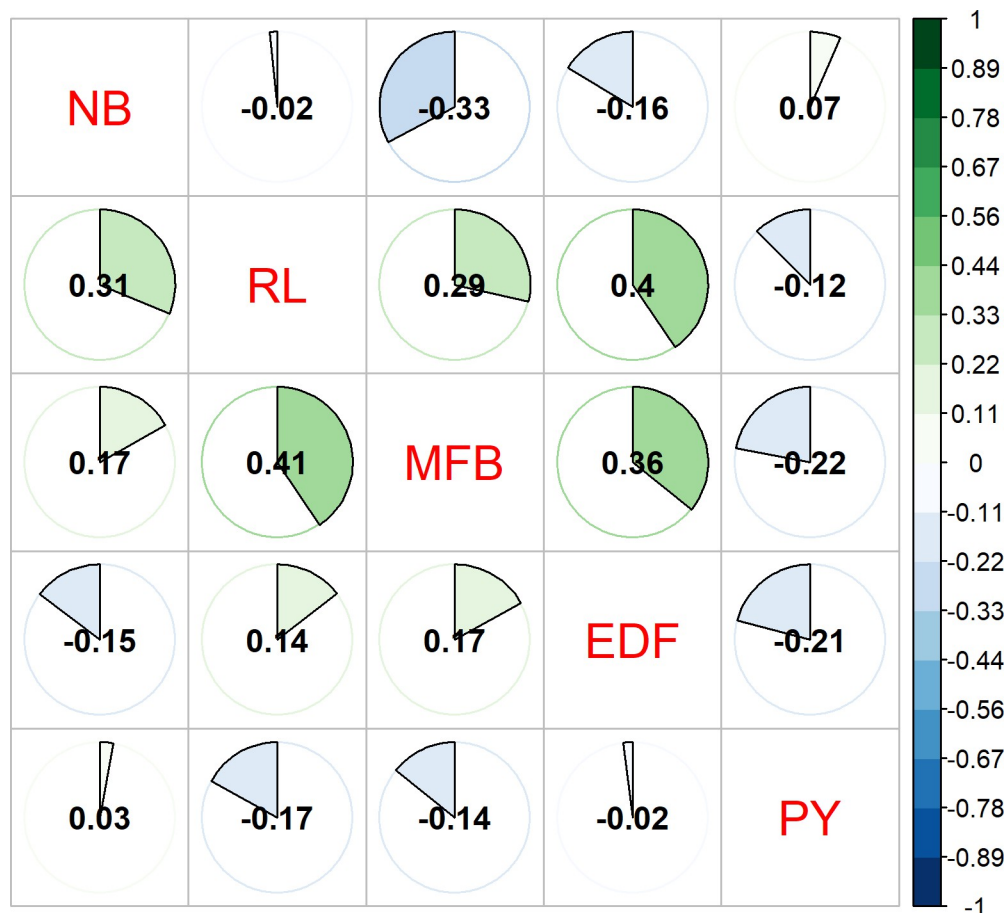


Fig 2. Genetic and phenotypic correlation. Genetic (upper diagonal) and phenotypic (lower diagonal) correlation between the traits RL (Rachis Length), EDF (Equatorial Fruit Diameter), MFB (Fruit Mass per Bunch), NB (Bunch Number) and PY (Pulp yield).

<https://doi.org/10.1371/journal.pone.0275407.g002>

population of the next selection cycle. The agreement between the G-BLUP ST and G-BLUP MT models was 0.68 and is classified as substantial (0.60–0.80) [44] (Fig 3). However, when we reduce the selected population to six individuals, to build the commercial seed donor population, the coefficient is reduced to an agreement of 0.50, considered moderate.

Even though the agreements of the Cohen's Kappa coefficient were classified from moderate (0.41–60) to substantial (0.61–80) [44], is observed in Fig 3, a great divergence in the ranking of individuals between the methods used. To support the comparison of the efficiency

Table 3. Predictive ability for G-BLUP ST and G-BLUP MT. Predictive ability ($r_{y,d}$) and their respective standard errors of the single-trait (G-BLUP ST), multi-trait (G-BLUP MT) models for RL (Rachis Length), EDF (Equatorial Fruit Diameter), MFB (Mass of Fruit per Bunch), NB (Number of Bunch) and PY (Pulp Yield).

Parameters	G-BLUP ST	G-BLUP MT
RL (cm)	0.29 (± 0.03)	0.22 (± 0.01)
MFB (kg)	0.27 (± 0.03)	0.26 (± 0.04)
NB (unit)	0.21 (± 0.06)	0.18 (± 0.05)
EDF (mm)	0.44 (± 0.02)	0.42 (± 0.01)
PY (%)	0.49 (± 0.06)	0.48 (± 0.07)

