Euterpe edulis seed germination parameters and genotype selection

Soler-Guilhen, José Henrique; Bernardes, Carolina de Oliveira; Marçal, Tiago de Souza; Oliveira, Wagner Bastos dos Santos; Ferreira, Marcia Flores da Silva; Ferreira, Adésio

GENETICS AND PLANT BREEDING

Euterpe edulis seed germination parameters and genotype selection

José Henrique Soler-Guilhen [1]
Universidade Federal do Espírito Santo, Brazil

Carolina de Oliveira Bernardes [1*]
Universidade Federal do Espírito Santo, Brazil
ORCID: http://orcid.org/0000-0003-1815-7201

Tiago de Souza Marçal [2]
Universidade Federal de Viçosa, Brazil

Wagner Bastos dos Santos Oliveira [3]
Universidade Federal de Santa Catarina, Brazil

Marcia Flores da Silva Ferreira [1]
Universidade Federal do Espírito Santo, Brazil

Adésio Ferreira [1]
Universidade Federal do Espírito Santo, Brazil

ABSTRACT.: The palm tree juçara (*Euterpe edulis* Mart) is considered one of the most important and threatened native species in the Atlantic Forest. The search for juçara seeds (*Euterpe edulis* Martius) has increased, generating a need for technologies favorable to their production, quality and conservation. The aim of this study was to describe the seed germination patterns of 45 juçara accessions by using a nonlinear model to examine the behavior of these patterns and select superior genotypes. After clustering 45 juçara accessions into four groups, four nonlinear models, namely, the logistic, Gompertz, von Bertalanffy and Weibull models, were tested based on their fit for each group of accessions. The Gompertz model presented the best fit to describe the germination growth curve of *E. edulis* seeds. Groups 3 and 4 presented the highest daily germination rates and were considered the most vigorous. Analysis of deviance was performed, revealing genetic variability among the genotypes in terms of seed germination and vigor. It was possible to find a nonlinear model that best explained our data, and this model may also be a promising model for other palm tree species. It is possible to select genotypes based on the characters we analyzed, due to the elevated genetic variability.

Keywords: Juçara, nonlinear models, genetic variability.

Introduction

Continuous forest fragmentation, which has contributed to the current high extinction rates in some species, has been especially reported in tropical regions (Pimm et al., 2014). The palm tree juçara (*Euterpe edulis* Mart) belongs to the Arecaceae family and is considered one of the most important and threatened native species found in the Atlantic Forest (Inácio, Lima, Lopes, Pessoa, & Teixeira, 2013). Juçara is distributed along the Brazilian coast from Rio Grande do Norte to Rio Grande do Sul (Henderson, 2000).

Euterpe edulis is considered one of the most important palm trees in Brazil and the one producing the best-flavored edible fruits (Cavalcante, Pivetta, Iha, & Takane, 2012). However, the intense exploitation of heart of palm, which is associated with deforestation, and the fact that *E. edulis* does not produce tillers have reduced its populations, endangering this species (Ganem, 2011).

The search for *E. edulis* seeds for commercial purposes, as an ornamental plant and for the forest regeneration has increased, which generates a need for technologies that are favorable to the production, quality and conservation of these seeds (Martins, Bovi, Nakagawa, & Machado, 2009) since palm tree propagation mainly happens through seeds (Cavalcante et al., 2012). Nevertheless, *E. edulis* seed germination has a low percentage and is considered slow and ununiform (Tavares, Ramos, Aguiar, & Kanashiro, 2008; Cavalcante et al., 2012).

Genetic factors affect seed germination by influencing seed physiological quality. However, other factors, such as seed size, maturation, dormancy, viability, age, and climate, may also influence seed germination (Casas et al., 2017; El-Keblawy, Shabana, Navarro, & Soliman, 2017; Penfield & MacGregor, 2017). Therefore, studies that aim to understand seed developmental patterns are important for breeding and conservation programs. This knowledge may be useful for supply sources that maintain genetic variation, to assist in the decision of whether seeds that are more vigorous should be used in breeding programs and to store seeds that may eventually be used in reforestation programs.

Statistical models may facilitate the understanding of some seed parameters; nonlinear models are some of the most widely used methods for continuous data. These models usually provide a good fit to the data and use fewer parameters than linear models. The germination growth curves determined by nonlinear regression will assist in the interpretation of parameters or functions and may facilitate studies on germination percentage, germination speed and germination average time (Paine et al., 2012).

