Reproductive phenology, morphoagronomic variability, and selection of mother plants of Hancornia speciosa Gomes (Apocynaceae)

Gabriela Corrêa Morais (✉ gabrielacm@usp.br)
USP ESALQ: Universidade de Sao Paulo Escola Superior de Agricultura Luiz de Queiroz
https://orcid.org/0000-0002-9532-9057

Rafael Tassinari Resende
Federal University of Goias: Universidade Federal de Goias

Lázaro José Chaves
Federal University of Goias: Universidade Federal de Goias

Research Article

Keywords: mangaba tree, native fruit tree, Cerrado biome, Brazilian Savanna, genetic gains, Harrison index.

Posted Date: April 3rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2735665/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Reproductive phenology, morphoagronomic variability, and selection of mother plants of

*Hancornia speciosa* Gomes (Apocynaceae)

Gabriela Corrêa Morais\(^1\) (0000-0002-9532-9057), Rafael Tassinari Resende\(^2,3\) (0000-0002-7139-708X), Lázaro José Chaves\(^2\) (0000-0002-4678-9014)

\(^1\)“Luiz de Queiroz” Superior College of Agriculture, University of São Paulo, Genetics Department, Piracicaba, São Paulo, Brazil (e-mail address: gabrielacm@usp.br)

\(^2\)Federal University of Goiás, Plant Breeding Sector, School of Agronomy, Goiânia, Goiás, Brazil

\(^3\)University of Brasília, Forest Engineering Department, Brasília, Federal District, Brazil

Abstract

*Hancornia speciosa* is a fruit tree species with potential for exploitation, which makes the adoption of strategies related to its conservation and breeding relevant. Our study aimed to investigate the reproductive phenology, estimate genetic variability, and identify superior *H. speciosa* genotypes among the accessions that compose the germplasm collection of the Federal University of Goiás, Brazil. Experimental data were obtained from 143 individual accessions. Evaluation of reproductive phenology was conducted on an arbitrary scale of points assigned to trees monthly. Morphological characterization was based on a list of 20 qualitative descriptors and 12 quantitative traits. The dissimilarity between and within botanical varieties and their correlation with the geographic distance of the origin populations were investigated. Analysis of variance components was performed using the Restricted Maximum Likelihood/Best Linear Unbiased Prediction procedure using two linear models. The selection of the \( \approx 10\% \) best individual accessions was performed considering genetic and genotypic values. With the exception of *H. speciosa* var. *speciosa*, the botanical varieties that compose the germplasm collection showed similar reproductive patterns, with a peak in fruit appearance in August. Morphological dissimilarity was smaller within botanical varieties. Among the botanical varieties, dissimilarity was generally positively correlated with the geographic distance. The adoption of a multivariate selection index procedure enables the prediction of genetic gains, resulting in the selection of genotypes from different similarity groups. Clone selection is the most promising strategy for the improvement of *H. speciosa*, ensuring a better response to selection in the short term than recurrent selection.

Keywords: mangaba tree, native fruit tree, Cerrado biome, Brazilian Savanna, genetic gains, Harrison index.

1. Introduction

*Hancornia speciosa* Gomes (Apocynaceae), popularly known as “mangaba tree” or “mangabeira,” is a native fruit species occurring in all Brazilian regions, but with greater incidence in areas covered by the Cerrado biome and coastal lowlands of the Northeast (Pereira et al. 2016). Its fruits (“mangaba”) have high nutritional value and chemical and physical characteristics that make them suitable for human consumption (Nascimento et al. 2014). Monachino (1945) described six botanical varieties for the species: *H. speciosa* var. *cuyabensis*, *H. speciosa* var. *gardneri*, *H. speciosa* var. *lundii*, *H. speciosa* var. *maximiliani*, *H. speciosa* var. *pubescens*, and *H. speciosa* var. *speciosa*. More recently, the species list of the Flora of Brazil 2020 project recognized only two varieties: *H. speciosa* var. *pubescens* Senso Monachino (1945) and *H. speciosa* var. *speciosa* (which includes all other forms) (Reflora 2022). However, other studies support the classification of the plants into two botanical varieties, maintaining the status of *H. speciosa* var. *speciosa* and grouping together *H. speciosa* var. *gardneri*, *H. speciosa* var. *pubescens*, and *H. speciosa* var. *cuyabensis* as a single botanical variety (Terto et al. 2022; Morais et al. 2022). *H. speciosa* var. *maximiliani* and *H. speciosa* var. *lundii* have not been reported in recent studies; however, following
Monachinos’s description, *H. speciosa* var. *maximiliani* would be joined with *H. speciosa* var. *speciosa* and *H. speciosa* var. *lundii* would be joined with other varieties (Morais et al. 2022).