<https://doi.org/10.1371/journal.pone.0275407.t003>

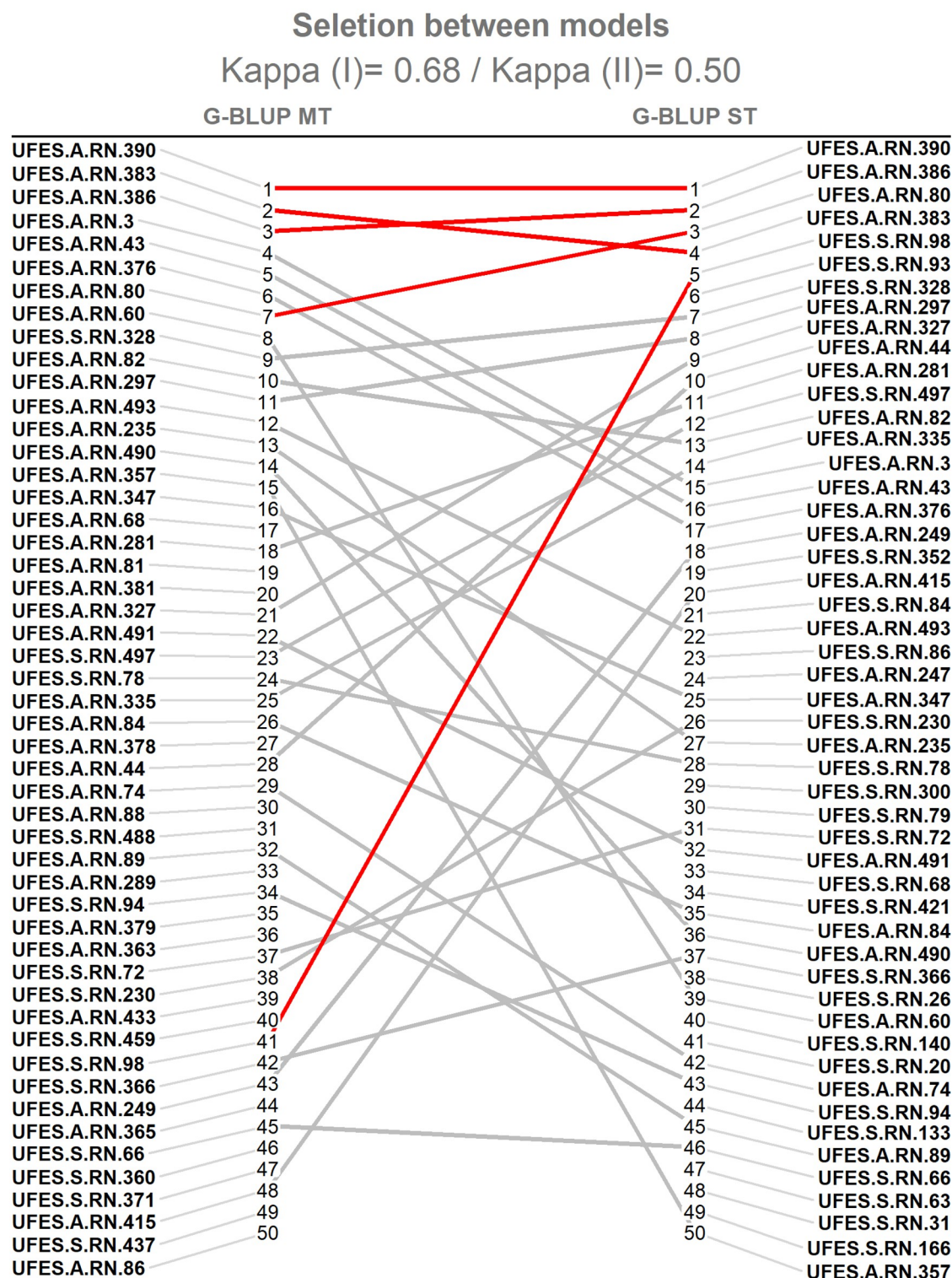


Fig 3. The agreement between genotypes selected by the G-BLUP MT and G-BLUP ST models. The Cohen's Kappa coefficient showing the agreement between genotypes selected by the G-BLUP MT and G-BLUP ST models. Kappa (I) - 20% of the best genotypes (50 genotypes) and Kappa (II)—the six best genotypes selected (represented by red lines).

<https://doi.org/10.1371/journal.pone.0275407.g003>

Table 4. Estimates of selection gains provided by the G-BLUP ST and G-BLUP MT models. Estimates of selection gains provided by the single-trait (G-BLUP ST) and multi-trait (G-BLUP MT) models for the selected genotypes (S.I.) for the population of the next selection cycle (50 genotypes), and the commercial seed donor population (six genotypes).

Trait	Population of the next selection cycle		Seed donor population	
	G-BLUP ST	G-BLUP MT	G-BLUP ST	G-BLUP MT
RL (cm)	31.25	31.64	30.58	32.58
MFB (kg)	2.06	1.62	2.40	1.76
NB (unit)	1.16	1.14	1.25	1.22
EDF (mm)	8.75	8.57	8.62	8.32
PY (%)	27.60	27.10	30.15	29.86
PFP (kg)	2.39	1.85	3.00	2.15

RL: Rachis Length, EDF: Equatorial Fruit Diameter, MFB: Mass of Fruit per Bunch, NB: Number of Bunch, and PY: Yield of Pulp; Production of Fruits per Plant (PFP).

<https://doi.org/10.1371/journal.pone.0275407.t004>

between the methods, the selection gains (SG) estimate was obtained through the heritability and phenotypic means of the selected individuals, and the results are shown in Table 4. In general, higher SG resulting from the G-BLUP ST model is observed for both populations.

The population mean, the phenotypic and genotypic behavior of the selected individuals, based on the information from the G-BLUP ST model and the field phenotypic, is shown in Fig 4. It is possible to observe that the phenotypic means of the six selected individuals, in the desired direction, have great performance in relation to the population average, being the improvement 31.13%, 51.05%, 98.40%, 11.56%, 15.57%, and 7.43% for NB, MFP, PFP, RL, PY and EDF, respectively.