This knowledge may also assist in the process of genotype selection, which, as in any phase of a breeding program, must be based on genetic averages, not phenotypic ones, because phenotypic characteristics vary according to the environment (Frank, Pluess, Howe, Sperisen, & Heiri, 2017). Therefore, we emphasize the need for an instrument that is efficient for genotype selection, such as the REML (Restricted Raximum Likelihood)/BLUP (best Linear Unbiased Prediction) method.

BLUP is a standard method for estimating random effects in a mixed model that was originally developed for the estimation of animal breeding values and is currently widely used in many areas of research (Piepho, Möhring, Melchinger, & Büchse, 2008). The REML method has been used as a computationally attractive choice for large data sets and complex linear mixed effects models. In other words, the REML/BLUP method allows one to divide the phenotypic variation into genetic, environmental and "genotype x environment" components (Resende, 2007).

The aim of this study was to use a nonlinear model identity test to compare seed germination patterns of 45 juçara accessions and select superior genotypes based on these characters.

Material and methods

Experimental characterization

We analyzed characteristics related to the seed vigor of 45 accessions of *E. edulis* from forest fragments located in the southern (Mimoso municipality) and Caparaó regions (Alegre 1 and 3, Guaçuí and Ibitirama municipalities) of the state of Espírito Santo. Since fruits of juçara were our object of study, instead of the heart of palm (the extraction of which is the most important reason why this plant is endangered), no licenses were needed for the collection of the plant material.

One hundred fruits were collected from each genotype, and their pulp was extracted. To determine the seeds' recalcitrance, the accessions (before germination) were disinfected with 0.5 % sodium hypochlorite for 15 minutes, and scarification was rapidly conducted (Teixeira, Vieira, Partelli, & Silva, 2011).

Afterwards, the seeds were sowed on sterilized sand in polystyrene trays and were maintained in a BOD (Biochemical Oxygen Demand) incubator for 60 days at 25°C (Brasil, 2009) with a photoperiod of 8 to 16 hours. Counts were made daily after the first plantlet emerged.

The analyzed characteristics were daily percentage of germination during an observation period of 60 days (G); germination speed (GSI); germination average time (GAT); first germination count (FGC) 20 days after planting; and germination percentage (GP) 60 days after planting (Brasil, 2009). The experiment was carried out following a completely randomized design with 45 accessions, four repetitions and parcels composed of 25 seeds.

Repeatability analysis

The Tocher optimization method [1] was used to group the most homogeneous accessions together as a way to indicate germination potential. This method is based on the Mahalanobis distance matrix (D^2) [2] (Cruz, Regazzi, & Carneiro, 2012) and was constructed based on the variables GSI, GAT, FGC, and GP:

$$d_{(ij)k} = d_{ik} + d_{jk}$$
$$\frac{d_{(group)k}}{\eta} \leq \alpha \quad [1]$$

$$D^2_{ii'} = d'_{ii'} S^{-1} d_{ii'} \text{ to } i \neq i' \quad [2]$$

$$S = \begin{bmatrix} \sigma^2_{e_{11}} & \sigma_{e_{12}} & \sigma_{e_{13}} & & \sigma_{e_{1p}} \\ \sigma_{e_{21}} & \sigma^2_{e_{22}} & \sigma_{e_{23}} & \cdots & \sigma_{e_{2p}} \\ \sigma_{e_{31}} & \sigma_{e_{32}} & \sigma^2_{e_{33}} & \cdots & \sigma_{e_{3p}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{e_{p1}} & \sigma_{e_{p2}} & \sigma_{e_{p3}} & & \sigma^2_{e_{pp}} \end{bmatrix}$$

where: $d_{(ij)k}$ is the distance from individual *k* and the group formed by individuals *ij*; *n* is the number of individuals in the group; *α* is the minimum distance of the possible groups (i.e., the entry criterion for a new accession to the group); D^2_{ii} is the Mahalanobis generalized distance among the genotypes i and i'; $d_{ii'}$ is the distance among the genotypes i and i' ($x_{ij}-x_{i'j}$); S^{-1} is the inverse of the variance and covariance residual matrix; σ^2_e is the residual variance of the variable p; and σ_e is the covariance among the variables p_i and p_j to i (j.

Four nonlinear models were used to model the germination process (G) of the accession groups as a function of time. The regression coefficient θ_1 was an asymptotic approximation of the maximum germination in all the models, and θ_3 was interpreted as the mean germination rate in the logistic, Gompertz and von Bertalanffy models (Online Resource 1).

