









# Selection of *Jatropha* genotypes for bioenergy purpose: an approach with multitrait, multiharvest and effective population size

Erina Vitório Rodrigues<sup>1</sup> , João Romero do Amaral Santos de Carvalho Rocha<sup>2</sup> , Rodrigo Silva Alves<sup>2</sup> , Paulo Eduardo Teodoro<sup>3,\*</sup> , Bruno Galvêas Laviola<sup>4</sup> , Marcos Deon Vilela de Resende<sup>5</sup> , Pedro Crescêncio Souza Carneiro<sup>2</sup> , Leonardo Lopes Bhering<sup>2</sup> 

1. Universidade de Brasília – Faculdade UnB Planaltina - Ciências da Vida e da Terra – Brasília (DF), Brazil.
2. Universidade Federal de Viçosa – Departamento de Biologia Geral – Laboratório de Biometria – Viçosa (MG), Brazil.
3. Universidade Federal de Mato Grosso do Sul – Departamento de Agronomia – Chapadão do Sul – (MS), Brazil.
4. Embrapa Agroenergia – Laboratório de Genética e Biotecnologia Brasília – Brasília (DF), Brazil.
5. Universidade Federal de Viçosa – Departamento de Estatística – Viçosa (MG), Brazil.

**ABSTRACT:** *Jatropha curcas* L. is a perennial plant with great potential for biodiesel production. Thus, the aim of this study was to select *J. curcas* genotypes for bioenergy purpose considering multitrait, multiharvest and effective population size. To this end, a data set with 70 *J. curcas* families, obtained through controlled crosses, evaluated in a randomized block design, with six replications, three plants per plot, spaced at 4 × 2 m was used. The following traits were evaluated: plant height in the 2014 and 2015 crop seasons; canopy projection in the row (2015 crop season); canopy projection between rows (2015 crop season); and grain yield (2015, 2016, and 2017 harvests). The mixed model methodology was used to estimate the variance components and to predict the genetic values. Superior genotypes were selected through Multitrait index based on factor analysis and genotype ideotype distance (FAI-BLUP index), considering the effective population size. According to this index, 98% of all additive genetic variability were summarized in three factors. By adopting the selection intensity that allows maintaining the effective population size equal to 30 (i.e., selection of 214 individuals), the predicted gains were 15.56, 21.75, 4.16, and 5.37% for the traits grain yield in the 2015 and 2016 harvests, canopy projection in the row, and canopy projection between rows in the 2015 crop season, respectively. Thus, the results suggest that the FAI-BLUP index can be successfully used in *J. curcas* breeding.

**Key words:** *Jatropha curcas* L. plant breeding, mixed models, genetic diversity.

Received:  
Feb. 5, 2020

Accepted:  
Jun. 8, 2020

Section Editor:  
Gabriel Constantino Blain

\*Corresponding author:  
eduteodoro@hotmail.com

## INTRODUCTION

*Jatropha curcas* L. (Euphorbiaceae) is a perennial plant native of Mexico and Central America that can be widely found in Latin America, Africa, India, and Southeast Asia (Pandey et al. 2012). This species can be used in soil recovery (Reubens et al. 2011), medicine (Pereira et al. 2018) and animal feeding (Makkar 2016) due to the identification of nontoxic varieties. *Jatropha curcas* also stands out for its great potential in biofuels production (Kumar et al. 2016; Laviola et al. 2017) due to its high yield and oil quality (Laviola et al. 2014). However, some technical-scientific improvements are required, especially regarding its genetics, for the use of this crop as renewable energy raw material.

Genetic variability in the population is fundamental for the selection of genotypes with high productive potential. Also, aggregating a set of traits of interest in the same genotype is a challenge to be overcome. Besides the high yield,



small plant is also essential because it facilitates *J. curcas* management. Therefore, *J. curcas* ideotype includes at least these two traits, which is a challenge for plant breeders. The selection for more than one trait involves genetic correlation and multicollinearity between them, which may lead to results misinterpretation (Rocha et al. 2018).

The use of accurate methods in the selection of perennial plants is crucial. These plants have long breeding cycles, and, therefore, several crops must be evaluated to assure the genetic superiority of an individual, which increases the costs. Also, the loss of individuals over the evaluation cycles generate unbalanced data. Thus, the restricted maximum likelihood/best linear unbiased prediction (REML/BLUP) procedure has been adopted for the genetic evaluation in perennial breeding, including *J. curcas* (Alves et al. 2018; Alves et al. 2019; Laviola et al. 2018; Peixoto et al. 2018).

Henderson and Quaas (1976) proposed the multiple-trait BLUP, that consists in the genetic evaluation of individuals for two or more traits simultaneously, considering the genetic and residual correlations between them (Alves et al. 2018). Besides considering several traits in the selection, this method also allows considering information of multiharvest that can be correlated with each other. This method has been effective in the selection of superior genotypes for multiple traits in *J. curcas* (Junqueira et al. 2016), *Eucalyptus* (Alves et al. 2018) e *Pennisetum purpureum* Schum. (Rocha et al. 2018).