The selection of the six best genotypes provides a change in phenotypic responses equivalent to 0.99 bunches, 1.70 kg, 10.43 kg, 7.08 cm, 5.09%, and -1.00 mm for the traits NB, MFP, PFP, RL, PY and EDF, respectively (Fig 4). With SG, it is expected an improvement on the averages in 0.41 bunches, -0.69 kg, 3.62 cm, 5.09%, and -0.47 mm, for NB, MFP, RL, PY, and EDF, respectively. For PFP, the genomic values were estimated by the product of the genomic values of MFB and NB, and the difference between the means of the selected and the general population was 3.71 kg (Fig 4).

Discussion

As there are few programs that are based on the improvement of *juçaizeiro*, there are also a few scientific works that are aimed at the application of selective techniques for individuals of this species. In this sense, the existing breeding programs are in the early stages of development, and thus, their techniques are basically based on classical methodologies with phenotypic information. In contrast to these processes, the present work innovates in being the first scientific study focused on genomic prediction of *Euterpe edulis*, with the objective of increasing selective accuracy and selection gains within the breeding cycle.

Genetic parameters, genetic correlation and phenotypic correlation

The h_u^2 found in the present work for the traits under study indicate that the morphometric traits of fruits (EDF and PY) have a greater potential for heredity and genetic control, when compared to the other productive parameters related to clusters (RL, MFB and NB). This behavior was also observed in *Euterpe oleracea* (açazeiro), a species of the same genus as *juçaizeiro* [45, 46].

It is expected that traits with higher h_u^2 , consequently, have a higher predictive ability [47]. The same was found by Legarra et al. (2008) [47], where the traits with the highest h_u^2 in the

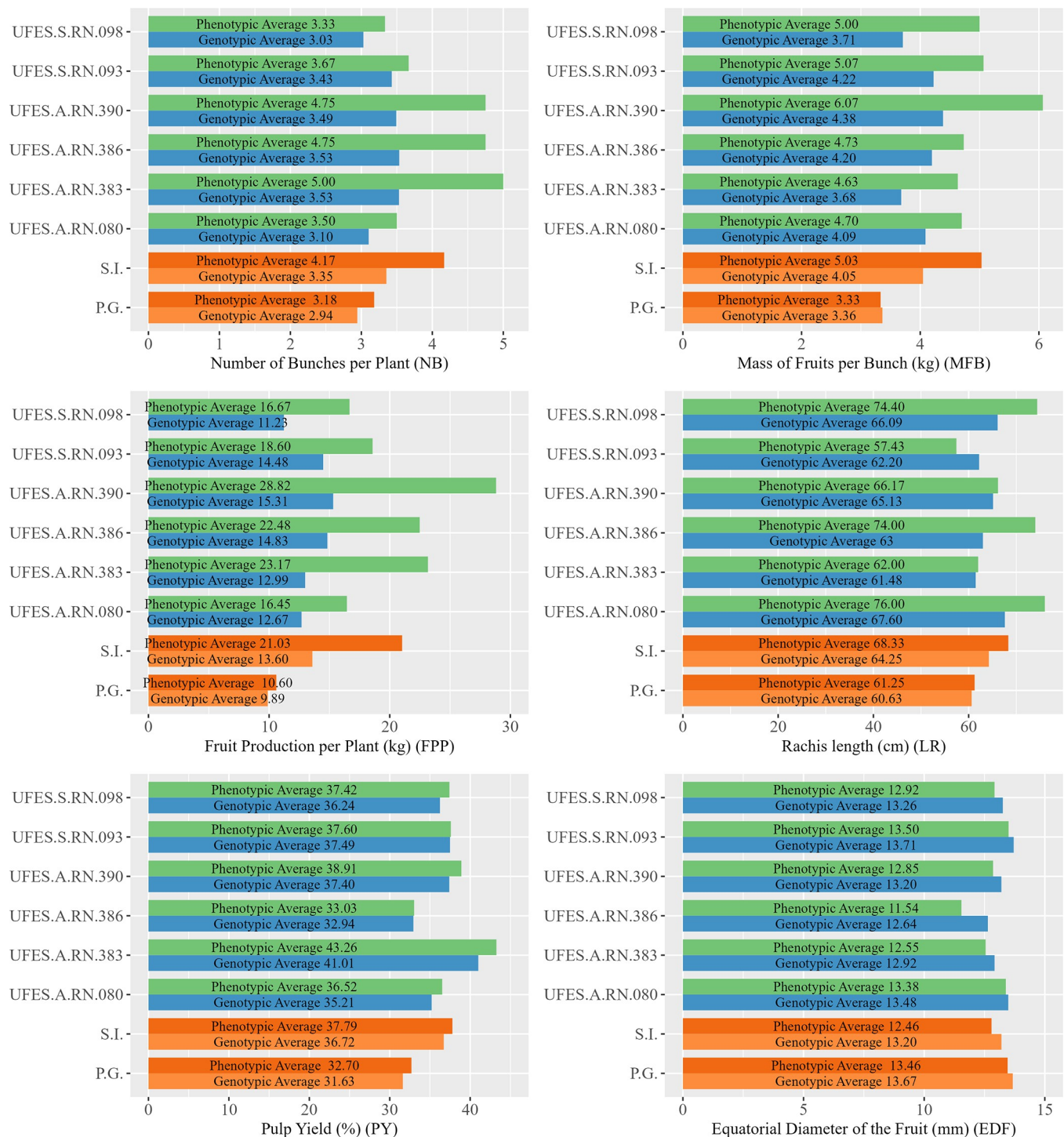


Fig 4. Phenotypic means observed in the field and GEBV's + mean of selected genotypes and general population. Phenotypic means observed in the field and GEBV's + mean of selected genotypes and general population. S.I.: commercial seed donor population, P.G.: general population.