The Gauss-Newton algorithm [3] was used to obtain the parameter estimates because of the interdependence of parameters in nonlinear models. This algorithm minimizes the sum of the squared errors (Björck, 1996). The initial estimates of the parameter vector were obtained by the logarithmic linearization of the models from Online Resource 1.

$$\theta^{i+1}=\theta^i+(J'J)^{-1}Jr(\theta^i) \quad [3]$$

$$r(\theta^i) = y - f(\theta^i,t)$$

$$\theta^i = \begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_k \end{bmatrix}^i$$

$$J = \begin{bmatrix} \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_1)}{\partial\theta_1} & \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_1)}{\partial\theta_2} & \cdots & \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_1)}{\partial\theta_k} \\ \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_2)}{\partial\theta_1} & \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_2)}{\partial\theta_2} & \cdots & \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_2)}{\partial\theta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_j)}{\partial\theta_1} & \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_j)}{\partial\theta_2} & \cdots & \dfrac{\partial f(\theta_1,\theta_2\cdots\theta_k,t_j)}{\partial\theta_k} \end{bmatrix}$$

where: θ^{i+1} is the actualized parameter vector; θ^i is the parameter vector of the *i*th interaction; J is the Jacobian gradient; $r(\theta^i)$ is the error vector; y is the observed germination percentage; and t is the level vector for the time factor.

Five indicators were used to evaluate the model fit: the determination coefficient (R^2) [4] (Kennedy, 2008), the mean squared error (MSE) [5] (Mello, Magalhaes, Breda, & Regazzi, 2008), the mean absolute deviation (MAD) [6] (Sarmento et al., 2006), the Akaike information criterion (AIC) [7] (Akaike, 1974) and the Bayesian information criterion (BIC) [8] (Schwarz, 1978).

$$R^2 = 1 - \frac{SSE}{TSS} \, [4]$$

$$MSE = \frac{1}{RDF} r(\theta^n)' r(\theta^n) \, [5]$$

$$AIC = 2k - 2\ln(\hat{L}) \, [7]$$

$$BIC = k\ln(n) - 2\ln(\hat{L}) \, [8]$$

where: SSE is the sum of the squared errors; TSS is the total sum of squares; RDF is the residual degrees of freedom; $r(^n)$ is the error vector of the last interaction; n is the number of observations; k is the number of model parameters; and $\ln(\hat{L})$ is the Napierian logarithm of the maximum likelihood function.

After determining the ideal model for describing the groups of families discriminated by the Tocher criterion, the germination tax (GT) function [9] was determined using the first derivative of the chosen model as a function of time. The Newton-Raphson algorithm (Björck, 1996) [10] was used with the GT function to determine the point of maximum germination tax. The initial value was determined through the functional analysis of the germination tax.

$$GT = \frac{\partial f(\theta_1, \theta_2 \cdots \theta_k, t)}{\partial t} \, [9]$$

$$GT' = \frac{\partial GT}{\partial t}$$

$$GT'' = \frac{\partial^2 GT}{\partial t^2}$$

where: t^{i+1} is the actualized approximation of the maximum point of the germination tax function; t^i is the approximation of the maximum point of the germination tax function in the *i*th interaction; GT' is the first derivative of the germination tax function; and GT" is the second derivative of the germination tax function.

The nonlinear model that best represented the family groups was submitted to a model identity analysis. This method was used to verify the equality of the biological parameters with the F statistic.

$$F = \frac{[SSE(\omega) - SSE(\Omega)]RDF(\Omega)}{[RDF(\Omega) - RDF(\omega)]SSE(\Omega)} \sim F[RDF(\Omega) - RDF(\omega), RDF(\Omega)]$$

where: SSE(()is the sum of the squared errors from the reduced model; SSE(()is the sum of the squared errors from the complete model; RDF(() is the residual degrees of freedom from the complete model; and RDF(() is the residual degrees of freedom from the reduced model.

The GENES program was used to estimate the Mahalanobis distance matrix and the cluster produced by the Tocher optimization method (Cruz, 2013).

Estimation of the genetic parameters

For the genetic parameters, estimations of deviance and variance components were carried out by the REML method, prediction of the genotypic values through the best linear unbiased prediction (BLUP) procedure and the calculation of average rank. All the analyses were carried out in R software (Team R, 2017).

A completely randomized design was used for the analysis of deviance (D), which may be written in the matrix form through equation [11]:

$$y = Xu + Zg + \varepsilon [11]$$

where: y is the vector of data observed for the variable to be analyzed; u is the estimate of the overall average through the best linear unbiased estimator (BLUE) method, assumed to be fixed; g is the vector of predicted genotypic effects, assumed to be random; # is the vector of random errors; and X and Z are the incidence matrixes for the fixed and random effects, respectively.