Because of its organoleptic characteristics when ripe, the fruits of *H. speciosa* are appreciated *in natura* (Pereira et al. 2016) and are used as an ingredient in the production of jelly, liqueur, pulp, juice, ice cream, and wine (Costa et al. 2011). Its various parts are also often used in traditional medicine to combat symptoms and diseases (Santos et al. 2013; Coutrim and Souza 2018), and the efficiency of some of these applications has been corroborated by pharmacological studies (Marinho et al. 2011; Penido et al. 2017). Despite its huge potential for exploitation, especially in the food and pharmaceutical industries, the production of mangaba results from predatory extractivism (Fonseca et al. 2018) and populations of this fruit species face drastic reductions due to the advance of anthropogenic activities in their areas of natural occurrence (Gilberto 2015).

In this context, the adoption of strategies that guarantee the proper conservation and use of mangaba trees has become relevant. Knowledge of the phenology of this species can help gather essential information about the dynamics of its reproductive patterns over time. In addition, efforts to select superior genotypes with high yields and fruit quality for use in breeding programs are opportune. Some studies have identified genetic variability in the mangaba tree germplasm collection at the Federal University of Goiás (Goiânia, Goiás state, Brazil) (Ganga et al. 2009; Almeida et al. 2019), qualifying it for use as a base population in a pre-breeding program. However, there is scarcity in the adoption of strategies for the selection of superior plants (Almeida et al. 2021), with emphasis on production and fruit quality, which are the main attractions for agribusiness and consumers.

In this context, our study aimed to investigate reproductive patterns, estimate genetic variability, and identify superior genotypes of *H. speciosa* for traits of agronomic importance among accessions of the germplasm collection of the Federal University of Goiás.

## 2. Material and Methods

### 2.1 Data collection

Experimental data were obtained from November 2020 to November 2021 in the *H. speciosa* germplasm collection of the School of Agronomy of the Federal University of Goiás, located at Goiânia, Goiás, Brazil (lat 16° 35’ 39″ S, long 49° 17’ 26″ W, altitude 720 m asl). According to the Köppen classification, the local climate is Aw tropical, with a rainy season from October to April, and a dry season between May and September. The average annual rainfall is 1,481 mm, the average annual temperature is 22.1°C (Lobato, 2016) and the soil is categorized as a dystrophic dark red latosol.

The germplasm collection was installed in the field in December 2005 with samples from natural subpopulations (populations hereafter) in the Brazilian states of Bahia, Tocantins, Minas Gerais, Goiás, Mato Grosso, and Mato Grosso do Sul, which represent almost the entire area of occurrence of the species in the Brazilian Cerrado (Fig. 1). The mother trees were found in native landscapes, in fragments of vegetation, and along roadsides. To classify the plants as belonging to the same population, a maximum distance of 30 km between them was considered (Ganga et al. 2009).
In the experimental area, a completely randomized block design with four replications was adopted, with one plant per plot spaced 6 m × 5 m (Ganga et al. 2009). At the time of data collection, 143 individual accessions remained in the germplasm collection, structured into 55 open-pollinated progenies originating from the sampled mother plants, representing 29 populations and four of the six botanical varieties according to the classification of Monachino (1945): *H. speciosa* var. *cuyabensis* (24), *H. speciosa* var. *gardneri* (89), *H. speciosa* var. *pubescens* (19), and *H. speciosa* var. *speciosa* (11).

2.2 Phenological data

Monthly evaluations were conducted between November 2020 and October 2021, considering the following phenological events: flowering (appearance of flower buds, anthesis, or senescent flowers) and fruiting (mature or immature fruits attached to the plant). The procedure was based on an arbitrary scale of points ranging from zero to ten, assigned to the trees according to the approximate number of fruits or flowers as follows: “0 points”: absence of flowers and/or fruits; “1 point”: 1 to 10 flowers and/or fruits; “2 points”: 11 to 40 flowers and/or fruits; “3 points”: 41 to 60 flowers and/or fruits; “4 points”: 61 to 80 flowers and/or fruits; “5 points”: 61 to 80 flowers and/or fruits; “6 points”: 121 to 140 flowers and/or fruits; “7 points”: 141 to 160 flowers and/or fruits; “8 points”: 161 to 180 flowers and/or fruits; “9 points”: 181 to 200 flowers and/or fruits; “10 points”: more than 200 flowers and/or fruits.

Monthly averages of points assigned to trees were obtained for each botanical variety, considering a period of 12 uninterrupted months. To evaluate the simultaneity of the appearance of flowers or fruits, we adopted the synchrony index proposed by Morellato et al. (1990), which establishes three categories: asynchrony (less than 20% of the total individuals manifest phenophase), low synchrony (20%...
to 60% of the plants are in phenophase), and high synchrony (more than 60% of the trees are in phenophase).

Here, we considered the assignment of a non-zero monthly score to flowering and/or fruiting in the phenological stage.

2.3 Dissimilarity between and within botanical varieties and its correlation with geographic distance

To determine the morphological dissimilarity among and within botanical varieties of *H. speciosa*, we performed phenotypic characterization considering 20 qualitative descriptors as follows: (i) whole trees: crown shape, trunk surface, branching pattern, tree growth habit, number of bifurcations at the height of the ramification, and branch pubescence; (ii) leaves: leaf shape, leaf base shape, leaf apex shape, leaf pubescence, number of leaves per node, phyllotaxis, and leaf texture; (iii) flowers: pedicel pubescence, calyx pubescence, corolla pubescence, and floral tube pubescence; (iv) fruit: fruit shape, color of fruit skin, and arrangement of fruits.