<https://doi.org/10.1371/journal.pone.0275407.g004>

present study (EDF and PY) also had the higher predictive ability (Table 3). The higher values of h_u^2 for the traits of EDF and PY in relation to NB; MFB and RL can demonstrate the dynamism of the behavior of the h_u^2 as a function of the hereditary and environmental behavior. It

is known that for quantitative traits, several environmental factors can affect phenotypic behavior. In this sense, it can be assumed that the genetic control of NB, MFB and RL is reduced due to the exposure of these traits for a longer period of time to environmental effects and more variable conditions that are not controlled, for example, direction of insertion of the bunch into the matrix plant, exposure to inclement weather, feeding the fauna, among others. The development of fruits in the bunch will still occur, for example, in a fraction of the time of the infructescence development cycle, which will be located in the same region of the plant, with a similar influence on luminosity, temperature, and availability of nutrients. Therefore, genetic control can be expected to be greater.

Published scientific studies evaluating associations between *Euterpe edulis* traits are scarce in the literature [12, 30, 48]. However, these analyses are fundamental for understanding the interactive behavior between traits, revealing linear cause and effect responses that are essential for the agricultural development of the crop and for breeding programs, as it allows the determination of different types of practical strategies.

The knowledge of correlations in the genetic scope is important in breeding programs because they are heritable and can be used in indirect selection [49]. However, the evaluation of phenotypic behavior is also essential, since selection carried out indirectly based on genotypic information, and can lead to unwanted results in the field. This occurs when the phenotypic associations have the opposite direction to the genetic correlations considered for use in indirect selection.

The (co)variances estimated by the G-BLUP MT, made it possible to estimate the Pearson correlation for genetic and phenotypic effects between the traits RL, EDF, MFB, NB and PY (Fig 2), allowing a counterpoint between these estimates. In this sense, we noticed that practically all the associations preserved their influence on the behavioral response, with the exception of the NB-RL and NB-MFB pairs, which had an inversion of the sense of their estimates. In this condition, indirect selection is compromised by the fact that selection by genotypic responses may not be expressed in the desired sense in the environment. With this, we can conclude that the analyses of phenotypic and genotypic associations are complementary and make it possible to determine practical actions, using indirect processes to enable increased gains in breeding programs.

Evaluating biometric characteristics of juçaizeiro fruits, Marçal et al. (2015) [30] reported a positive genetic association between EDF and fresh seed mass (0.75). Knowing that the juçaizeiro fruit is mostly composed of seed, and the pulp makes up a small portion of it—increasing EDF can reduce the percentage of pulp yield. This fact is confirmed by the association observed in Fig 2, which shows inversely proportional associations between EDF and PY (-0.17 phenotypic; -0.21 genetic), that is, the increase in fruit size results in a reduction in the percentage of pulp per fruit, harmful to industrial processing.

The results observed by Oliveira et al. (2015) [48] for the genetic correlation in *Euterpe edulis*, are close to those observed in the present work between the characteristics NB, PY and RL. In both situations, the magnitudes of the associations can be classified from very weak to weak [43]. Oliveira et al. (2015) [48] reported associations of 0.20, -0.10 and -0.02 for NB-PY, PY-RL and NB-RL, respectively, while the present work obtained estimates of 0.07, -0.12 and -0.02, respectively. Corroborating these results, Farias Neto et al. (2016) [50] also found a weak correlation between NB-RL (-0.01) for *Euterpe oleracea*.

Regarding the correlation results observed in Fig 2, we can determine that the use of indirect selection among the evaluated characteristics would be inefficient. Therefore, the effect that the characteristics presented among themselves were mostly low, making the use of this practice unfeasible due to the small gains that could be obtained.

Predictive ability

In order to increase selection efficiency, choosing the best methodology for analyzing and predicting the genetic values of individuals is one of the fundamental aspects to be taken into account in a breeding program. In this study, we compare G-BLUP ST e G-BLUP MT models for the evaluation of genomic prediction in a base population of *juçaizeiro*. The existence of genetic correlations between the selection traits is the basis for the advantages presented by G-BLUP MT [51] and the absence of correlation could lead to equivalent or even superior results when using G-BLUP ST [51]. Therefore, we can conclude that, depending on the analyzed traits and the existence of correlations, it is expected that the use of the GBLUP MT model will increase the accuracy of the evaluation predictions [52].

However, the superiority of G-BLUP MT over G-BLUP ST for $r_{y,\hat{u}}$ was not observed in the present work. The results obtained by the G-BLUP MT model showed similar values of $r_{y,\hat{u}}$ when compared to the G-BLUP ST model, which may have been caused by the linear associations that varied from low to moderate intensity (≤ 0.42) (Fig 2) and, high standard of error associated with these estimates (S1 Table). As reported by Calus and Veerkamp (2011) [53], who observed an improvement in $r_{y,\hat{u}}$ when marker information was added to the model for the traits that had associations less than 0.50. In this sense, it is expected that the flow of genetic and phenotypic information provides higher $r_{y,\hat{u}}$ when compared to G-BLUP ST.