Based on model [11], the deviance (D) was calculated by equation [12] after the maximization of the restricted likelihood function [13] in an iterative process, with the Dempster, Laird, and Rubin (1977) EM algorithm, which used the difference between the broad-sense heritability estimates (1×10^{-5}) as a stopping criterion throughout the iterative steps (Resende, 2007):

$$D = -2\ln(L)[12]$$

$$\ln(L) = -\frac{1}{2}\ln|X'V^{-1}X| - \frac{1}{2}\ln|V| - \frac{1}{2}(y - Xu)' V^{-1}(y - Xu)[13]$$

where: ln(L) is the maximum point of the logarithmic function of maximum restricted likelihood; y is the vector of data observed from the variable to be analyzed; u is the estimate of the overall average through the best linear unbiased estimator (BLUE) method; X is the incidence matrix for the fixed effects; and V is the matrix of y variance and covariance.

For the estimation of likelihood ratio test (LRT) statistics, the following estimator [14] was used (Resende, 2007):

$$LTR = \left| -2\ln(\hat{L}_G) + 2\ln(\hat{L}_{MC}) \right| [14]$$

where: (L^$_G$) is the estimation algorithm for the maximum point of the restricted likelihood function for the reduced model (without the genotypic effects); and ln(L^$_{MC}$) is the estimation algorithm for the

maximum point of the restricted likelihood equation for the complete model (with the genotypic effects).

The error () and genotypic () variance components were predicted based on model [11]. The error () and genotypic () variances were obtained by the REML method, which corrects for the bias generated by the loss of degrees of freedom and always produces positive estimates for the variance components (Resende, 2007) and the estimates of other genetic parameters from the completely randomized model (Online Resource 2).

Selection of the best genotypes

The average rank, adapted from Mulamba and Mock (1978), was determined with the BLUP values of the genetic parameters. The average rank is a selection index based on the average of the individual ranks obtained by BLUP. Through this tool, it was possible to select genotypes based on multiple characters simultaneously.

Results and discussion

Due to the complexity of plant growth, several parameters must be considered to infer plant behavior, and growth analysis is one of the most accessible and precise tools for obtaining this knowledge. Therefore, understanding biological processes, such as growth, is essential for developing management plans that are suitable for each species. In this study, four nonlinear models were used with the aim of identifying which one had the best fit for describing the germination growth curves of 45 *E. edulis* accessions.

The accessions were clustered (considering the variables GSI, GAT, FGC, and GP) into four different groups based on the Tocher optimization criterion obtained from the Mahalanobis distance matrix (Online Resource 3). Tocher optimization is a simple clustering method that separates genotypes into groups such that the average distances within the groups are always lower than the average distances between the groups (Cruz et al., 2012). This method has been widely and successfully used in genetic divergence studies among accessions (Cantelli et al., 2016; Costa et al., 2016; Manuel et al., 2016; Hoogerheide et al., 2017).

We evaluated the quality of the four nonlinear models' adjustments with each of the four groups of accessions to select the model that provided the best fit for our type of data (Online Resource 4). The four nonlinear models were chosen because they are widely used to explain continuous data (Chatterjee, Chatterjee, Majumdar, & Chakrabarti, 2015; Tjørve & Tjørve, 2017), such as the ones provided by the seed germination growth curves studied here.

Regarding the parameters R^2, MSE, MAD, AIC, and BIC, which were used to describe the quality of the models, values of R^2 close to 1 and the lowest values of MSE, MAD, AIC and BIC are required for a model to

be considered the best. Due to the differences observed in all the studied models, the evaluation of nonlinear model fit based only on R^2 values may not be the best option because R^2 values were not able to discriminate these differences.

The Gompertz model is considered to have the best fit to our data because it presented the lowest evaluator values, which indicate the best model fit. Therefore, this was the model chosen to describe the germination pattern of *E. edulis* seeds. This model and its modifications have been applied to several types of studies such as those on bacterial growth curves (Chatterjee et al., 2015), the kinetics of biogas production (Yono, Syaichurrozi, & Sumardiono, 2014), pest management models (Tian, Zhang, & Sun, 2016), and tumor growth (Bolton, Alain, Cloot, & Schalk, 2015), and according to our results, we recommend that it be used in the study of seed germination patterns of other palm trees.

The von Bertalanffy model presented the highest values of β_1 for all the groups. However, the confidence intervals of the von Bertalanffy model for $\beta 1$ are the most elevated. In addition, the von Bertalanffy model presented the lowest R^2 value and the highest MSE, MAD, AIC, and BIC values, what indicates that this is the least appropriate model for describing the groups we examined (Online Resource 5).

Parameter ($_3$, which represents the average germination tax in the logistic, Gompertz and von Bertalanffy models, was most accurately estimated by the logistic model.