Traditionally, researchers have used selection indexes in breeding programs aiming at selecting a set of information simultaneously through a linear combination to maximize gains in characteristics of economic interest (Cruz et al. 2012). However, this technique is conditioned to the inconvenience of multicollinearity between the characteristics evaluated, which can cause difficulties in interpreting the results and, consequently, erroneous conclusions (Prunier et al. 2015). To overcome this limitation, Rocha et al. (2018) proposed a multitrait index based on factor analysis and genotype-ideotype distance. This index considers the correlation structure obtained from the data and guides the selection of ideotype.

To date, no studies have reported the use of multitrait, multiharvest and effective population size in the genetic selection of *J. curcas*. Thus, the aim of this study was to select *J. curcas* genotypes for bioenergy purpose considering multitrait, multiharvest and effective population size.

## MATERIAL AND METHODS

### Experimental data

The experiment was carried out at the Embrapa Cerrados experimental area, in Planaltina, DF (lat. 15°35'30" S, long. 47°42'30" W, at 1007 m asl). The climate of the region is classified as Aw type (tropical with dry winter and rainy summer), according to the Köppen classification, with an average annual temperature of 22 °C, relative humidity of 73%, and average rainfall of 1100 mm. The soil is predominantly classified as red latosol, with high clay content.

Controlled crosses between 42 *J. curcas* genotypes that formed the Embrapa Agroenergia's Active Germplasm Bank in a unbalanced diallel (3 × 3) were performed, only 70 segregating families were obtained (Table 1), since some crosses did not result in viable seeds, which were evaluated in a randomized block design, with six replications, three plants per plot, spaced at 4 × 2 m, totalizing 1206 individuals. The following traits were evaluated: plant height in the 2014 and 2015 crop seasons; canopy projection in the row (2015 crop season); canopy projection between rows (2015 crop season); and grain yield (2015, 2016, and 2017 harvests).

### Statistical analyses

#### Statistic model

The REML/BLUP procedure was used to estimate the variance components and to predict the genetic values, according to Patterson and Thompson (1971) and Henderson (1975). The statistical model associated with the evaluation of individuals of full-sib families is given by Eq. 1:

$$y = Xr + Za + Wp + Tf + e \quad (1)$$

where  $y$  is the phenotypic data vector;  $r$  is the vector of the replication effects (assumed as fixed) added to the overall mean;  $a$  is the vector of the individual additive genetic effects (assumed as random);  $p$  is the vector of the plot effects (assumed as random);  $f$  is the vector of dominance effects of full-sib families (assumed as random); and  $e$  is the vector of errors (random).  $X$ ,  $Z$ ,  $W$ , and  $T$  represent the incidence matrices for  $r$ ,  $a$ ,  $p$  and  $f$ , respectively.

**Table 1.** Identification of *J. curcas* families and their respective parents.

Family	Male parent	Female parent	Family	Male parent	Female parent				
1	BAG 240-I-2	16	BAG 113-I-4	13	36	DIAL 7	29	DIAL 6	32
2	BAG 270-II-2	14	BAG 147-I-2	7	37	DIAL 6	28	DIAL 5	29
3	BAG 300-I-1	21	BAG 167-I-2	8	38	DIAL 5	28	DIAL 5	28
4	BAG 167-I-5	4	BAG 133-I-1	20	39	JTROPT 10	10	BAG 276-II-2	41
5	BAG 300-I-1	21	BAG 144-I-1	3	40	BAG 167-I-2	19	BAG 279-I-4	8
6	BAG 279-I-4	19	BAG 144-I-1	3	41	DIAL 6	26	JATROPT 4	29
7	BAG 167-I-5	4	BAG 263-I-1	9	42	DAG 167-I-5	1	BAG 199-II-5	4
8	TP 811-I-5	22	BAG 167-I-2	8	43	BAG 167-II-5	3	BAG 144-I-1	2
9	BAG 279-I-4	19	BAG 156-I-5	6	44	DIAL 7	28	DIAL 5	32
10	BAG 240-I-2	16	BAG 107-I-4	43	45	DIAL 6	29	DIAL 6	29
11	BAG 113-I-4	13	BAG 263-I-1	9	46	DIAL 7	32	DIAL 7	32
12	BAG 270-II-2	14	BAG 133-I-1	20	47	JATROPT 10	5	BAG 283-I-1	41
13	BAG 191-I-5	17	BAG 167-I-2	8	48	BAG 171-I-5	26	JATROPT 4	25
14	BAG 113-I-4	13	BAG 133-I-1	20	49	DIAL 7	22	TP 811-I-5	32
15	DIAL 2-V-3	23	JATOPT 8	39	50	BAG 167-I-5	3	BAG 144-I-1	4
16	DIAL 5	28	DIAL 6	29	51	DIAL 3	2	BAG 167-II-5	31
17	TP 811-I-5	22	DIAL 3	31	52	DIAL 16	29	DIAL 6	30
18	BAG 191-I-5	17	DIAL 3	31	53	DIAL 16	28	DIAL 5	30
19	BAG 191-I-5	17	DIAL 7	32	54	DIAL 6	27	JATROPT 6	29
20	BAG 191-I-5	17	DIAL 10	33	55	DIAL 16	5	DIAL 5	16
21	BAG 167-II-5	2	BAG 167-II-5	2	56	BAG 171-I-5	27	JATROPT 6	25
22	BAG 167-I-5	4	BAG 147-I-2	7	57	BAG 156-I-5	3	BAG 144-I-1	6
23	BAG 270-II-2	14	BAG 263-I-1	9	58	BAG 147-I-2	13	BAG 113-I-4	7
24	BAG 300-I-1	21	BAG 156-I-5	6	59	DIAL 7	12	BAG 241-II-4	32
25	BAG 266-II-4	11	BAG 156-I-5	6	60	BAG 199-II-5	1	BAG 199-II-5	1
26	BAG 266-II-4	11	BAG 144-I-1	3	61	BAG 167-I-5	4	BAG 167-I-5	4
27	DIAL 5	28	BAG 171-I-5	25	62	BAG 113-I-4	17	BAG 191-I-5	13
28	JATROPT 4	26	DIAL 16	30	63	BAG 156-I-5	6	BAG 156-I-5	6
29	TP 811-I-5	22	DIAL 10	33	64	DIAL 16	32	DIAL 7	30
30	BAG 283-I-1	5	BAG 167-II-5	2	65	DIAL 3	12	BAG 241-II-4	31
31	BAG 283-I-1	5	BAG 167-I-5	4	66	BAG 283-I-1	5	BAG 283-I-1	5
32	TP 811-I-5	22	BAG 107-I-4	43	67	BAG 167-I-2	11	BAG 266-II-4	8
33	TP 811-I-5	22	BAG 113-I-4	13	68	DIAL 16	27	INTER 6	30
34	BAG 240-I-2	16	BAG 167-I-2	8	69	JATROPT 10	2	BAG 167-II-5	41
35	BAG 144-I-1	3	BAG 144-I-1	3	70	DIAL 3	1	BAG 199-II-5	31