Furthermore, eight quantitative traits were converted into ordinal qualitative variables based on the intervals provided by the difference between the highest and lowest values assumed by the trait for the formation of the three fictitious discrete classes. The quantitative descriptors analyzed qualitatively were height of the first branch, crown diameter, leaf length, leaf width, petiole length, petiole thickness, internode length, and fruit peduncle length.

We adopted Harrison’s model (1968) provided by Oliveira et al. (2009), which deals with a vector of qualitative variables \( y' = y_1, y_2, ..., y_l \), in which the \( i \)-th component assumes levels or classes. The levels of each qualitative descriptor were converted into binary variables, and the trees \( o_i \) were compared pairwise for a set of binary variables arising from each qualitative criterion. As they were heterogeneous in terms of the number of classes (e.g., five levels for crown shape, three levels for leaf base shape, and two levels for branch pubescence), weighting was performed by dividing the logarithm of the number of levels of each qualitative criterion \( \sum \ln l_i \) by the summation of the logarithms of the number of levels of all variables evaluated \( \sum \ln l_i \). For the indicator function of coincidence of binary variables \( I \) the following were considered: “1,” if \( y_i(o_1) = y_i(o_2) \) and “0,” if \( y_i(o_1) \neq y_i(o_2) \) or NA for missing data. The similarities obtained from Harrison’s model, \( s(o_1, o_2) = \sum I[y_i(o_1), y_i(o_2)] \), were converted into measures of dissimilarity \( d = 1 - s \), organized in a 143 x 143 matrix.

To investigate the possible correlation between geographic distance and morphological dissimilarity between and within botanical varieties, a similarity matrix based on the geographic coordinates of the populations of origin of the accessions was also developed. For this, the distances in kilometers between the sites of origin of each of the trees computed pairwise were converted into values that varied from zero to one, proportional to the geographic distances, assigning “1” to trees from the same locality and “0” for trees that came from the most geographically distant populations.

2.4 Variance components and genetic parameters

We estimated variance components and related genetic parameters for the following quantitative traits: plant height (HEI), trunk circumference 10 cm above ground level (CIR), fruit volume (FRV), and estimated fruit yield (FRY). To calculate the fruit volume (FRV), we used the equation adopted by Resende et al. (2020) and Yunes et al. (2015), which considers the radial and longitudinal circumferences...
of fruits. The estimated fruit yield (FRY) was obtained from the product of the fruit mass and number of fruits per plant (developed or immature fruits present on the tree in September 2021, a period marked by the appearance of “falling fruits”). The mass and volume of the fruits were measured at the Laboratory of Native Plants of the Federal University of Goiás using 5-10 ripe fruits per plant (according to availability).

Analysis of variance was performed using the Restricted Maximum Likelihood/Best Linear Unbiased Prediction (REML/BLUP) method with two models. In the first model (model 1), the effects of families, populations, and botanical varieties in a randomized complete block design were considered as follows:

\[ Y = Xb + Z_1 f + Z_2 p + Z_3 v + \varepsilon, \]

where \( Y \) = vector of the observed value of the variable; \( X \) = incidence matrix under fixed effects; \( Z_1 \) = incidence matrix under random effects; \( f \) = family effect, random; \( p \) = population effect, random; \( v \) = effect of the botanical variety, random; \( b \) = block effect, fixed; and \( \varepsilon \) = experimental error associated with \( Y \), random.

Based on Model 1, the components of genetic variance at the levels among botanical varieties \( (\sigma_{var}^2) \), among populations within varieties \( (\sigma_{pop}^2) \), and among families within populations \( (\sigma_{fam}^2) \), were estimated.

An alternative mixed hierarchical linear model based on similarity matrices was also adopted in the estimation of variance components, considering the additive effects of half-sib families (matrix \( a \)), the geographic distance between populations (matrix \( c \)), and Harrison’s index (1968) (matrix \( h \)), in a randomized complete block design, as follows:

\[ Y = Xb + W_1 a + W_2 c + W_3 h + \varepsilon, \]

where \( Y \) = vector of observed values of variable \( Y \); \( b \) = block effect, fixed; \( X \) = incidence matrix under fixed effects; \( W_1 \) = incidence matrix under random effects; \( a \) = additive effect of half-sib families, random with \( \sigma_a^2 \sim N(0, A \sigma_a^2) \); \( c \) = effect of geographic coordinates, random with \( \sigma_c^2 \sim N(0, C \sigma_c^2) \); \( h \) = effect of the Harrison index (1968), random with \( \sigma_h^2 \sim N(0, H \sigma_h^2) \); and \( \varepsilon \) = experimental error associated with \( Y \), random.

From the best linear model used (1 or 2), the breeding values (EBV) and genotypic values (EGV) were estimated, considering each individual accession and each of the quantitative traits evaluated.