Germination growth curves were fit for the four *E. edulis* accession groups based on the chosen model (Gompertz). These curves showed the germination percentage (G) and germination tax (GT) as a function of time. Regarding the germination percentage, in general, accessions in Group 2 presented the lowest percentages (approximately 30%), and the accessions in Group 4 presented the highest percentages (approximately 90%). The germination taxes from Groups 1 and 2 were similar and constant through time. Group 3 presented an increase in the germination tax around the 20th day, then a decrease, and the tax remained constant thereafter. Group 4 also presented an increase in the germination tax, which occurred on the 30th day, before decreasing and remaining constant thereafter (Figure 1).
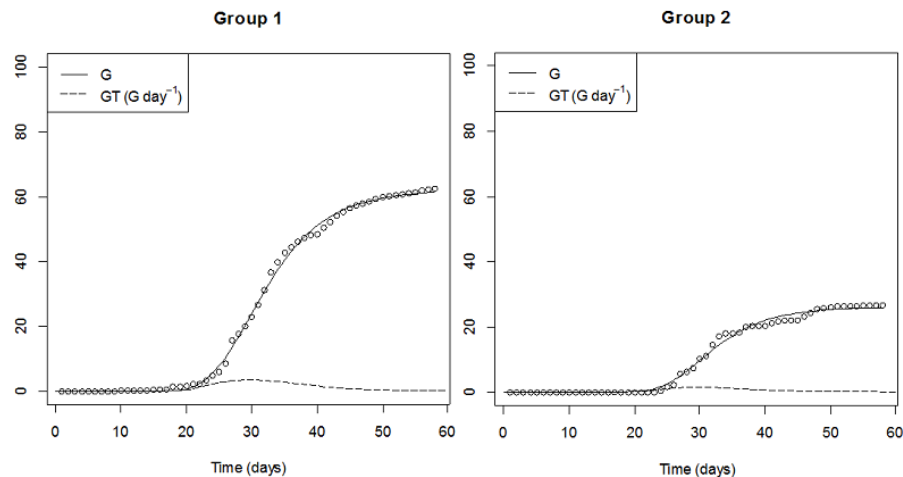
Figure 1

Germination percentage (G) and germination tax (GT) as a function of
time for the four *E. edulis* accession groups. Parameters were obtained
by the Gompertz growth model and its first derivative, respectively.

The vigor of seeds was estimated according to their maximum
germination tax. Groups 1, 2, 3, and 4 presented their maximum
germination points (MP) at 29.64, 29.86, 21.18, and 28.47 days,
respectively. Their maximum germination rates (MGRs) per day were
3.58, 1.66, 5.00, and 8.50, respectively (Table 1).

Table 1

Estimates of the maximum germination point (MP), maximum germination
rate (MGR) and inflection point (IP) of the germination percentage
curves as a function of time for the four *E. edulis* accession groups.

Groups	MP (days)	MGR (G day^{-1})	IP (G)
1	29.64	3.58	22.83
2	29.86	1.66	9.58
3	21.18	5.00	27.72
4	28.47	8.50	34.37

Seeds from groups 3 and 4 may be considered the most vigorous
because they had the highest daily germination rates. According to this
information, it is possible to perform a projection at the arboretum level,
at which the critical seed production period can be estimated, for example.
This would allow a schedule to be produced that would help determine
the time when the arboretum needs to be organized.

Variability in seed germination among accession groups was observed,
and similar results (i.e., variation among individuals) were found for other
palm trees (Berton, Filho, Siqueira, & Colombo, 2013; Silva-Cardoso &
Souza, 2017).

The analysis of deviance revealed significant differences among
genotypes for all the variables evaluated by a χ^2 test at a 1 % level of
probability (Table 2). The genetic variance of *E. edulis* seeds was elevated
for all the analyzed characteristics and higher than the environmental
variance (Table 3). The existence of genetic variability in germination and

vigor among the genotypes indicates the possibility for selection for these characters.

Selection is only possible and justifiable when there is a significant difference among genotypes. The higher the variability, the higher the probability of finding genotypes with favorable characteristics or that reunite these characteristics. Thus, the elevated genetic variability found in *E. edulis* for the measured germination parameters offers good opportunities to identify superior genotypes in terms of these characteristics.

Table 2

Analysis of deviance (ANADEV) for the germination speed index (GSI), first germination count (FGC), germination percentage (GP) and germination average time (GAT) performed with restricted maximum likelihood (REML) for the 45 *E. edulis* accessions.