The slightly lower performance of G-BLUP MT for predictive abilities can be explained by the low genetic and residual correlations between traits, as reported by Runcie and Cheng (2019) [54], who attribute such results to the fact that low correlations can cause imprecision in the estimates of genetic and residual covariance parameters may result in reduced model performance. However, as the $r_{y,\hat{u}}$ were close, there is no evidence of such negative effects on the covariance estimates, deducing that the flow of information between the characteristics was not sufficient to improve the $r_{y,\hat{u}}$ of the model G-BLUP MT in relation to the G-BLUP ST, considering that marker information was used by both models.

The $r_{y,\hat{u}}$ for GBLUP-ST and GBLUP-MT performed statistically equivalent for most traits, with the exception of the RL, which showed $r_{y,\hat{u}}$ superior to G-BLUP-ST (0.29) compared to G-BLUP MT (0.22). The results diverge from patterns observed by other studies [27, 51, 52, 55–59], in which superiority is observed for the G-BLUP MT model. As it was designed to benefit from the existence of genetic correlations between traits, it is expected that the G-BLUP MT model will present better results than the G-BLUP ST [11, 53]. Due to the similarity of the results to the predictive ability of the models G-BLUP ST and G-BLUP MT, it is not possible, just for this parameter, to indicate the best model for the selection of *juçaizeiro* genotypes.

Genotype selection and expected genetic advancement

To build the next selective cycle breeding population, and in order to maintain genetic variability, a selection intensity of approximately 20% was defined. In this condition, the Kappa agreement observed among the approaches evaluated was classified as substantial (0.68), considering that its value was between 0.60 and 0.80 [44]. This similarity between the approaches can be explained by the fact that as the variables under study present low to moderate correlations [43], both methods ended up leading to similar results.

We should add that the high Kappa agreement between the methods is also associated with the conditions of the phenotypic data set used in the study, which has a low degree of imbalance. Thus, the predictive abilities of the approaches did not have a great impact on the selection of different genotypes between models. However, the influence on the ranking of the best individuals was strongly impacted, as shown in Fig 3.

When considering the six best genotypes, the selection of the best genotypes had greater divergence between the models, in which the Kappa value was 0.50, evidencing the need to determine the best model to carry out the selection process, aiming at greater future gains. In this way, the models are chosen according to their ability to predict more accurately the GEBVs of the analyzed individuals. However, in the present study, $r_{y,\hat{u}}$ was not sufficient to determine the best method, due to similar estimates observed.

Due to the similarities of the $r_{y,\hat{u}}$ and the great divergence in the ranking for the Kappa agreement, the choice of the best model for the selection of the juçaizeiro genotypes was based on the SG. In order to standardize the comparison between the models and eliminate the differences in the GEBVs associated with the form of prediction, the SG of each trait was estimated by the corrected phenotypic means and heritability from each model for the populations selected for the next selection cycle and seed donors. In this sense, the G-BLUP ST was chosen for the selection of juçaizeiro genotypes, because this model provided higher SG, in addition to being a simpler and more usual method for breeding programs.

The production of fruits per plant (PFP) is a feature that serves both the productive sector and the industrial sector, as it provides greater economic return to producers and increases the amount of raw material for industries that currently suffer from a shortage of this base product for their operations. In this sense, the selection differential amounts to an increase of 98.40%, which indicates a great advance for the culture due to the possibility of achieving great gains in productivity, resulting in an incentive to install of culture.

As the juçaizeiro is a wild and perennial species, which presents a series of characteristics intrinsic to its development, the practical activities of conventional breeding techniques become more difficult. In general, when evaluating the structure of a breeding program aimed at juçaizeiro, it is expected that the gains obtained per unit of time will be reduced. In this sense, the present study stands out for applying genomic prediction techniques in *Euterpe edulis*, a wild open-pollinated species, aiming to increase gains per unit of time.

Therefore, perhaps the biggest advantage brought by this methodology for a wild breeding population, as is the case of *Euterpe edulis*, is the use of the genomic matrix to bring knowledge of the genetic relationship between individuals for statistical analysis, which was previously unknown. Thus, the establishment of this information provides more accurate estimates and, consequently, greater reliability to the program. Furthermore, the use of genomic prediction may allow a reduction in the selection cycle by being able to predict the behavior of non-evaluated genotypes.

All the traits evaluated in the study influence the economic potential of the species, covering the interests of the productive sector and the industrial sector. Additionally, in function of the alterations observed by the selection of the six superior genotypes, we can conclude that the genomic prediction using the G-BLUP ST was efficient to provide alterations of the means in the desired directions. That is, obtaining an increase in the average phenotypic response of the selected genotypes for the characters of NB, MFB, PFP and PY, and a reduction for EDF.

Conclusion

Our results showed that the GBLUP-ST genomic prediction was more efficient in selecting the best genotypes, the selection provided substantial gains in the desired direction for multiple traits. Thus, the six selected genotypes (UFES.A.RN.390, UFES.A.RN.386, UFES.A.RN.080, UFES.A.RN.383, UFES.S.RN.098 e UFES.S.RN.093) can be used as commercial seed donor genotypes for the development of seedlings for the implantation of productive orchards, which will meet the demands of the productive sector and the consumer market.

Supporting information

S1 Table. Genetic (upper diagonal) and phenotypic (lower diagonal) correlation and respective standard errors between the traits RL (Rachis Length), EDF (Equatorial Fruit Diameter), MFB (Fruit Mass per Bunch), NB (Bunch Number) and PY (Pulp).
(DOCX)

S1 Data. Genomic estimated breeding values for single and multi- trait.
(XLSX)

S2 Data. Genomic relationship matrix between individuals for additive effects.
(XLSX)

S1 File. Scripts used.
(TXT)

Acknowledgments

We thank the companies, Açaí Juçara and Bonaloti, for supporting research development and we are also grateful to Pedro and Vicente Bortoloti, the owners of the managed area. We would like to thank all colleagues at the Biometrics and Genetics and Plant Improvement laboratories from the Universidade Federal do Espírito Santo, for helping to process and organize the samples.

