The cupuaçu genetic improvement program at Embrapa Amazônia Ocidental

Aparecida das Graças Claret de Souza*1; Marcos Deon Vilela de Resende2; Sebastião Eudes Lopes da Silva1; Nelcimar Reis Sousa1
1Embrapa Amazônia Ocidental, Caixa Postal 319, CEP 69011-970, Manaus, AM, Brazil, 2Embrapa Florestas, Estrada da Ribeira, km 111, CEP 83411-000, Colombo, PR, Brazil (* Corresponding Author. E-mail: claret@cpaa.embrapa.br)

ABSTRACT

The objectives of the cupuaçu breeding program developed by Embrapa Amazônia Ocidental are to develop cultivars with high productivity and disease resistance, mainly from witch broom. The program emphasises population improvement through recurrent selection and the deriving of varieties and clones for commercial planting. It started with characterization and evaluation of germplasm collection, and a process of selection of promising genotypes. For the design of the program and definition of a breeding strategy, genetic parameters such as individual narrow and broad heritability and repeatability were estimated for several characteristics. Estimates of these parameters have shown great variability as a consequence of production fluctuations between years, genotype x environment interaction and change of traits with plant age. However, the results show excellent possibilities for the genetic improvement of the species.

KEY WORDS: Perennial crops, breeding strategies, selection, cultivar development, genetic parameters.

INTRODUCTION

Cupuaçu (Theobroma grandiflorum), a native fruit tree from the Amazon, belongs in the Sterculiaceae and has become a new option for the agricultural business. The pulp is used to make industrialised juice, ice cream, candy, jam, nectar, yoghurt and biscuits among others. From the nuts, the cupulate (Nazaré et al., 1990), a product similar to the chocolate, can be obtained. The plant can also be used for the production of aroma, items of personal hygiene and cosmetics.

Until the beginning of the 1970's, the emerging cupuaçu culture was carried out in a small scale in domestic orchards. Nowadays, however, the cultivation has spread to greater areas, mainly in the states of Pará, Amazonas and Rondônia. In the state of Amazonas the planted area increased from 331 hectares in 1991 to 9,733 hectares in 1999 (INDICADORES, 1992; IDAM,1997; 1999). By and large, this increase in cultivated areas occurred without the use of appropriate technology and market considerations. Its use in agroforestry systems also contributed to the increased of cultivation. As a consequence, the demand for technologies specialised in giving an enterprise character to the activity is increasing. Among the several things that have limited industrialisation of cupuaçu, the lack of improved cultivars is considered a major limitation.

The witch broom, caused by the fungus Crinipellis perniciosaea, is recognized as the main disease of the cupuaçu, being responsible for considerable economic losses by affecting production and by increasing the number of employees involved in sanitary pruning.

The Embrapa Amazônia Ocidental research project on the cupuaçu genetic resources and breeding aims at the development of improved cultivars with great productivity, insects and disease resistance, specially to witch broom, and at contributing to the development of an agriculture based on fruit cultivation.

GENERAL FEATURES OF THE CUPUAÇU PLANT

The cupuaçu tree has 7 meters of diameter at the crown and is between 4 to 10 meters high (Figure 1). The leaves are simple pink and pubescent when young, turning green when mature. The bisexual flowers are the largest in the genus, varying in color from white to red.

There are morphological barriers against self or wind pollination, leading to a dependence on pollinating insects for a safe natural reproduction of the species. Plant propagation can be sexual or through cloning, with fruit production beginning two years after planting.
The cupuaçu fruit comes in various shapes which vary from 9 to 15 cm in diameter, 10 to 40 cm in length and 300g to 4000g in weight (1500g average), and is composed of (Figure 2) pulp (24% to 50%), seed (10% to 29%), bark (39% a 52%) and placenta (2% to 4%). The number of seeds per fruit vary from 15 to 50. In seedless fruits, pulp percentage varies from 60% to 68%.

**GENETIC RESOURCES**

*Theobroma grandiflorum* grow spontaneously in the eastern part of Amazonia, on terra firme forests and in the high floodplain, in the southern and eastern parts of the state of Pará (PA) (Cuatrecasas, 1964; Ducke, 1940). Cultivated even before the discovery of America (Ducke, 1946), it is difficult to ensure with precision if the individuals are sub-spontaneous (Moraes et al., 1994). The cupuaçu plant grows in high temperatures, with an annual mean of 21.6°C to 27.5°C, relative humidity of 77% to 88% and rainfall in the range of 1900 mm to 3100 mm (Diniz et al., 1984).