Variation sources	GSI	FGC (%)	GP (%)	GAT
Genotypes	-23.995[+]	1325.404[+]	1348.940[+]	874.882[+]
Complete model	-192.945[++]	1071.309[++]	1126.075[++]	618.855[++]
LRT (χ^2)	168.950**	254.095**	222.866**	256.027**

LRT - likelihood ratio test, +Deviance of the model adjusted without the referenced effect, ++Deviance of the adjusted complete model, ** Significant based on a $\chi2$ test with 1 degree of freedom at a 1% level of probability.

Table 3

Estimates of genetic parameters for the germination speed index (GSI), first germination count (FGC), germination percentage (GP) and germination average time (GAT) obtained with restricted maximum likelihood (REML) for the 45 *E. edulis* accessions.

Parameters	GSI	FGC (%)	GP (%)	GAT
$\hat{\sigma}_g^2$	0.257	538.243	595.745	43.512
$\hat{\sigma}_e^2$	0.060	58.045	84.031	4.617
$\hat{\sigma}_f^2$	0.317	596.288	679.776	48.128
$\hat{h}_g^2$	0.812	0.903	0.876	0.904
*CVg (%)	63.640	97.845	38.431	26.370
*CVe (%)	30.610	32.132	14.433	8.589
*CVg/CVe	2.079	3.045	2.663	3.070
Overall average	0.797	23.711	63.511	25.015
Accuracy	0.961	0.976	0.972	0.976

CVg = coefficient of variation expressed by genotype; CVg = coefficient of variation expressed by environment.

This fact can be confirmed by the heritability, which presented a high magnitude for all the evaluated indexes, indicating good genetic control of the expression of the seed characteristics, i.e., the higher the genetic variation relative to the environmental variation, the higher the heritability. Therefore, the elevated heritability values described for the characters may be related to the high genetic variability in the studied population.

The genetic variation coefficient, which expresses the genetic fraction of the variance measured in the experiment as a percentage (Vasconcelos, Reis, Sediyama, & Cruz, 2012), was 63.640, 97.845, 38.431, and 26.370% for the variables GSI, FGC, GP, and GAT, respectively, which shows the high genetic variability of these characteristics in the population, especially for GSI and FGC (Table 3). Thus, there are good opportunities for genetic advancement via genotype selection for higher germination percentages with faster and more uniform germination.

The high heritability of the evaluated indexes may also be related to the fact that maternal genetic effects determine several seed characteristics, which are important determinants of germination and dormancy. In angiosperms, these effects may arise from the maternal inheritance of plastids; from the additional maternal genetic contribution to the endosperm; from the seed coat and other maternal tissues that envelop the embryo; and from hormones, proteins, transcripts and nutrients passed from the mother plant to the seeds during their development (Laossi, Noguera, & Barot, 2010; Rix et al., 2012).

Therefore, even for allogamous species, seeds from the same plant present similar germination characteristics inherited from the mother plant that are different from those presented by seeds from other plants (Laurentin & Benítez, 2014). This increases genetic variation and decreases the environmental effect, generating elevated heritability values. High-magnitude heritability values were also observed in *Acrocomia aculeata* (Berton et al., 2013), *Lupinus angustifolius* (Beyer et al., 2015) and soybean germinative characters (Vasconcelos et al., 2012).

The CVg/CVe ratio was 2.079, 3.045, 2.663, and 3.070 for the parameters GSI, FGC, GP, and GAT, respectively. If the ratio is greater than or equal to 1, as it was for the four evaluated characteristics, then the genetic variation represents the higher part of the CV of the estimated values, making the selection process easy (Yokomizo & Farias Neto, 2003). In other words, the CVg/CVe ratio demonstrates how much of the overall variance is caused by the genotype (Vasconcelos et al., 2012) and is an indication of the easiness of genotype selection for the desired characteristic (Yokomizo & Farias Neto, 2003). The CVg/CVe values in this study indicate that there is sufficient genetic variation to facilitate the selection of superior plants based on the measured characteristics.

The measure of accuracy that reports the correlation between the true genotypic value and either the estimated genotypic value or the one predicted through experiments was very high for all the characteristics (between 0.961 and 0.976), which indicates the high quality of the experiment and guarantees precise selection of the evaluated and selected genotypes in the studied population.

According to the predicted genotypic values of the variables GSI, FGC, GP, and GAT of *E. edulis* accessions obtained through the BLUP method, the GP values varied from 10.85 to 94.89% (Online Resource 6). Genotype AL1P09, which presented the highest GP, presented low performance in terms of the other evaluated characteristics. In AL3P03, 93.92% of seeds germinated, and this genotype had average performance in terms of the other characteristics. Genotype GU1P06 had a GP of 90.06% and presented good performance in terms of the other characteristics.

Genotype AL1P02 presented the best values for GSI and GAT, followed by MI1P01 and GU1P06. Genotypes present similar results for these characteristics because, generally, the higher the germination speed, the lower the germination time.