2.5 Selection index and response to selection

The identification of superior individuals was based on the following ideotypes: low plant, with voluminous fruits, and high productivity. For the selection of accessions that combined these characteristics simultaneously, the selection index based on the sum of ranks proposed by Mulamba and Mock (1978) was used. The classification of individuals for the trait set was based on EBV and EGV obtained from the best mixed hierarchical linear model (1 or 2). For individuals with the lowest EBV-based indices, the response to selection for the next recombination cycle was calculated with selection in a single sex (maternal). The response to selection for cloning was estimated for those with the lowest rank for the EGV. In both cases, the 15 individual accessions with the best performance (≈10%) were selected.
Once the 15 best individual accessions were identified by a selection index based on EBV and EGV, the status number ($N_S$) of both subsets was calculated, which corresponded to the inverse of the matrix $A$ mean of the selected units. The values of $N_S$ were also calculated considering the selection of superior individuals for each quantitative trait separately.

3. Results

3.1 Reproductive phenology

Except H. speciosa var. speciosa, the botanical varieties present in the germplasm collection exhibited similar reproductive patterns during the evaluation period (Fig. 2).

Fig. 2 Monthly averages of the points assigned to the botanical varieties of the germplasm collection for the appearance of flowers (flowering) and fruits (fruited) between November 2020 and October 2021. “var”: botanical variety. “cb”: H. speciosa var. cuyabensis. “gn”: H. speciosa var. gardneri. “pc”: H. speciosa var. pubescens; “sc”: H. speciosa var. speciosa. The continuous lines correspond to the mean values for each variety.

In H. speciosa var. cuyabensis, a high level of synchrony for the appearance of flowers was achieved in September, with flowering peaking in October (average of 5.29 points on the grade scale and 91.67% of individuals with flowers). The fruiting body of H. speciosa var. cuyabensis was highly synchronous between May and November, with 100% of its representatives having at least one fruit.
between June and September. The lowest percentage of individuals in that phenophase and the lowest monthly average occurred in February (16.67% and 0.16 points, respectively), and the peak of fruit appearance at different stages of maturation was reached in August, with an average of 8.46 points.

For *H. speciosa* var. *gardneri*, an asynchrony in flowering was detected between November and August. A high synchrony for flowering occurred in September, with 89.89% of the individuals of the botanical variety having at least one flower. October was marked by the climax of flower appearance, with a monthly average of 5.37 points. A high synchrony for fruiting occurred between May and October, with 94.38% of individuals having at least one fruit in August and September. The peak of fruit appearance occurred in August and the monthly average for that phenophase was 6.76 points.

Flowering of *H. speciosa* var. *pubescens* were asynchronous between November and August. High synchrony for this phenophase occurred in September, and the highest monthly average for the appearance of flowers occurred in October (4.79). In May, 52.63% of the individuals in this botanical variety had at least one fruit (low synchrony). From June to October, flowering was highly synchronous, with 94.74% of the individuals in this phenophase in the months of August and September. The peak fruit appearance in *H. speciosa* var. *pubescens* occurred in August (a monthly average of 3.68 points).

In *H. speciosa* var. *speciosa*, flower buds, anthesis, or senescent flowers were observed only in September (apex of flowering, with an average of 0.91 points) and October (average of 0.63 points). Fruiting climax occurred in August and September; in both months, the average phenophase was 0.72 points. In addition to flowering, fruiting did not show high synchrony and the highest percentage of *H. speciosa* var. *speciosa* with at least one fruit was only 54.55% in September and October.

### 3.2 Morphological dissimilarity among and within botanical varieties

Based on Harrison's model (1968), it was possible to detect greater dissimilarity among the botanical varieties of *H. speciosa*, considering the qualitative traits evaluated (Fig. 3). The highest dissimilarity index was observed for *H. speciosa* var. *pubescens* and *H. speciosa* var. *speciosa* (pc x sc) (Fig. 3a). These data also confirm the great dissimilarity between *H. speciosa* var. *speciosa* and all other varieties included in this study, corroborating what can be seen visually in the field (Fig. 4). Within the botanical varieties, the dissimilarity was less pronounced (Fig. 3a). In this study, we highlighted *H. speciosa* var. *speciosa*, *H. speciosa* var. *cuyabensis*, and *H. speciosa* var. *pubescens* as having the lowest dissimilarities within. Both among and within botanical varieties, the dissimilarity index was generally positively correlated with the geographic distances of their populations of origin, up to approximately 750 km (Fig. 3b). Thus, it is expected that mangaba trees belonging to the same variety or different botanical varieties are morphologically more divergent for the traits evaluated here as we increase the physical distance between the sites where they occur.
Fig. 3 Dissimilarity among and within botanical varieties of mangabeira according to Harrison’s model (1968). A: dissimilarity among (above) and within (below) botanical varieties. B: relationship between the geographic distances of the source populations (in km) and the dissimilarity between (above) and within (below) botanical varieties. “cb”: H. speciosa var. cuyabensis. “gn”: H. speciosa var. gardneri. “pc”: H. speciosa var. pubescens; “sc”: H. speciosa var. speciosa

Fig. 4 Details of the morphology of representatives of the four botanical varieties of H. speciosa present in the germplasm collection of Federal University of Goiás. A: H. speciosa var. cuyabensis. B: H. speciosa var. gardneri. C: H. speciosa var. pubescens. D: H. speciosa var. speciosa
3.3 Variance components and genetic parameters

Plant height (HEI) varied from 2.60 m to 11.00 m among accessions, with an average of 6.17 m. The mean circumference of the trunk 10 cm above ground level (CIR) was 63.30 cm, with minimum and maximum values of 22 cm and 158 cm, respectively. The fruit volume (FRV) ranged from 5.98 cm$^3$ to 69.23 cm$^3$, with an average value of 25.59 cm$^3$. The estimated fruit yield (FRY) was equivalent, on average, to 17.92 kg/plant, with a minimum value of 0.02 kg/plant and a maximum of 111.18 kg/plant.