There are six germoplasm collections of the species in Brazilian Amazonia. The collection at Embrapa Amazônia Ocidental in Manaus, state of Amazon (AM), has 128 clones and 119 half-sib families, 54% of which were collected in Central Amazonia (Negro/Solimões rivers-AM), 19% in the Upper Solimões-AM, 16% in the Lower Amazonas -AM and 11% in Bragantina –PA, (Souza and Sousa, 1999). The INPA collection has 79 entries from Tucuruí and from Pré-Amazônia of Maranhão (Martel, 1999). In Pará, the collection of the Embrapa Amazônia Oriental has materials from Amazonas, Pará and Amapá, mostly concentrated in two sites: Belém, with 46 clones, and Tomé-açu, with 49 progenies (Alves et al., 1996). In the state of Acre, Rio Branco, the Embrapa collection is based on 12 entries collected by seed in Acre (Ledo and Costa, 1999). In Rondônia, Porto Velho, the Embrapa collection contains 36 entries collected by seed (Ribeiro, 1998) and in Amapá, Macapá, the Embrapa collection started in 1997, with 50 entries collected by seed throughout the state (Farias Neto, 1998).

**BREEDING PROGRAM**

The cupuaçu breeding program developed at Embrapa Amazônia Ocidental started with the characterisation and evaluation of the germplasm collection. Currently, the promising genotypes are being selected to provide improved material in the short and long term. High genetic gains are expected due to the great genetic variability evidenced by the traits under selection.

The breeding strategy emphasises techniques that ensure quick and continuous genetic gains through the increase of gene frequency and simultaneous maintenance of genetic variability in the breeding population.

The breeding and genetic conservation programs are

---

**Table 1.** In natura pulp content of the cupuaçu fruit.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humidity (%)</td>
<td>82.2</td>
</tr>
<tr>
<td>Acidity (%)</td>
<td>2.4</td>
</tr>
<tr>
<td>°Brix</td>
<td>13.3</td>
</tr>
<tr>
<td>PH</td>
<td>3.1</td>
</tr>
<tr>
<td>Vitamin C (mg%)</td>
<td>26.2</td>
</tr>
<tr>
<td>Pectin (%)</td>
<td>0.39</td>
</tr>
<tr>
<td>Volatiles (%)</td>
<td>89.0</td>
</tr>
<tr>
<td>Amino-acids (mg % de N)</td>
<td>21.9</td>
</tr>
<tr>
<td>Ethereal extract (%)</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Source: Souza et al. (1999); Nazaré (1997).
based on trials carried out with 119 open pollinated progenies and 128 clones. These half sib families and clones provide an effective population size ($N_e$) of around 600 unrelated individuals. This $N_e$ is enough to ensure suitable genetic gain rates in short and long-term recurrent selection programs.

The breeding program will be periodically adjusted taking into account the changes in the breeding objectives in terms of products and plant architecture.

**SHORT AND MIDTERM IMPROVEMENT**

**Germplasm collection, characterization and evaluation**

In the germplasm collection phase, the main objective was to ensure enough genetic variability to support and sustain the breeding program of the species. Collections took place in small farms in the Amazon, considering agronomic traits, witch broom disease occurrence and plant age, which should be over ten years old to provide a better evaluation of disease occurrence.

The trials were conducted in randomised complete block designs with four replications using five plants per single line plot (for the progeny tests) and two plants per single line plot (for the clone tests). Productive traits such as number of fruits per plant, total weight of fruits per plant, and qualitative traits, including the chemical composition, were evaluated. Other important traits such as resistance/tolerance to witch broom disease and plant architecture were also evaluated on an annual basis.

The characterization and phenotypic evaluation of the germplasm followed the descriptors list as provided by Souza (1996).

**Clone Selection**

Individual plant selection among and within half-sib families were selected based on characterisation, evaluation and genetic distance. These plants were propagated giving rise to the clones to be evaluated. Agronomic and quality traits were emphasised.

**Clone trials**

The selected clones are evaluated in comparison trials and used in diallel crosses aiming to verify the inter-clone compatibility and general combining ability. The best clones in the clone trials will constitute the new cultivars.