Genotypes MI1P01 and GU1P06 also presented elevated germination percentages in the first count (FGC), with values of 82.417 and 70.732%, respectively, and AL3P01 reached a value of 73.653%.

Fragments AL1, AL3, GU1, and MI1 contained genotypes with the best and worst performances, with similar averages, whereas GU1 presented the best averages of all the characteristics. Genotypes from fragment IB1 were positioned, for the most part, between the ones with the worst performance, and the fragment had the worst averages of the characteristics.

The elevated population variability indicates that the genotypes exhibited high, average and low performance for the evaluated characteristics, and the ones with the best performance are targets for selection. The genotypes with high performance in terms of GP presented very high values, considering the higher values reported in the literature are approximately 70 and 80% (Tavares et al., 2008; Cavalcante et al., 2012).

The average rank index (Mulamba & Mock, 1978) was used to select genotypes based on all the evaluated characteristics, aiming to obtain simultaneous gains, which allows the most equilibrated genotypes to be selected. Genotype GU1P06 presented the best performance, with a genetic gain of 557.14%. Genotypes MI1P01, GU1P04, and GU1P03 also presented high performance, with genetic gains of over 400%.

A total of 27 genotypes presented selection gains of over 50%, and among the ten best genotypes, five belong to fragment GU1, indicating the presence of favorable alleles in this population involved in the genetic control of the studied variables (Online Resource 7).

In addition to its low percentage, *E. edulis* seed germination is considered slow and ununiform. Tavares et al. (2008) reported a very low GSI value (0.15), which may be related to the genotype used. Cavalcante et al. (2012) observed a GSI of 1.6, which is lower than the values identified for genotypes AL1P02 and MI1P01 and close to the value for genotype GU1P06, which exhibited good performance in terms of all the characteristics.

However, the selection of genotypes based on all the evaluated characteristics is fundamental for species breeding involving germination. Selection based on one unique characteristic, in spite of leading to a superior final product measured based only on this characteristic, may lead to low performance in terms of the others (Cruz et al., 2012). Thus, the average rank index promoted genotype selection based on good germination percentage and faster and more uniform germination, which are the most-reported problems for juçara seeds.

Conclusion

The methodology proposed in the present work may be considered appropriate for classifying nonlinear models with the best fit. The Gompertz model most accurately described the germination curve of *E. edulis* seeds.

In general, the groups presented variable germination rates. Group 4, however, had the highest germination rate and was considered to be the most promising and vigorous group.

High genetic variability and high heritability were found for *E. edulis* accessions, which makes the selection of genotypes possible. However, the selection of genotypes must be based on as many characters as possible, which is fundamental for species breeding involving germination.

Acknowledgements

References

Akaike, H. A. (1974). New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723. DOI: 10.1109/TAC.1974.1100705

Berton, L. H. C., Filho, J. A. A., Siqueira, W. J., & Colombo, C. A. (2013). Seed germination and estimates of genetic parameters of promising macaw palm (*Acrocomia aculeata*) progenies for biofuel production. *Industrial Crops Products*, *51*, 258-266. DOI: 10.1016/j.indcrop.2013.09.012

Beyer, H., Schmalenberg, A. K., Jansena, G., Jürgens, H. U., Uptmoor, R., Broer, I., ... Ordon, F. (2015). Evaluation of variability, heritability and environmental stability of seed quality and yield parameters of *L. angustifolius*. *Field Crops Research*, *174*, 40-47. DOI: 10.1016/j.fcr.2014.12.009

Björck, A. (1996). *Numerical methods for least squares problems*. Philadelphia, US: Siam.

Bolton, L., Alain, H. J. J., Cloot, W., & Schalk, W. (2015). Schoombie. A proposed fractional-order Gompertz model and its application to tumour growth data. *Mathematical Medicine and Biology*, *32*(2), 187-207. DOI: 10.1093/imammb/dqt024

Brasil. Ministério da Agricultura e Reforma Agrária, Secretaria Nacional de Defesa Agropecuária. (2009). *Regras para análise de sementes*. Brasília, DF: MAPA.