For the linear models used in the analysis of variance components, the Akaike Information Criterion (AIC) values of Model 2 were lower than those of Model 1 for all quantitative traits. AIC values are used as criteria for choosing a statistical model among a set of competitors, with the lowest value being the most parsimonious (Bozdogan, 1987). Thus, our results show that Model 2, which considers the effects of matrices A, C, and H, can better explain the variation observed in HEI, CIR, FRV and FRY (Table 1).

Table 1. Variance components, percentages of variance components and AIC values considering linear models 1 and 2.

<table>
<thead>
<tr>
<th>Character</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effect</td>
<td>Var (%)</td>
</tr>
<tr>
<td>HEI</td>
<td>Family</td>
<td>0.5005</td>
</tr>
<tr>
<td></td>
<td>Population</td>
<td>1.6156</td>
</tr>
<tr>
<td></td>
<td>Variety</td>
<td>0.1097</td>
</tr>
<tr>
<td></td>
<td>Residue</td>
<td>1.6121</td>
</tr>
<tr>
<td>AIC</td>
<td>92.5797</td>
<td></td>
</tr>
<tr>
<td>CIR</td>
<td>Family</td>
<td>85.6383</td>
</tr>
<tr>
<td></td>
<td>Population</td>
<td>147.5337</td>
</tr>
<tr>
<td></td>
<td>Variety</td>
<td>100.7743</td>
</tr>
<tr>
<td></td>
<td>Residue</td>
<td>387.1701</td>
</tr>
<tr>
<td>AIC</td>
<td>120.7983</td>
<td></td>
</tr>
<tr>
<td>FRV</td>
<td>Family</td>
<td>25.1310</td>
</tr>
<tr>
<td></td>
<td>Population</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>Variety</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>Residue</td>
<td>89.5590</td>
</tr>
<tr>
<td>AIC</td>
<td>116.5942</td>
<td></td>
</tr>
<tr>
<td>FRY</td>
<td>Family</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>Population</td>
<td>74.5510</td>
</tr>
<tr>
<td></td>
<td>Variety</td>
<td>16.9194</td>
</tr>
<tr>
<td></td>
<td>Residue</td>
<td>316.0309</td>
</tr>
<tr>
<td>AIC</td>
<td>80.1080</td>
<td></td>
</tr>
</tbody>
</table>


Based on Model 1, it was possible to identify the existence of genetic variation among varieties, among populations within varieties, and among families within populations for the trait HEI, with the percentages of the variances in relation to the total variance corresponding to the 2.86%, 42.10%, and 13.04%, respectively. The same occurred with CIR, with variations among varieties, populations, and
families equivalent to 13.97%, 20.46%, and 11.88% of the total variation, respectively. For FRV, genetic
variation was detected among families (21.91% in relation to the total variance), with null variances among
populations and among botanical varieties. For FRY, the genetic variance among varieties and among
populations corresponded, respectively, to 4.15% and 18.30% of the total variance, with zero variance
among families.

According to Model 2, it was possible to verify the existence of additive genetic variation
among half-sib families for all quantitative attributes evaluated. For HEI and CIR, most of the variance was
due to the effect of the H matrix, which considers Harrison’s similarity index constructed from qualitative
characters. For FRV, most of the variation was explained by the additive genetic variation of families
(matrix A) and for FRY, most of which was due to the geographic distance among populations (matrix C).

### 3.4 Selection index and response to selection

The best accessions (≈10%) of the germplasm collection of the Federal University of Goiás,
selected based on the Mulamba and Mock selection index, are listed in Table 2 and in Supplementary
Information (Table S1).

Table 2. Summary of the best ≈10% accessions based on Mulamba and Mock (1978) selection index,
response to selection for the next recombination cycle ($R_S^{EBV}$) and response to selection for cloning
($R_S^{EGV}$).