The significance of the random effects of the statistical model was evaluated by the likelihood ratio test (LRT) (Wilks 1938), using the  $\chi^2$  statistic, with one degree of freedom and 0.05 significance level. The narrow-sense heritability ( $h_a^2$ ) and the broad-sense heritability were obtained, respectively, by Eq. 2 and 3 (Resende 2016):

$$h_a^2 = \hat{\sigma}_a^2 / \hat{\sigma}_{Phen}^2 \quad (2)$$

$$h_g^2 = (\hat{\sigma}_a^2 + 4\hat{\sigma}_d^2) / \hat{\sigma}_{Phen}^2 \quad (3)$$

where  $\hat{\sigma}_a^2$  is the estimate of the additive genetic variance,  $\hat{\sigma}_{Phen}^2$  is the estimate of the phenotypic variance and  $\hat{\sigma}_d^2$  is the estimate of the dominance genetic variance.

The solutions for the genetic values of the nonparent individuals were based on Eq. 4 (Resende et al. 2014):

$$\hat{a}_{ijk} = \mu + 0.5(\hat{a}_p + \hat{a}_m) + \left[ \left( \frac{0.5 \hat{\sigma}_a^2}{0.5 \hat{\sigma}_a^2 + \hat{\sigma}_e^2} \right) (y_{ijk} - \hat{b}_k - 0.5(\hat{a}_p + \hat{a}_m) - \hat{f}_j - \hat{p}_{ik}) \right] \quad (4)$$

where  $\hat{a}_{ijk}$  is the estimate of the additive genetic value of individual  $i$  of family  $j$  in block  $k$ ;  $\mu$  is the overall mean of individuals;  $\hat{a}_p$  and  $\hat{a}_m$  are the predicted additive genetic values for parents  $p$  and  $m$ , respectively;  $\hat{\sigma}_e^2$  is the estimate of residual variance;  $y_{ijk}$  is the phenotypic observation of individual  $i$  of family  $j$  in block  $k$ ;  $\hat{b}_k$  is the estimate of the effect of block  $k$ ;  $\hat{f}_j$  is the predicted dominance genetic value for family  $j$ ; and  $\hat{p}_{ik}$  is the predicted plot effect (individual  $i$  in block  $k$ ). For individuals without phenotypic observation (not available), the additive genetic value of the individual was given by the mean additive effects of the parents, added to the overall mean of individuals.

## Effective population size

The effective population size ( $N_e$ ) is the ideal size of a population that could generate the same amount of inbreeding or variance in allele frequencies as that of the study population (Kimura and Crow 1963). The  $N_e$  estimator is given by Eq. 5 (Resende 2015):

$$N_{ep} = \frac{2N_f \bar{k}_f}{\frac{\sigma_f^2}{k_f} + \bar{k}_f + 1} \quad (5)$$

where  $N_f$  is the number of families selected;  $\bar{k}_f$  is the mean number of individuals selected per family; and  $\sigma_f^2$  is the variance of the number of individuals selected per family (Resende 2015).

## Effective family size

The estimator of the effective number of families selected ( $N_{ef}$ ) is given by Eq. 6 (Resende 2015):

$$N_{ef} = \left[ \frac{(\sum_{j=1}^n k_{fj})^2}{\sum_{j=1}^n k_{fj}^2} \right] \quad (6)$$

where  $k_{fj}$  is the number of individuals selected from family  $j$  (Robertson 1961).

## Genetic diversity of selected families

The genetic diversity between the families selected was calculated based on Eq. 7 (Wei and Lindgren 1996):

$$D = \left[ \frac{N_{ef}}{N_{f0}} \right] \quad (7)$$

where  $N_{f0}$  is the original number of families.