**LONG TERM BREEDING PROGRAM**

**Recurrent selection**

After evaluation of the base population, the selected individuals are used for the production of improved material through cloning or seed propagated plants, as well as to form the breeding population for long term improvement. Basically, a recurrent selection program within populations was adopted. A more detailed breeding strategy for cupuaçu is presented in Figure 4.

**GENETIC PARAMETERS ESTIMATION AND BREEDING AND GENOTYPIC VALUES PREDICTION**

For the design of an efficient breeding strategy, estimates of genetic parameters such as narrow and broad sense heritabilities and repeatability for all traits are required. By the same token, efficient selection requires rankings of several individuals based on their predicted breeding and/or genotypic values. Genetic parameters estimates are essential in the prediction procedures as well.

The standard and unified procedure for simultaneous genetic parameters estimation and genetic values prediction in perennial crops is the restricted maximum likelihood (REML) / best linear prediction (BLUP) at an individual level. The advantages of using the REML/BLUP procedure in perennial species are:

(i) simultaneous accounting for environmental effects, variance components estimation and breeding values prediction;

(ii) can be applied to unbalanced data;

(iii) more precise estimates arising from the use of a greater number of observations;
(iv) unbiased comparison of individuals across time and space;
(v) use of complex data structures such as repeated measures, overlapping generations and different locals and years of evaluation;
(vi) better statistical properties than methods based on the analysis of variance;
(vii) use of predicted breeding values in the planning of crosses;
(vii) full use of genealogy and accumulated data over generations.

In this study, the REML/BLUP procedure is described according to the sequence reported by Resende (2000; 2002) and Resende et al. (2000) to match it with the data structure of the cupuaçu breeding program.

As for the half sib progenies evaluated in the randomised complete block designs with several plants per plot and several measures per individual, the following linear mixed model is suitable for genetic evaluation.

\[ y = Xb + Za + Wc +Tp + e, \]

where: \( y, b, a, c, p \) and \( e \): vectors of observations, of block

![Diagram](image-url)

**Figure 4.** Short and long term cupuaçu breeding and genetic conservation programs at Embrapa Amazonia Ocidental.
effects (fixed), of additive genetic values (random), of plot effects (random), of random permanent effects (non additive genetic and environmental within plot effects) and of random errors effects, respectively.

$X$, $Z$, $W$ and $T$: incidences matrix for b, a and c, respectively.

Under this model, the fixed effects of blocks, measures and measure x block interaction, can be adjusted only in one fixed effect: block-measure combination.

Distributions and structures of means and variances

\[
\begin{align*}
\gamma &\sim N(Xb, V) \\
\sigma^2 &= \sigma^2_x \\
\sigma^2_\varepsilon &\sim N(0, \sigma^2_\varepsilon) \\
\rho &\sim N(0, \sigma^2_\rho) \\
\phi &\sim N(0, \sigma^2_\phi) \\
\text{Cov} (a, c') &= 0; \quad \text{Cov}(a, \rho') = 0; \quad \text{Cov}(a, \theta') = 0; \\
\text{Cov} (\rho, c') &= 0; \quad \text{Cov}(\rho, \theta') = 0; \quad \text{Cov}(c, \theta') = 0
\end{align*}
\]

\[
\begin{pmatrix}
y \\
y \\
y \\
E \\
c \\
\rho \\
\theta \\
\end{pmatrix}
= \begin{pmatrix}
Xb \\
Xb \\
Xb \\
\begin{pmatrix}
ZG & W & I & TP & R
\end{pmatrix} \\
\begin{pmatrix}
GZ & W & I & TP & R
\end{pmatrix} \\
\begin{pmatrix}
PT & 0 & 0 & P & 0
\end{pmatrix} \\
\begin{pmatrix}
R & 0 & 0 & 0 & R
\end{pmatrix}
\end{pmatrix}
\]

, where:

\[
P = I \sigma^2_\rho \\
V = ZA \sigma^2_x Z' + WI \sigma^2_\varepsilon W' + TI \sigma^2_\rho T' + \sigma^2_\theta
\]