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Access</th>
<th>Fam</th>
<th>Pop</th>
<th>Var</th>
<th>Traits under selection</th>
<th>CIR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HEI</td>
<td>FRV</td>
</tr>
<tr>
<td>1º</td>
<td>G122</td>
<td>53</td>
<td>chs</td>
<td>gn</td>
<td>4.96</td>
<td>42.69</td>
</tr>
<tr>
<td>2º</td>
<td>G069</td>
<td>37</td>
<td>cha</td>
<td>cb</td>
<td>5.86</td>
<td>37.08</td>
</tr>
<tr>
<td>3º</td>
<td>G079</td>
<td>40</td>
<td>ron</td>
<td>cb</td>
<td>6.00</td>
<td>40.62</td>
</tr>
<tr>
<td>4º</td>
<td>G071</td>
<td>37</td>
<td>cha</td>
<td>cb</td>
<td>5.93</td>
<td>39.48</td>
</tr>
<tr>
<td>5º</td>
<td>G104</td>
<td>47</td>
<td>cox</td>
<td>gn</td>
<td>5.59</td>
<td>27.42</td>
</tr>
<tr>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td></td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>15º</td>
<td>G132</td>
<td>56</td>
<td>cac</td>
<td>gn</td>
<td>6.41</td>
<td>40.06</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>10</td>
<td>9</td>
<td>3</td>
<td>$R_S^{EBV}$ (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-4.29</td>
<td>22.34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Access</th>
<th>Fam</th>
<th>Pop</th>
<th>Var</th>
<th>Traits under selection</th>
<th>CIR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HEI</td>
<td>FRV</td>
</tr>
<tr>
<td>1º</td>
<td>G071</td>
<td>37</td>
<td>cha</td>
<td>cb</td>
<td>4.88</td>
<td>44.08</td>
</tr>
<tr>
<td>2º</td>
<td>G069</td>
<td>37</td>
<td>cha</td>
<td>cb</td>
<td>5.63</td>
<td>42.69</td>
</tr>
<tr>
<td>3º</td>
<td>G122</td>
<td>53</td>
<td>chs</td>
<td>gn</td>
<td>5.03</td>
<td>48.78</td>
</tr>
<tr>
<td>4º</td>
<td>G067</td>
<td>36</td>
<td>chs</td>
<td>cb</td>
<td>4.35</td>
<td>35.68</td>
</tr>
<tr>
<td>5º</td>
<td>G070</td>
<td>37</td>
<td>cha</td>
<td>cb</td>
<td>4.41</td>
<td>28.62</td>
</tr>
<tr>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td>(...)</td>
<td></td>
<td>(...)</td>
<td>(...)</td>
</tr>
<tr>
<td>15º</td>
<td>G118</td>
<td>51</td>
<td>alc</td>
<td>gn</td>
<td>5.46</td>
<td>25.92</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>$R_S^{EGV}$ (%)</td>
<td></td>
</tr>
</tbody>
</table>
“Var”: botanical variety. “HEI”: plant height (in m). “CT10”: trunk circumference 10 cm of ground level
gardneri. “cb”: H. speciosa var. cuyabensis. “RS_{EBV}\%”: response to selection for the next recombination
cycle. “RS_{EGV}\%”: response to selection for cloning.

Based on EBV, 15 trees comprising 10 families, nine populations, and three botanical
varieties were selected. For HEI, the response to selection for the next cycle (RS_{EBV}) was equivalent to –
4.29%. For FRV and FRY, RS_{EBV} was equal to 22.34% and 6.57%, respectively. The 15 individual
accessions selected from EGV belonged to eight families, six populations, and two botanical varieties of
the mangaba tree. For HEI, the response to selection for cloning (RS_{EGV}) was –4.52%. For FRV and FRY,
RS_{EGV} was 45.99% and 51.12%, respectively.

The selection of accessions using the Mulamba and Mock index considering the EBV
resulted in a slightly higher status number (N_S = 12.16) than that obtained using EGV (N_S = 11.25) (Fig.
5). However, in both cases, the status number values obtained approached the number of selected
individuals.
Fig. 5 Simulations of selection response for the next recombination cycle ($R_S^{EBV}$) and selection response for cloning ($R_S^{EGV}$) versus number of selected individuals versus status number values ($N_s$), considering four quantitative traits, individually. “EBV”: estimated genetic value. “EGV”: estimated genotypic value. “CIR”: trunk circumference 10 cm of ground level (in cm). “HEI”: plant height (in m); “FRV”: fruit volume (in cm$^3$). “FRY”: estimated fruit yield (in kg/plant)

4. Discussion

$H$. speciosa var. cuyabensis, $H$. speciosa var. gardneri, and $H$. speciosa var. pubescens showed a similar flowering pattern of high synchronicity from September to October, with an apex in October. The flowering of a year corresponds to the fruiting of the following year, with a long-time interval between the appearance of flower buds and fruit development. Two well-defined seasons prevail in the Cerrado biome: rainy (from October to March) and dry (from April to September) (Dos Santos et al. 2020). Phenological studies involving plant species from the Neotropical savannah demonstrate a pattern of flowering based mainly on precipitation, with greater flower production at the end of the dry season and beginning of the rainy season (Leão-Araújo et al. 2019; Godim et al. 2020; Novaes et al. 2020), similar to the mangaba tree. A greater production of flowers in a short period (September and October) guarantees a concentration of nectar available to the pollinating agents of $H$. speciosa, which are usually visited by bees, butterflies, and moths (Darrault and Schlindwein, 2005). Thus, the establishment of a commercial orchard and the increase in production of this native fruit species require proximity to areas containing other plant
species capable of sustaining pollinator populations and providing them with food in periods when it is not
in the flowering phase.