## Genetic selection based on multitrait and multiharvest

Superior genotypes were selected through multitrait index based on factor analysis and genotype-ideotype distance (FAI-BLUP index) (Rocha et al. 2018) which is based on the theoretical foundations of the structural equation models. This index associates exploratory (based on the data) and confirmatory (based on the ideotype) factor analysis techniques. It also considers the correlation structures obtained from the data and directs the selection, seeking for genotypes similar to those idealized by the breeders.

The ideotype used in the FAI-BLUP index for plant height in the 2014 and 2015 crop seasons was defined as the lowest value among the predicted genetic values for these crop seasons. Conversely, for canopy projection in the row and canopy projection between rows, evaluated in the 2015 crop seasons, and grain yield, assessed in the 2015, 2016, and 2017 harvests, the highest value among the predicted genetic values for the respective traits and years was considered. Only the traits with significant additive genetic effect (GCA) were considered for the index composition. The selection intensity was determined based on the effective population size ( $N_{ep}$ ) suitable for the formation of the breeding population ( $N_{ep} = 30$ ), according to Resende (2015). The variance components, genetic and nongenetic parameters and genetic values were obtained using the Selegen REML/BLUP software (Resende 2016). The genetic selection was made using the FAI-BLUP index routine, available by Rocha et al. (2018), using the R software (R Development Core Team 2019).

## RESULTS

### Additive and dominance genetic effects

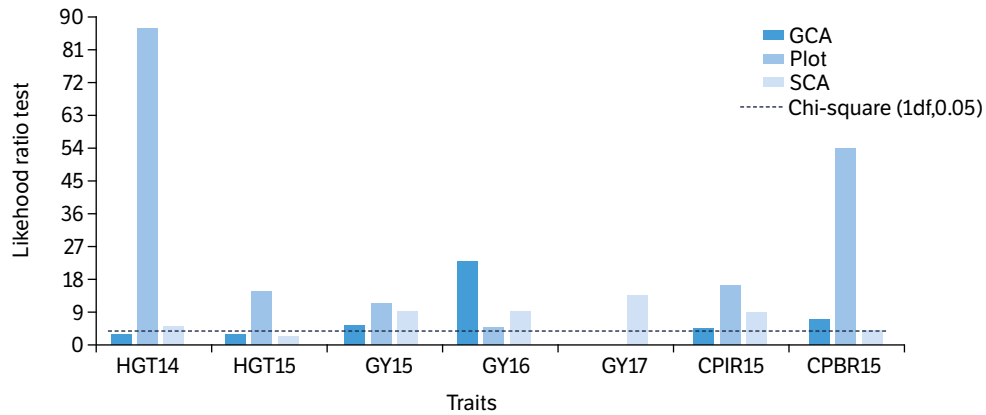
The significance of the random effects of the statistical model was evaluated using the likelihood ratio test (LRT), based on Resende (2016). This test indicated a nonsignificant effect of the general combining ability (GCA) for traits HGT14, HGT15, and GY17 (Table 2). A nonsignificant effect was observed for the specific combining ability (SCA) for the traits HGT15 and CPBR15 and plot effect for the trait GY17. For the other effects, the likelihood ratio test revealed a significant effect ( $p < 0.05$ ) (Table 1). Since no additive genetic variance was observed for traits HGT14, HGT15, and GY17, they were not used in the FAI-BLUP index composition.

Additive genetic variance estimates ( $\hat{\sigma}_a^2$ ) for traits HGT14, HGT15, and GY17 do not statistically differ from zero (Table 1). Consequently, the narrow-sense heritability values ( $h_a^2$ ) do not differ from zero. However, the estimates for the other traits are higher than zero, indicating the existence of genetic variability, which allows the selection of superior genotypes. The estimate of dominance genetic variance ( $\hat{\sigma}_d^2$ ) for trait HGT15 was the only one that did not differ from zero (Fig. 1).

**Table 2.** Likelihood ratio test and estimates of additive genetic variance ( $\sigma_a^2$ ), variance between plots ( $\sigma_{plot}^2$ ), dominance variance ( $\sigma_d^2$ ), residual variance ( $\sigma_{res}^2$ ), and phenotypic variance ( $\sigma_{phen}^2$ ); narrow- ( $h_a^2$ ) and broad-sense heritability ( $h_g^2$ ); and overall mean ( $\mu$ ).

Component	HGT14	HGT15	GY15	GY16	GY17	CPBR15	CPBR15
$\sigma_a^2$	0.0023 <sup>ns</sup>	0.0085 <sup>ns</sup>	538.9404*	47178.6899*	2693.7172 <sup>ns</sup>	0.0178*	0.0332*
$\sigma_{plot}^2$	0.0072*	0.0175*	581.2140*	13411.4361*	3094.9317 <sup>ns</sup>	0.0228*	0.0549*
$\sigma_d^2$	0.0014*	0.0039 <sup>ns</sup>	342.9830*	10656.2450*	48962.8530*	0.0112*	0.0112 <sup>ns</sup>
$\sigma_{res}^2$	0.0137	0.1129	3439.4748	129351.7451	599050.4355	0.1309	0.1473
$\sigma_{phen}^2$	0.0247	0.1429	4902.6123	200598.1161	653801.9374	0.1827	0.2467
$h_a^2$	0.0929	0.0598	0.1099	0.2352	0.0041	0.0974	0.1345
$h_g^2$	0.3264	0.1700	0.3898	0.4477	0.3037	0.3430	0.3162
$\mu$	0.5202	1.9236	92.3256	779.4103	891.3007	1.5361	1.5328