Author Contributions

Conceptualization: Marcia Flores da Silva Ferreira, Adésio Ferreira.

Data curation: Guilherme Bravim Canal, Cynthia Aparecida Valiati Barreto, Camila Ferreira Azevedo, Moysés Nascimento, Adésio Ferreira.

Formal analysis: Guilherme Bravim Canal, Cynthia Aparecida Valiati Barreto, Camila Ferreira Azevedo, Moysés Nascimento, Adésio Ferreira.

Funding acquisition: Marcia Flores da Silva Ferreira, Adésio Ferreira.

Investigation: Guilherme Bravim Canal, Camila Ferreira Azevedo, Marcia Flores da Silva Ferreira, Adésio Ferreira.

Methodology: Guilherme Bravim Canal, Francine Alves Nogueira de Almeida, Iasmine Ramos Zaidan, Diego Pereira do Couto, Camila Ferreira Azevedo, Moysés Nascimento, Adésio Ferreira.

Project administration: Adésio Ferreira.

Resources: Guilherme Bravim Canal, Adésio Ferreira.

Supervision: Camila Ferreira Azevedo, Moysés Nascimento, Adésio Ferreira.

Writing – original draft: Guilherme Bravim Canal, Cynthia Aparecida Valiati Barreto.

Writing – review & editing: Guilherme Bravim Canal, Marcia Flores da Silva Ferreira, Adésio Ferreira.

References

1. Gaiotto FA, Grattapaglia D, Vencovsky R. Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart.). *J Hered*. 2003; 94: 399–406.

2. Henderson A, Galeano G, Bernal R. Field guide to the palms of the Americas. Princeton University Press; 2019.
3. Pereira AG, da Silva Ferreira MF, da Silveira TC, Soler-Guilhen JH, Canal GB, Alves LB, et al. Patterns of genetic diversity and structure of a threatened palm species (*Euterpe edulis* Arecaceae) from the Brazilian Atlantic Forest. *Heredity* (Edinb). 2022; 129: 161–168.
4. Carvalho LMJ, Esmerino AA, Carvalho JLV de. Jussai (*Euterpe edulis*): a review. *Food Sci Technol*. 2022;42.
5. dos Santos GM, Maia GA, de Sousa PHM, da Costa JMC, de Figueiredo RW, do Prado GM. Correlação entre atividade antioxidante e compostos bioativos de polpas comerciais de açaí (*Euterpe oleracea* Mart). *Arch Latinoam Nutr*. 2008; 58: 187–192.
6. Reitz R. Palmeiras In: REITZ R. Flora Ilus Catarinense Itajaí Herbário Barbosa Rodrigues. 1974.
7. Coelho GM, Santos AS, de Menezes IPP, Tarazi R, Souza FMO, Silva M das GCPC, et al. Genetic structure among morphotypes of the endangered Brazilian palm *Euterpe edulis* Mart (Arecaceae). *Ecol Evol*. 2020; 10: 6039–6048.
8. Lorenzi H, Souza HM de, Costa JTM, Cerqueira LSC de, Ferreira EJJ. Palmeiras brasileiras e exóticas cultivadas. 2004.
9. Wendt T, da Cruz DD, Demuner VG, Guilherme FAG, Boudet-Fernandes H. An evaluation of the species boundaries of two putative taxonomic entities of *Euterpe* (Arecaceae) based on reproductive and morphological features. *Flora-Morphology, Distrib Funct Ecol Plants*. 2011; 206: 144–150.
10. Maciel L de O, de Moura NF, Leonardi A. Cadeia produtiva do açaí juçara na região do litoral norte do Rio Grande do Sul. *Rev Teor e Evidência Econômica*. 2019; 25: 29–53.
11. Schulthess AW, Wang Y, Miedaner T, Wilde P, Reif JC, Zhao Y. Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. *Theor Appl Genet*. 2016; 129: 273–287. <https://doi.org/10.1007/s00122-015-2626-6> PMID: 26561306
12. Silva JZ da, Reis MS dos, Schulz M, Borges G da SC, Gonzaga LV, Costa ACO, et al. Fenologia Reprodutiva e Produção de Frutos em *Euterpe edulis* (Martius). *Ciência Florest*. 2018; 28: 295–309. <https://doi.org/10.5902/1980509831655>
13. Meuwissen THE, Hayes BJ, Goddard M. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 2001; 157: 1819–1829. <https://doi.org/10.1093/genetics/157.4.1819> PMID: 11290733
14. Alkimim ER, Caixeta ET, Sousa TV, Resende MDV, da Silva FL, Sakiyama NS, et al. Selective efficiency of genome-wide selection in *Coffea canephora* breeding. *Tree Genet Genomes*. 2020; 16: 1–11.
15. de Resende MD V, Ramalho MAP, Guilherme SR, de FB Abreu Â. Multigeneration index in the within-progenies bulk method for breeding of self-pollinated plants. *Crop Sci*. 2015; 55: 1202–1211.
16. El-Kassaby YA, Cappa EP, Liewlaksaneeyanawin C, Klápště J, Lstibůrek M. Breeding without Breeding: Is a Complete Pedigree Necessary for Efficient Breeding? Ingvarsson PK, editor. *PLoS One*. 2011; 6: e25737. <https://doi.org/10.1371/journal.pone.0025737> PMID: 21991342
17. Bush D, Thumma B. Characterising a *Eucalyptus cladocalyx* breeding population using SNP markers. *Tree Genet Genomes*. 2013; 9: 741–752. <https://doi.org/10.1007/s11295-012-0589-1>
18. Gamal El-Dien O, Ratcliffe B, Klápště J, Porth I, Chen C, El-Kassaby YA. Implementation of the realized genomic relationship matrix to open-pollinated white spruce family testing for disentangling additive from nonadditive genetic effects. *G3 Genes, Genomes, Genet*. 2016; 6: 743–753. <https://doi.org/10.1534/g3.115.025957> PMID: 26801647
19. El-Dien OG, Ratcliffe B, Klápště J, Porth I, Chen C, El-Kassaby YA. Multienvironment genomic variance decomposition analysis of open-pollinated Interior spruce (*Picea glauca* x *engelmannii*). *Mol Breed*. 2018; 38: 26. <https://doi.org/10.1007/s11032-018-0784-3> PMID: 29491726
20. Klápště J, Suontama M, Dungey HS, Telfer EJ, Graham NJ, Low CB, et al. Effect of Hidden Relatedness on Single-Step Genetic Evaluation in an Advanced Open-Pollinated Breeding Program. *J Hered*. 2018. <https://doi.org/10.1093/jhered/esy051> PMID: 30285150
21. Thavamanikumar S, Arnold RJ, Luo J, Thumma BR. Genomic studies reveal substantial dominant effects and improved genomic predictions in an open-pollinated breeding population of *Eucalyptus pellita*. *G3 Genes, Genomes, Genet*. 2020; 10: 3751–3763.
22. Sousa TV, Caixeta ET, Alkimim ER, Oliveira ACB, Pereira AA, Sakiyama NS, et al. Early Selection Enabled by the Implementation of Genomic Selection in *Coffea arabica* Breeding. *Front Plant Sci*. 2019;9. <https://doi.org/10.3389/fpls.2018.01934> PMID: 30671077
23. Massman JM, Gordillo A, Lorenzana RE, Bernardo R. Genomewide predictions from maize single-cross data. *Theor Appl Genet*. 2013; 126: 13–22. <https://doi.org/10.1007/s00122-012-1955-y> PMID: 22886355