Mixed model equations

\[
\begin{pmatrix}
X'X & X'Z & X'W & X'T \\
Z'X & Z'Z + A^{-1} \lambda_1 & Z'W & Z'T \\
W'X & W'Z & W'W + \lambda_2 & W'T \\
T'X & T'Z & T'W & T'T + \lambda_3
\end{pmatrix}
\begin{pmatrix}
\hat{b} \\
\hat{a} \\
\hat{c} \\
\hat{p}
\end{pmatrix}
= \begin{pmatrix}
X'y \\
Z'y \\
W'y \\
T'y
\end{pmatrix}
\]

, where:

\[
\lambda_1 = \frac{1 - \rho}{h^2} \sigma^2_x, \quad \lambda_2 = \frac{1 - \rho}{c^2} \sigma^2_\varepsilon, \quad \lambda_3 = \frac{1 - \rho}{\rho^2} \sigma^2_\rho.
\]

\[
h^2 = \frac{\sigma^2_x}{\sigma^2_x + \sigma^2_\varepsilon + \sigma^2_\rho + \sigma^2_\theta}: \text{individual narrow sense heritability in the block in one measure;}
\]

\[
\rho = \frac{\sigma^2_x + \sigma^2_\varepsilon + \sigma^2_\rho}{\sigma^2_x + \sigma^2_\varepsilon + \sigma^2_\rho + \sigma^2_\theta}: \text{individual repeatability in the block;}
\]

\[
\rho^2 = \frac{\sigma^2_\rho}{\sigma^2_x + \sigma^2_\varepsilon + \sigma^2_\rho + \sigma^2_\theta}: \text{determination coefficient of within plot permanent effects;}
\]

\[
c^2 = \frac{\sigma^2_\varepsilon}{\sigma^2_x + \sigma^2_\varepsilon + \sigma^2_\rho + \sigma^2_\theta}: \text{correlation due to common environmental effects in the plot.}
\]

Iterative estimators of variance components by REML

\[
\hat{\sigma}^2_x = [\hat{a}' A^{-1} \hat{a} + \hat{\sigma}^2_\varepsilon \text{tr} (A^{-1} C^{32})] / q
\]

\[
\hat{\sigma}^2_\varepsilon = [\hat{c}'c + \hat{\sigma}^2_\varepsilon \text{tr} C^{32}] / s
\]

\[
\hat{\sigma}^2_\rho = [\hat{\rho}' \hat{\rho} + \hat{\sigma}^2_\rho \text{tr} C^{44}] / q, \text{ where:}
\]

\[
C^{-1} = \begin{bmatrix}
C_{11} & C_{12} & \ldots & C_{1q} \\
C_{12} & C_{22} & \ldots & C_{2q} \\
\vdots & \vdots & \ddots & \vdots \\
C_{1q} & C_{2q} & \ldots & C_{qq}
\end{bmatrix}^{-1}
\]

The peculiarity of the data structure of perennial plants comes from the repeated observations or measures taken from each individual. Therefore, it must take into consideration the parameters of heritability and repeatability during the breeding values prediction procedures.

The same model can be used for clonal values prediction by replacing the a vector by a g vector that takes into account the total genotypic value. Similarly, variance $a$ must be changed to variance $g$ in the estimators and predictors.

The main objective of the cupuaçu breeding program is to maximise fruit production per year using a selection based on the mean of several harvests. Thus the heritability ($h_m^2$) level of several harvests, which is related to the heritability of one harvest and to repeatability, becomes an important characteristic as shown by the expression:

\[
h_m^2 = \frac{m h^2}{1 + (m - 1) \rho}.
\]

The repeatability ($\rho_m$) of several harvests is given by the formula:

\[
\rho_m = \frac{m \rho}{1 + (m - 1) \rho}.
\]

Some estimates of these parameters for the traits fruit production per plant and number of fruits per plant
are shown in the Table 2.

Heritability of just one harvest is moderate for both traits (Table 2). These results are expected for quantitative traits and are in agreement with the results published for cacao (Dias and Resende, 2001). However, these individual heritabilities are enhanced by using several annual harvests per plant, reaching values of 0.69 and 0.67 for number of fruits per plant and fruit production per plant, respectively.

Individual repeatability estimates were low for both traits (Table 2), showing the presence of fluctuations in productivity between years, among individuals.