HGT14: plant height in the 2014 harvest; HGT15: plant height in the 2015 harvest; GY15: grain yield in the 2015 harvest; GY16: grain yield in the 2016 harvest; GY17: grain yield in the 2017 harvest; CPBR15: canopy projection in the row in the 2015 harvest; and CPBR15 canopy projection between rows in the 2015 harvest. ns and \*: non-significant and significant at 5% probability by the chi-square test was used with one degree of freedom.



**Figure 1.** Likelihood ratio test (bar plots). HGT14: plant height in the 2014 harvest; HGT15: plant height in the 2015 harvest; GY15: grain yield in the 2015 harvest; GY16: grain yield in the 2016 harvest; GY17: grain yield in the 2017 harvest; CPIR15: canopy projection in the row in the 2015 crop; and CPBR15 canopy projection between rows in the 2015 harvest. All bars above the dashed red line indicate a significant effect, considering the chi-square test, with one degree of freedom, at a 0.05 significance level 0.05 ( $\chi^2_{1,0.05}=3.84$ ).

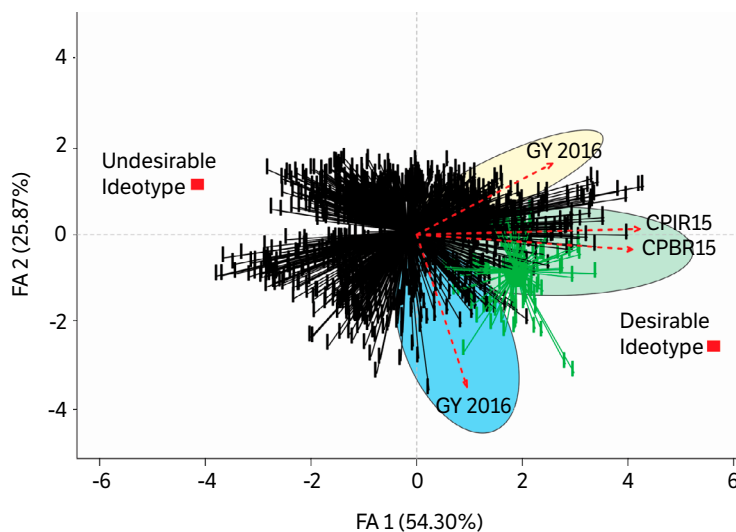
Narrow-sense ( $h_a^2$ ) and broad-sense ( $h_g^2$ ) heritabilities estimates increased for GY from the 2015 harvest to the 2016 harvest (132.21 and 14.85%, respectively). The  $h_a^2$  estimate of CPBR15 was higher than that of CPIR15. Conversely, the opposite occurred for the  $h_g^2$  estimates. Finally, the overall mean for GY increased over the harvests.

### Effective population and family sizes and genetic diversity of selected families

The 214 individuals selected belong to 28 out of the 70 families evaluated (Table S1). The families selected belong to a gene pool that includes 22 out of 33 parents evaluated. Thus, the effective number of families selected ( $N_{ef}$ ) was 16.38. Regarding the genetic diversity of these families, the estimated coefficient of genetic diversity (D) was 0.23.

### Individual genetic selection based on multitrait and multiharvest information (FAI-BLUP index)

The FAI-BLUP index was used to select superior genotypes for bioenergy purpose, considering multitrait and multiharvest information, simultaneously. The 1206 individuals (green and black dashes) ranked are shown in Fig. 2. Figure 2 shows



**Figure 2.** Two-dimensional plot of the selected *Jatropha curcas* genotypes (green dashes) and the desirable and undesirable ideotypes (red squares). The variables clustered in the same colored oval represent one factor according to the FAI-BLUP index.

that the 214 best genotypes (green dashes) were closer to the desirable ideotype. According to the FAI-BLUP index, 98% of all additive genetic variability were summarized in three factors. Factor 1 clustered the traits canopy projection in the row and canopy projection between rows in the 2015 crop season, explaining 54% of the variability. Factor 2 clustered the trait grain yield in the 2015 harvest, explaining 26% of the total variability. In its turn, factor 3 clustered the trait grain yield in the 2016 harvest, explaining 18% of the total variability. These results indicate a high correlation between traits of the first factor and the presence of genotypes  $\times$  harvests interaction for grain yield.

By adopting the selection intensity that allows maintaining the  $N_{ep} = 30$  (i.e., selection of 214 individuals), the predicted gains were 15.56, 21.75, 4.16, and 5.37% for the traits grain yield in the 2015 and 2016 harvests, canopy projection in the row, and canopy projection between rows in the 2015 crop season, respectively.

## DISCUSSION

The GCA is fundamental to identifying and selecting superior genotypes and quantifying the breeding potential of the studied population. Consequently, it allows the direct selection, which aims at increasing the frequency of favorable alleles (Resende 2015). The GCA effect was prevalent to the SCA effect only for traits GY16 and CPBR15 (Table 1). Results indicate that these traits can be improved by applying simple selection techniques, and the strategy to be used is intrapopulation recurrent selection, highlighting the additive genetic effects.