24. Jarquin D, Howard R, Graef G, Lorenz A. Response Surface Analysis of Genomic Prediction Accuracy Values Using Quality Control Covariates in Soybean. *Evol Bioinforma*. 2019; 15: 117693431983130. <https://doi.org/10.1177/1176934319831307> PMID: 30872917
25. Gaire R, Arruda MP, Mohammadi M, Brown-Guedira G, Kolb FL, Rutkoski J. Multi-trait genomic selection can increase selection accuracy for deoxynivalenol accumulation resulting from fusarium head blight in wheat. *Plant Genome*. 2022. <https://doi.org/10.1002/tpg2.20188> PMID: 35043582
26. Lynch M, Walsh B. Genetics and analysis of quantitative traits. 1998.
27. Bhatta M, Gutierrez L, Cammarota L, Cardozo F, Germán S, Gómez-Guerrero B, et al. Multi-trait genomic prediction model increased the predictive ability for agronomic and malting quality traits in barley (*Hordeum vulgare* L.). *G3 Genes, Genomes, Genet*. 2020; 10: 1113–1124.
28. Lado B, Vázquez D, Quincke M, Silva P, Aguilar I, Gutiérrez L. Resource allocation optimization with multi-trait genomic prediction for bread wheat (*Triticum aestivum* L.) baking quality. *Theor Appl Genet*. 2018; 131: 2719–2731. <https://doi.org/10.1007/s00122-018-3186-3> PMID: 30232499
29. Sapkota S, Boatwright JL, Jordan K, Boyles R, Kresovich S. Multi-Trait Regressor Stacking Increased Genomic Prediction Accuracy of Sorghum Grain Composition. *Agronomy*. 2020; 10: 1221. <https://doi.org/10.3390/agronomy10091221>
30. Marçal T de S, Ferreira A, Oliveira WB dos S, Guilhen JHS, Ferreira MF da S. Correlações genéticas e análise de trilha para caracteres de fruto da palmeira juçara. *Rev Bras Frutic*. 2015; 37: 692–698.
31. Doyle JJ. Isolation of plant DNA from fresh tissue. *Focus (Madison)*. 1990; 12: 13–15.
32. Carvalho MS, Ferreira MF da S, Oliveira WB dos S, Marçal T de S, Guilhen JHS, Mengarda LHG, et al. Genetic diversity and population structure of *Euterpe edulis* by REML/BLUP analysis of fruit morphology and microsatellite markers. *Crop Breed Appl Biotechnol*. 2020;20.
33. Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, et al. Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms. 2012. pp. 67–89. https://doi.org/10.1007/978-1-61779-870-2_5 PMID: 22665276
34. Sansaloni C, Franco J, Santos B, Percival-Alwyn L, Singh S, Petroli C, et al. Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nat Commun*. 2020; 11: 4572. <https://doi.org/10.1038/s41467-020-18404-w> PMID: 32917907
35. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2021.
36. Carvalho HF, Galli G, Ferrão LFF, Nonato JVA, Padilha L, Perez MM, et al. The effect of bienniality on genomic prediction of yield in arabica coffee. *Euphytica*. 2020; 216: 101. <https://doi.org/10.1007/s10681-020-02641-7>
37. Covarrubias-Pazaran G. Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS One*. 2016; 11: e0156744. <https://doi.org/10.1371/journal.pone.0156744> PMID: 27271781
38. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci*. 2008; 91: 4414–4423. <https://doi.org/10.3168/jds.2007-0980> PMID: 18946147
39. Falconer DS, Mackay TFC. Introduction to quantitative genetics. Essex. UK Longman Gr. 1996
40. Mulamba NN, Mock JJ. Improvement of yield potential of the ETO blanco maize (*Zea mays* L.) population by breeding for plant traits [Mexico]. *Egypt J Genet Cytol*. 1978.
41. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960; 20: 37–46.
42. Resende MD V, Silva FF, Azevedo CF. Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão aleatória, seleção genômica, QTL-QWAS, estatística espacial e temporal, competição, sobrevivência. Suprema, Visconde do Rio Branco. 2014;1.
43. Evans G, Heath A, Lalljee M. Measuring left-right and libertarian-authoritarian values in the British electorate. *Br J Sociol*. 1996; 93–112.
44. Munoz SR, Bangdiwala SI. Interpretation of Kappa and B statistics measures of agreement. *J Appl Stat*. 1997; 24: 105–112. <https://doi.org/10.1080/02664769723918>
45. Farias Neto JT de, Resende MDV de, Oliveira M do SP de, Nogueira OL, Falcão PNB, Santos NSA dos. Estimativas de parâmetros genéticos e ganhos de seleção em progênies de polinização aberta de açaizeiro. *Rev Bras Frutic*. 2008; 30: 1051–1056.
46. Teixeira DHL, Oliveira M do SP de, Gonçalves FMA, Nunes JAR. Correlações genéticas e análise de trilha para componentes da produção de frutos de açaizeiro. *Rev Bras Frutic*. 2012; 34: 1135–1142.
47. Legarra A, Robert-Granié C, Manfredi E, Elsen J-M. Performance of Genomic Selection in Mice. *Genetics*. 2008; 180: 611–618. <https://doi.org/10.1534/genetics.108.088575> PMID: 18757934