**REPEATABILITY IN CLONE SELECTION**

With balanced data, repeatability in clone selection can be expressed as the mean of several ramets level. Using data from fruit production of four consecutive years, starting at the age of six, Souza and Sousa (2001) presented the following estimates: (i) repeatability of the mean at the clone level: 0.56; (ii) repeatability of the mean of four measures or coefficient of determination: 0.83. Consequently, it will take eight years of evaluation to have a selection with 90% of determination.

**NUMBER OF HARVESTS FOR AN ACCURATE SELECTION**

The efficiency of using several harvests rather than only one is given by $\left[ m/[1+(m-1) r]\right]^{1/2}$. Table 3 shows these efficiencies for several levels of individual repeatability.

**Table 2.** Estimates of the parameters: individual narrow sense heritability ($\hat{h}^2$), individual repeatability ($\hat{\rho}$), heritability of several ($m = 7$) harvests ($\hat{h}_m^2$), repeatability of several ($m = 7$) harvests ($\hat{\rho}_m$), correlation due to common environmental effects in the plot for the traits fruit production per plant (FP) and number of fruits per plant (NF).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>NF</th>
<th>FP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{h}^2$</td>
<td>0.26</td>
<td>0.24</td>
</tr>
<tr>
<td>$\hat{\rho}$</td>
<td>0.27</td>
<td>0.25</td>
</tr>
<tr>
<td>$\hat{h}_m^2$</td>
<td>0.69</td>
<td>0.67</td>
</tr>
<tr>
<td>$\hat{\rho}_m$</td>
<td>0.72</td>
<td>0.70</td>
</tr>
<tr>
<td>$\hat{c}^2$</td>
<td>0.002</td>
<td>0.002</td>
</tr>
</tbody>
</table>

At least four annual harvests or repeated measures are necessary to provide an efficient cupuaçu selection (individual repeatability around 0.25). In addition, the increments for each additional measure are less than ten percent. With four measures the individual heritabilities enhance from around 0.25 to around 0.55 for both traits, which is a representative high value.

**Table 3.** Efficiency values (EV) for mass selection based on several ($m$) numbers of harvests, considering different values of the individual repeatability ($r$).

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$m$</th>
<th>EV</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>1</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.04</td>
</tr>
<tr>
<td>0.85</td>
<td>2</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.05</td>
</tr>
<tr>
<td>0.80</td>
<td>2</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.07</td>
</tr>
<tr>
<td>0.75</td>
<td>2</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.11</td>
</tr>
<tr>
<td>0.70</td>
<td>2</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.14</td>
</tr>
<tr>
<td>0.65</td>
<td>2</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.16</td>
</tr>
<tr>
<td>0.60</td>
<td>2</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.20</td>
</tr>
<tr>
<td>0.55</td>
<td>2</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.23</td>
</tr>
<tr>
<td>0.50</td>
<td>2</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.26</td>
</tr>
<tr>
<td>0.45</td>
<td>2</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.30</td>
</tr>
<tr>
<td>0.40</td>
<td>2</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.35</td>
</tr>
<tr>
<td>0.35</td>
<td>2</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.40</td>
</tr>
<tr>
<td>0.30</td>
<td>2</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.45</td>
</tr>
<tr>
<td>0.25</td>
<td>2</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.67</td>
</tr>
</tbody>
</table>
RESUMO

Melhoramento genético de cupuaçu na Embrapa Amazônia Ocidental

Os objetivos do melhoramento genético de cupuaçu na Embrapa Amazônia Ocidental são orientados para obtenção de cultivares com elevada produtividade e resistência a doenças, principalmente vassoura de bruxa. O programa enfatiza o melhoramento de populações via seleção recorrente e o desenvolvimento de variedades e clones melhorados destinados aos plantios comerciais. O programa iniciou-se com a coleta, caracterização e avaliação do germoplasma, e está passando pelo processo de seleção de genótipos mais promissores. Para o direcionamento do programa e definição de uma estratégia de melhoramento, parâmetros genéticos como as herdabilidades individuais no sentido restrito e amplo bem como a repetibilidade individual tem sido estimadas para os vários caracteres. Estimativas destes parâmetros tem revelado grande variação nos valores encontrados em virtude, provavelmente, da alternância da produção, normalmente verificada na cultura, das interações genótipos x ambientes e sobretudo da evolução das variáveis com a idades das plantas. Porém os resultados revelam excelentes possibilidades para o melhoramento da espécie.

REFERENCES


Received: January 18, 2002;
Accepted: February 08, 2002.