Conversely, dominance effects (SCA) were prevalent in the genetic control of HGT14, GY15, GY17, and CPIR15, suggesting differences in the genetic compositions of the parents (Table 2). Biabani et al. (2012) observed higher SCA for plant height when compared with GCA in diallel crosses with *J. curcas*. Results suggest that the methods must prioritize the capitalization of heterosis (genetic dominance and divergence) to improve the population considering these traits, and the breeding strategy should be recurrent interpopulation selection to improve the best hybrids.

Narrow-sense individual heritability ( $h_a^2$ ) quantifies the relative contribution of additive genes in the expression of the trait (Laviola et al. 2012). All estimates obtained can be considered low, according to the classification proposed by Resende (2002). Similar magnitudes values were reported by Laviola et al. (2010) for grain yield trait measured at 12 months in 110 *Jatropha* accessions. Spinelli et al. (2015) also found lower  $h_a^2$  estimates for this trait in 16 *Jatropha* half-sib families in the second, third, and fourth years after planting. In general, the increase in the  $h_a^2$  and  $h_g^2$  estimates of the 2015 harvest for the 2016 harvest indicates that the GY16 is the most suitable trait for selecting the most productive *Jatropha* genotypes.

The FAI-BLUP index was established based on the traits that showed GCA values statistically higher than zero, i.e., there were genetic variability. The first two factors explained 80% of the total variability among individuals (Fig. 2). The clustering of CPBR15 and CPIR15 traits in factor 1 indicates a high correlation between them (Fig. 2), as already reported by Laviola et al. (2010; 2012). The fact that the traits GY15 and GY16 are allocated in different factors demonstrates the genotypes  $\times$  harvests interaction for grain yield, i.e., the genotypes showed different yields in these years. These results are similar to those obtained by Laviola et al. (2013), Teodoro et al. (2016), and Alves et al. (2018), who verified low repeatability among the performance of *Jatropha* genotypes for grain yield in the early years.

The clustering of GY15 in factor 2 (26% of the total variability) can be explained by the fact that, in the early years, some *Jatropha* genotypes either produce very few or no grains at all, generating a great phenotypic variability for this trait. However, the phenotypic variability of this trait was predominantly composed of plot variance (Fig. 1). In the 2016 harvest, most of the genotypes produced grains, reducing plot variance and increasing genetic variance (Table 1).

The selection index used in this study allowed positive and high gains with selection for the traits GY15, GY16, CPBR15, and CPIR15. According to Rocha et al. (2018), the FAI-BLUP index has several advantages over the classic index proposed by Smith (1936) and Hazel (1943), since the former is based on the theory of the structural equation model. Moreover, the FAI-BLUP index does not assign weights and is free from multicollinearity. The estimates of selection gains obtained in this study are aligned with those of other works (Bhering et al. 2013; Spinelli et al. 2015) that used indices to select *J. curcas* genotypes to produce bioenergy.

It should be noted that the FAI-BLUP index was successfully used by Silva et al. (2018) to assist the selection of sorghum hybrids that simultaneously meet favorable traits for the production of second-generation bioethanol. Oliveira et al. (2019) selected five sorghum hybrids with higher potential for energy cogeneration using the FAI-BLUP index. Woyann et al. (2019) used the FAI-BLUP index in recombinant soybean inbred lines to identify those that best related to the ideotype for the biodiesel production.

It should be highlighted that the gains predicted with the selection in this study were obtained with restriction in the effective population size ( $N_{ep} = 30$ ) and, for that, 214 individuals were selected. An adequate effective population size is of paramount importance for the success in new cycles of recombination and selection in any breeding program because the effective population size must be maintained at a reliable level in each generation to avoid losing favorable alleles (Resende 2002).

## CONCLUSION

This study applied the mixed model methodology and the FAI-BLUP index for genetic selection, with restriction in the effective population size, of *J. curcas* genotypes. The results suggest that this procedure can be successfully used in *J. curcas* breeding.

## FUNDERS

Conselho Nacional de Desenvolvimento Científico e Tecnológico  
[<http://doi.org/10.13039/501100003593>]

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior  
[<http://doi.org/10.13039/501100002322>]

## AUTHOR'S CONTRIBUTION

Conceptualization, Rodrigues E. V. and Teodoro P. E.; Methodology, Rodrigues E. V., Teodoro P. E., Rocha J. R. A. S. C. and Alves R. S.; Writing – Original Draft, Rodrigues E. V., Teodoro P. E., Rocha J. R. A. S. C. and Alves R. S.; Writing – Review and Editing, Bhering L. L.; Funding Acquisition, Laviola B. G.; Resources, Laviola B. G.; Supervision, Resende M. D. V., Carneiro P. C. S. and Bhering L. L.