48. Oliveira WB dos S, Ferreira A, Guilhen JHS, Marçal T de S, Ferreira MF da S, Senra JF de B. Path analysis and genetic diversity of *Euterpe edulis* Martius for vegetative and fruit traits. *Sci For*. 2015; 43: 303–311.
49. Cruz CD, Regazzi AJ, Carneiro PCS. Modelos biométricos aplicados a melhoramento. UFV, Viçosa. 2012.
50. Farias Neto JT de, Oliveira M do SP de, Yokomizo GKI. Ganho esperado na seleção de progênies de polinização aberta de *Euterpe oleracea* para produção de frutos. *Sci For*. 2016;44. <https://doi.org/10.18671/scifor.v44n109.23>
51. Jia Y, Jannink J-L. Multiple-Trait Genomic Selection Methods Increase Genetic Value Prediction Accuracy. *Genetics*. 2012; 192: 1513–1522. <https://doi.org/10.1534/genetics.112.144246> PMID: 23086217
52. Tsuruta S, Misztal I, Aguilar I, Lawlor TJ. Multiple-trait genomic evaluation of linear type traits using genomic and phenotypic data in US Holsteins. *J Dairy Sci*. 2011; 94: 4198–4204. <https://doi.org/10.3168/jds.2011-4256> PMID: 21787955
53. Calus MPL, Veerkamp RF. Accuracy of multi-trait genomic selection using different methods. *Genet Sel Evol*. 2011; 43: 1–14. <https://doi.org/10.1186/1297-9686-43-26> PMID: 21729282
54. Runcie D, Cheng H. Pitfalls and Remedies for Cross Validation with Multi-trait Genomic Prediction Methods. *G3 Genes|Genomes|Genetics*. 2019; 9: 3727–3741. <https://doi.org/10.1534/g3.119.400598> PMID: 31511297
55. Rutkoski J, Benson J, Jia Y, Brown-Guedira G, Jannink J-L, Sorrells M. Evaluation of Genomic Prediction Methods for Fusarium Head Blight Resistance in Wheat. *Plant Genome*. 2012; 5: 51–61. <https://doi.org/10.3835/plantgenome2012.02.0001>
56. Guo G, Zhao F, Wang Y, Zhang Y, Du L, Su G. Comparison of single-trait and multiple-trait genomic prediction models. *BMC Genet*. 2014; 15: 1–7.
57. Okeke UG, Akdemir D, Rabbi I, Kulakow P, Jannink J-L. Accuracies of univariate and multivariate genomic prediction models in African cassava. *Genet Sel Evol*. 2017; 49: 88. <https://doi.org/10.1186/s12711-017-0361-y> PMID: 29202685
58. Aroju SK, Cao M, Trollove M, Barrett BA, Inch C, Eady C, et al. Multi-Trait Genomic Prediction Improves Predictive Ability for Dry Matter Yield and Water-Soluble Carbohydrates in Perennial Ryegrass. *Front Plant Sci*. 2020;11. <https://doi.org/10.3389/fpls.2020.01197> PMID: 32849742
59. Tsai H-Y, Cericola F, Edriss V, Andersen JR, Orabi J, Jensen JD, et al. Use of multiple traits genomic prediction, genotype by environment interactions and spatial effect to improve prediction accuracy in yield data. Zhang A, editor. *PLoS One*. 2020; 15: e0232665. <https://doi.org/10.1371/journal.pone.0232665> PMID: 32401769