The climax of the appearance of fruits at different stages of maturation for all botanical
varieties evaluated occurred in August shortly before the rainy season. Although native species of the
Cerrado fruit are predominant in the rainy season (Maruyama et al. 2019), it is common for the appearance
of fruits to be more pronounced during the dry-rain transition (Silva et al. 2011). In *H. speciosa*, the “falling
fruits” (which come off the tree when ripe) appear in September and October, and the seeds are released at
the beginning of the most favorable season for the establishment and development of seedlings due to the
higher soil moisture (Salazar et al. 2011). In general, this behavior is intrinsically related to species whose
seeds are not dormant and germinate quickly when subjected to ideal water, light, nutrient, and temperature
conditions (Ramos et al. 2017). *H. speciosa* var. *speciosa* predominantly inhabits the northeastern region
of the Cerrado, typically in sandy soils and under low rainfall (Chaves et al. 2020). Silva et al. (2017)
evaluated four populations of this variety in a Cerrado-Restinga ecotonal zone with low fertility soil in the
state of Maranhão, Brazil, and obtained mean values of 338, 330, 280, and 172 fruits per plant. Thus, the
relative inferiority of *H. speciosa* var. *speciosa* regarding the production of flowers and fruits detected in
our study can be explained by the fact that the environmental conditions of the experimental field of the
germplasm collection at the Federal University of Goiás are different from those of their naturally occurring
sites and are unfavorable to their development.

The highest morphological dissimilarity observed here for *H. speciosa* var. *pubescens* and *H.
*speciosa* var. *speciosa* partially reinforces the findings of Flores et al. (2018), who evaluated the chemical
composition of the leaves of four botanical varieties of mangaba trees and reported a more pronounced
differentiation between *H. speciosa* var. *pubescens* and *H. speciosa* var. *speciosa*, as well as a chemical
similarity between *H. speciosa* var. *cuyabensis* and *H. speciosa* var. *gardneri*. Our results showed a greater
morphological similarity between trees belonging to the same botanical variety of mangabeira, according
to the classification proposed by Monachino (1945). According to Chaves et al. (2020), the quantitative
divergence within botanical varieties of this fruit species is mainly due to genetic drift, while the divergence
among varieties is due to both drift and divergent selection.

A final botanical review proposed by the Flora do Brasil 2020 project recognized only two
botanical varieties of the mangaba tree: *H. speciosa* var. *pubescens* (comprising individuals with pubescent
structures) and *H. speciosa* var. *speciosa* (encompassing all other forms) (Reflora 2022). However, the
findings by Collevatti et al. (2018) in a study involving the molecular characterization of mangabeira using
neutral molecular markers did not corroborate this more recent classification, suggesting the existence of
at least four different groups of *H. speciosa*. Furthermore, when estimating genetic parameters from
microsatellite markers, Collevatti et al. (2016) obtained low but significant genetic differentiation among
the same botanical varieties studied here. Our results also showed that, although *H. speciosa* var. *pubescens*
and *H. speciosa* var. *speciosa* are more dissimilar, the varieties proposed by Monachino (1945) diverge
phenotypically when compared to each other, mainly in terms of morphological attributes related to leaves,
flowers, and branches. Considering the proposal to classify *H. speciosa* individuals into only two botanical
varieties, it would be more coherent to separately maintain *H. speciosa* var. *speciosa* and group the other
varieties into a single clade based on morphological traits. The same pattern was defended by Terto et al. (2022), based on molecular markers.

In our study, mangaba trees belonging to the same variety or different botanical varieties are morphologically more divergent as we increase the physical distance between the places where they occur. Fajardo et al. (2018) compared five populations of *H. speciosa* from the coast of Rio Grande do Norte and found no correlation between the genetic and geographic distances. In contrast, a positive and significant correlation between molecular genetic differentiation and geographic distance of up to 400 km was observed among 34 mangabeira tree populations from the Cerrado biome, constituting evidence of isolation by distance (Colevatti et al. 2018).

The use of the Mulamba and Mock selection index is promising in genetic breeding programs for fruit species (Barth et al. 2020; Oliveira et al. 2018), with advantages such as relative simplicity and ease of application. The ideotype defined here to classify the trees (short plants, with voluminous fruits, and high productivity) considered attributes that are usually sought in other fruit trees (Castro et al. 2020; Costa et al. 2020), aiming at great aptitude for commercial cultivation and the consumer market. Although CIR was not considered in the selection index, the $RSEBV$ for this quantitative attribute was $-1.76\%$, while the response to $RSEGv$ was $1.84\%$. As the average of the population selected for HEI based on EBV was also reduced, the decrease in the average trunk circumference may not imply a compromise in the stability of the trees. The choice to adopt selection at the individual level is justified by the fact that this strategy is simpler than the selection of families, in addition to allowing the exploration of the additive variance among and within families.