## REFERENCES

- Alves, R. S., Teodoro, P. E., Peixoto, L. A., Rocha, J. R. A. S. C., Silva, L. A., Laviola, B. G., Resende, M. D. V. and Bhering, L. L. (2019). Multiple-trait BLUP in longitudinal data analysis on *Jatropha curcas* breeding for bioenergy. *Industrial Crops and Products*, 130, 558-561. <https://doi.org/10.1016/j.indcrop.2018.12.019>
- Alves, R. S., Peixoto, L. A., Teodoro, P. E., Silva, L. A., Rodrigues, E. V., Resende, M. D. V., Laviola, B. G. and Bhering, L. L. (2018). Selection of *Jatropha curcas* families based on temporal stability and adaptability of genetic values. *Industrial Crops and Products*, 119, 290-293. <https://doi.org/10.1016/j.indcrop.2018.04.029>
- Bhering, L. L., Barrera, C. F., Ortega, D., Laviola, B. G., Alves, A. A., Rosado, T. B. and Cruz, C. D. (2013). Differential response of *Jatropha* genotypes to different selection methods indicates that combined selection is more suited than other methods for rapid improvement of the species. *Industrial Crops and Products*, 41, 260-265. <https://doi.org/10.1016/j.indcrop.2012.04.026>



- Biabani, A., Rafii, M. Y., Saleh, G., Shabanmofrad, M. and Latif, M. A. (2012) Combining ability analysis and evaluation of heterosis in *Jatropha curcas* (L.). *Australian Journal of Crop Science*, 6, 1030-1036.
- Cruz, C. D., Regazzi, A. J. and Carneiro, P. C. S. (2012). Modelos biométricos aplicados ao melhoramento genético. Viçosa: Suprema.
- Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics*, 28, 476-490.
- Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31, 423-447. <https://doi.org/10.2307/2529430>
- Henderson, C. R. and Quaas, R. L. (1976). Multiple trait evaluation using relatives' records. *Journal of Animal Science*, 43, 1188-1197. <https://doi.org/10.2527/jas1976.4361188x>
- Junqueira, V. S., Peixoto, L. A., Laviola, B. G., Bhering, L. L., Mendonça, S., Costa, T. S. A. and Antoniassi, R. (2016). Correction: Bayesian multi-trait analysis reveals a useful tool to increase oil concentration and to decrease toxicity in *Jatropha curcas* L. *PLoS ONE*, 11, e0161046. <https://doi.org/10.1371/journal.pone.0161046>
- Kimura, M. and Crow, J. F. (1963). The measurement of effective population number. *Evolution*, 17, 279-288. <https://doi.org/10.2307/2406157>
- Kumar, P., Srivastava, V. C. and Jha, M. K. (2016). *Jatropha curcas* phytotomy and applications: Development as a potential biofuel plant through biotechnological advancements. *Renewable and Sustainable Energy Reviews*, 59, 818-838. <https://doi.org/10.1016/j.rser.2015.12.358>
- Laviola, B. G., Rosado, T. B., Bhering, L. L., Kobayashi, A. K. and Resende M. D. V. (2010). Genetic parameters and variability in physic nut accessions during early developmental stages. *Pesquisa Agropecuária Brasileira*, 45, 1117-1123. <https://doi.org/10.1590/S0100-204X2010001000010>
- Laviola, B. G., Alves, A. A., Gurgel, F. L., Rosado, T. B., Rocha, R. B. and Albrecht, J. C. (2012). Estimates of genetic parameters for physic nut traits based in the germplasm two years evaluation. *Ciência Rural*, 42, 429-435. <https://doi.org/10.1590/S0103-84782012000300008>
- Laviola, B. G., Oliveira, A. M. C., Bhering, L. L., Alves, A. A., Rocha, R. B., Gomes, B. E. L. and Cruz, C. D. (2013). Estimates of repeatability coefficients and selection gains in *Jatropha* indicate that higher cumulative genetic gains can be obtained by relaxing the degree of certainty in predicting the best families. *Industrial Crops and Products*, 51, 70-76. <https://doi.org/10.1016/j.indcrop.2013.08.016>
- Laviola, B. G., Silva, S. D. A., Juhász, A. C. P., Rocha, R. B., Oliveira, R. J. P., Albrecht, J. C., Alves, A. A. and Rosado, T. B. (2014). Desempenho agrônomo e ganho genético pela seleção de pinhão-mansão em três regiões do Brasil. *Pesquisa Agropecuária Brasileira*, 49, 356-363. <https://doi.org/10.1590/S0100-204X2014000500005>
- Laviola, B. G., Rodrigues, E. V., Teodoro, P. E., Peixoto, L. A. and Bhering, L. L. (2017). Biometric and biotechnology strategies in *Jatropha* genetic breeding for biodiesel production. *Renewable and Sustainable Energy Reviews*, 76, 894-904. <https://doi.org/10.1016/j.rser.2017.03.116>
- Laviola, B. G., Teodoro, P. E., Peixoto, L. A. and Bhering, L. L. (2018). Parental selection in diallel crosses of *Jatropha curcas* using mixed models. *Acta Scientiarum. Agronomy*, 40, e35008. <https://doi.org/10.4025/actasciagron.v40i1.35008>
- Makkar, H. P. S. (2016). State-of-the-art on detoxification of *Jatropha curcas* products aimed for use as animal and fish feed: A review. *Animal Feed Science and Technology*, 222, 87-99. <https://doi.org/10.1016/j.anifeedsci.2016.09.013>
- Oliveira, I. C. M., Marçal, T. S., Bernardino, K. C., Ribeiro, P. C. O., Parrella, R. A. C., Carneiro, P. C. S., Schaffert, R. E. and Carneiro, J. E. S. (2019). Combining ability of biomass sorghum lines for agroindustrial characters and multitrait selection of photosensitive hybrids for energy cogeneration. *Crop Science*, 59, 1554-1566. <https://doi.org/10.2135/cropsci2018.11.0693>
- Pandey, V. C., Singh, K., Singh, J. S., Kumar, A., Singh, B. and Singh, R. P. (2012). *Jatropha curcas*: A potential biofuel plant for sustainable environmental development. *Renewable and Sustainable Energy Reviews*, 16, 2870-2883. <https://doi.org/10.1016/j.rser.2012.02.004>

- Patterson, H. D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58, 545-554. <https://doi.org/10.1093/biomet/58.3.545>
- Peixoto, L. A., Teodoro, P. E., Silva, L. A., Rodrigues, E. V., Laviola, B. G. and Bhering, L. L. (2018). *Jatropha* half-sib family selection with high adaptability and genotypic stability. *PLoS ONE*, 13, e0199880. <https://doi.org/10.1371/journal.pone.0199880>
- Pereira, I. R., D'Abadia, P. L., Prado, A. D. L., Matos, F. S., Nabout, J. C., Gonçalves, P. J. and Almeida, L. M. (2018). Trends and gaps in the global scientific literature about *Jatropha curcas* L. (Euphorbiaceae), a tropical plant of economic importance. *Semina: Ciências Agrárias*, 39, 7-18. <https://doi.org/10.5433/1679-0359.2018v39n1p7>
- Prunier, J. G., Colyn, M., Legendre, X., Nimon, K. F. and Flamand, M. C. (2015). Multicollinearity in spatial genetics: separating the wheat from the chaff using commonality analyses. *Molecular Ecology*, 24, 263-283. <https://doi.org/10.1111/mec.13029>
- R Development Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Accessed Jan. 7, 2020]. Available at: <https://www.r-project.org>
- Resende, M. D. V. (2002). *Genética biométrica e estatística no melhoramento de plantas perenes*. Colombo: Embrapa Florestas.
- Resende, M. D. V., Silva, F. F. and Azevedo, C. F. (2014). *Estatística matemática, biométrica e computacional*. Viçosa: Suprema.
- Resende, M. D. V. (2015). *Genética quantitativa e de populações*. Viçosa: Suprema.
- Resende, M. D. V. (2016). Software Selegen-REML/BLUP: a useful tool for plant breeding. *Crop Breeding and Applied Biotechnology*, 16, 330-339. <https://doi.org/10.1590/1984-70332016v16n4a49>
- Reubens, B., Achten, W. M. J., Maes, W. H., Danjon, F., Aerts, R., Poesen, J. and Muys, B. (2011). More than biofuel? *Jatropha curcas* root system symmetry and potential for soil erosion control. *Journal of Arid Environments*, 75, 201-205. <https://doi.org/10.1016/j.jaridenv.2010.09.011>
- Robertson, A. (1961). Inbreeding in artificial selection programmes. *Genetics Research*, 2, 189-194. <https://doi.org/10.1017/S0016672300000690>
- Rocha, J. R. A. S. C., Machado, J. C. and Carneiro, P. C. S. (2018). Multitrait index based on factor analysis and ideotype design: proposal and application on elephant grass breeding for bioenergy. *GCB Bioenergy*, 10, 52-60. <https://doi.org/10.1111/gcbb.12443>
- Silva, M. J., Carneiro, P. C. S., Carneiro, J. E. S., Damasceno, C. M. B., Parrella, N. N. L. D., Pastina, M. M., Simeone, M. L. F., Schaffert, R. E. and Parrella, R. A. C. (2018). Evaluation of the potential of lines and hybrids of biomass sorghum. *Industrial Crops and Products*, 125, 379-385. <https://doi.org/10.1016/j.indcrop.2018.08.022>
- Smith, H. F. (1936). A discriminant function for plant selection. *Annals of Eugenics*, 7, 240-250.
- Spinelli, V. M., Dias, L. A. S., Rocha, R. B. and Resende, M. D. V. (2015). Estimates of genetic parameters with selection within and between half-sib families of *Jatropha curcas* L. *Industrial Crops and Products*, 69, 355-361. <https://doi.org/10.1016/j.indcrop.2015.02.024>
- Teodoro, P. E., Costa, R. D., Rocha, R. B. and Laviola, B. G. (2016). Contribuição de caracteres agrônômicos para a produtividade de grãos em pinhão-mansão. *Bragantia*, 75, 51-56. <https://doi.org/10.1590/1678-4499.314>
- Wei, R.-P. and Lindgren, D. (1996). Effective family number following selection with restrictions. *Biometrics*, 52, 525-535. <https://doi.org/10.2307/2532892>
- Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Annals of Mathematical Statistics*, 9, 60-62. <https://doi.org/10.1214/aoms/1177732360>
- Woyann, L. G., Meira, D., Zdziarski, A. D., Matei, G., Milioli, A. S., Rosa, A. C., Madella, L. A. and Benin, G. (2019). Multiple-trait selection of soybean for biodiesel production in Brazil. *Industrial Crops and Products*, 140, 111721. <https://doi.org/10.1016/j.indcrop.2019.111721>