The great genetic variability observed in *H. speciosa* allows the adoption of two strategies for its breeding: recurrent selection, which involves the selection and intercrossing of sexually propagated and open-pollinated superior plants, as well as the planting of clones generated by vegetative propagation (Chaves, 2006, Pereira et al. 2006). For recurrent selection purposes, it is recommended that an improved population be obtained by harvesting the open-pollinated seeds of the superior individuals identified here, which characterizes selection in the female sex (Almeida et al. 2021). The highest selection gains obtained based on EGV in our study showed that clone planting is the most promising strategy for the genetic breeding of *H. speciosa* when compared to recurrent selection, allowing the acquisition of greater genetic gains in the short term as expected. According to Chaves (2006), after the selection of superior matrices of mangabeira trees (such as those identified here), they must be artificially reproduced via asexual propagation for the production of clones that will be planted and evaluated in an experimental scheme. The best clones should then be propagated vegetatively for broader evaluation, enabling the recommendation of superior materials for planting. It is important to highlight the need to use different clones, given the self-incompatibility characteristics of the species.

The $NS$ is a measure of representativeness indicative of the genetic diversity maintained after selection, and is related to the number of unrelated genotypes in the subgroup (Lindgren et al. 1996). The adoption of a greater intensity of selection allows for greater genetic gains but promotes a reduction in the status number (Andersson et al. 1998). Thus, it is necessary for breeders to seek a strategy that approaches the ideal scenario for both measures (Fig. 5). A higher status number using the Mulamba and Mock index considering the EBV ($NS = 12.16$) than that obtained using EGV ($NS = 11.25$) can be explained by the fact
that the selection of the \( \approx 10\% \) best individuals from EBV brings together a greater number of families, populations, and botanical varieties than the group selected from EGV, reducing the number of related individuals. However, in both cases, it was possible to select dissimilar genotypes, increasing the possibility of allele regrouping in new favorable combinations.

5. Conclusions

The botanical varieties of \textit{H. speciosa} present in the germplasm collection of the Federal University of Goiás flower between September and October, and the peak of immature fruits attached to trees occurs in August. The morphological dissimilarity was greater among the botanical varieties of \textit{H. speciosa} var. \textit{pubescens}, with \textit{H. speciosa} var. \textit{speciosa} having the greatest dissimilarity. Both among and within botanical varieties, this dissimilarity is generally associated with the geographic distance between their populations of origin.

There was genetic variation among the accessions of the germplasm collection for quantitative traits with agronomic importance related to size (plant height and trunk circumference) and fruits (fruit volume and estimated fruit yield). Clone selection is the most promising strategy for the improvement of \textit{H. speciosa}, ensuring a better response to selection in the short term than recurrent selection. However, this strategy depends on the availability of a viable asexual propagation system, which still requires advances for the species.

6. Acknowledgements

We thank the Postgraduate Program in Genetics and Plant Breeding of the Federal University of Goiás for their opportunity to conduct this study. We are also grateful to Erasmo Pereira do Vale Junior for his assistance in data collection in the field.

7. References


Gilberto KY (2015) A mangabeira e os principais aspectos de seu melhoramento genético na Embrapa Amapá. Embrapa Amapá, Macapá


Mulamba NN, Mock JJ (1978) Improvement of yield potential of the ETO blanco maize (*Zea mays* L.) population by breeding for plant traits. Egypt J Genet Cytol 7:40-51


Resende RT, Kuki KN, Correa TR, Zaidan UR, Mota PHS, Telles LAA, Gonzales DGE, Motoike SY,
Rezende MDV, Leite HG, Lorenzon AS. (2020) Data-based agroecological zoning of Acrocomia
aculeata: GIS modeling and ecophysiological aspects into a Brazilian representative occurrence area. Ind

Salazar A, Goldstein G, Franco AC, Miralles-Wilhelm F (2011) Timing of seed dispersal and dormancy,
rather than persistent soil seed-banks, control seedling recruitment of woody plants in Neotropical

Santos ACB, Silva MAP, Santos MAF, Leite TR (2013) Levantamento químico e farmacológico de
/10.1590/S1516-05722013000300019

Silva IA, da Silva DM, de Carvalho GH, Batalha MA (2011) Reproductive phenology of Brazilian
https://doi.org/10.1007/s13595-011-0071-5

Silva LPV, Araújo JRG, Rocha AE, Carvalho MJN, Braun H, Mesquita ML (2017) Characterization of
org/10.1590/0100-29452017746

of Hancornia speciosa Gomes in the Cerrado and Atlantic Forest domains reveal two botanical varieties.
Research Square (Preprint). https://doi.org/10.21203/rs.3.rs-1719926/v1


Statements & Declarations

Funding

This work was supported by Conselho Nacional de Desenvolvimento Científico e
Tecnológico (CNPq, Brazil) (MCTIC/CNPq, n°442061/2019-6). GCM received a graduate scholarship by
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes – Brazil). LJC has been
continuously supported by a research grant from CNPq – Brazil.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

Lázaro José Chaves designed and conceived the study. Gabriela Corrêa Morais collected data
in the field and laboratory. Rafael Tassinari Resende performed the statistical analyses. The first draft of
the manuscript was written by Gabriela Corrêa Morais and the other authors commented on your previous
versions. All authors read and approved the final manuscript.
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryinformation.docx