Cantelli, D. A. V., Hamawaki, O. T., Rocha, M. R., Nogueira, A. P. O., Hamawaki, R. L., Sousa, L. B., & Hamawaki, C. D. L. (2016). Analysis of the genetic divergence of soybean lines through hierarchical and Tocher optimization methods. *Genetics and Molecular Research*, *15*(4), 1-13. DOI: 10.4238/qmr.15048836

Casas, R. R., Willis, C. G., Pearse, W. D., Baskin, C. C., Baskin, J. M., & Cavender-Bares, J. (2017). Global biogeography of seed dormancy is determined by seasonality and seed size: a case study in the legumes. *New Phytologist*, *214*, 1527-1536. DOI: 10.1111/nph.14498

Cavalcante, M. Z. B, Pivetta, K. F. L., Iha, L. L., & Takane, R. J. (2012). Temperatura, escarificação mecânica e substrato na germinação de

sementes das palmeiras juçara e açaí. *Brazilian Journal of Agricultural Science*, 7(4), 569-573. DOI: 10.5039/agraria. v7i4a1684.

Chatterjee, T., Chatterjee, B. K., Majumdar, D., & Chakrabarti, P. (2015). Antibacterial effect of silver nanoparticles and the modeling of bacterial growth kinetics using a modified Gompertz model. *Biochimica Biophysica Acta*, 1850(2), 299-306. DOI: 10.1016/j.bbagen.2014.10.022.2015

Costa, R. B., Silva, J. C., Skowronski, L., Constantino, M., Pistori, H., & Pinto, J. V. C. (2016). Genetic Divergence in *Eucalyptus camaldulensis* progenies in the Savanna Biome in Mato Grosso, Brazil. *PloS ONE*, 11(9), e0163698. DOI: 10.1371/ journal.pone.0163698

Cruz, C. D. (2013). GENES - a software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum. Agronomy*, 35(3), 271-276. DOI: 10.4025/actasciagron.v35i3.21251

Cruz, C. D., Regazzi, A. J., & Carneiro, P. C. S. (2012). *Modelos biométricos aplicados ao melhoramento genético*. (4a ed.). Viçosa, MG: UFV.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1), 1-38. DOI: 10.1.1.133.4884

El-Keblawy, A., Shabana, H. A., Navarro, T., & Soliman, S. (2017). Effect of maturation time on dormancy and germination of *Citrullus colocynthis* (Cucurbitaceae) seeds from the Arabian hyper-arid deserts. *BMC Plant Biology*, 17(263), 1-10. DOI: 10.1186/s12870-017-1209-x

Frank, A., Pluess, A. R., Howe, G. T., Sperisen, C., & Heiri, C. (2017). Quantitative genetic differentiation and phenotypic plasticity of European beech in a heterogeneous landscape: Indications for past climate adaptation. *Perspectives in Plant Ecology, Evolution and Systematics*, 26(1), 1-13. DOI: 10.1016/j.ppees.2017.02.001

Ganem, R. S. (2011). *Conservação da biodiversidade*: legislação e políticas públicas. Brasília, DF: Câmara dos Deputados.

Henderson, A. (2000). The genus *Euterpe* in Brazil. *Sellowia*, 49(52), 1-22.

Hoogerheide, E. S. S., Filho, J. A. A., Vencovsky, R., Zucchi, M .I., Zago, B. W., & Pinheiro, J. B. (2017). Genetic variability of garlic accessions as revealed by agro-morphological traits evaluated under different environments. *Genetic Molecular Research*, 16(2). DOI: 10.4238/gmr16029612.

Inácio, M. R. C., Lima, K. M. G., Lopes, V. G., Pessoa, J. D. C., & Teixeira, A. G. H. (2013). Total anthocyanin content determination in intact açaí (*Euterpe oleracea* Mart.) and palmitero-juçara (*Euterpe edulis* Mart.) fruit using near infrared spectroscopy (NIR) and multivariate calibration. *Food Chemistry*, 136(3), 1160-1164. DOI: 10.1016/j.foodchem.2012.09.046

Kennedy, P. (2008). *A guide to econometrics*. Massachusetts, US: Blackwell Publishing.

Laossi, K. R., Noguera, D. C., & Barot, S. (2010). Earthworm-mediated maternal effects on seed germination and seedling growth in three annual plants. *Soil Biology Biochemistry*, 42(2), 319-323. DOI: 10.1016/ j.soilbio.2009.11.010

Laurentin, H., & Benítez, T. (2014). Inheritance of seed coat color in sesame. *Pesquisa Agropecuária Brasileira*, 49(4), 290-295. DOI: 10.1590/ S0100-204x2014000400007

Manuel, B. L. R., Chagas, E. A., Smiderle, O., Carlos, A. R., Chagas, P. C., Mota Filho, A. B., ... Cordeiro, A. C. C. (2016). Genetic divergence among Camu-Camu plant populations based on the initial characteristics of the plants. *Journal of Agricultural Science*, 8(11), 51-58. DOI: 10.5539/jas.v8n11p51

Martins, C. C., Bovi, M. L. A., Nakagawa, J., & Machado, C. G. (2009). Secagem e armazenamento de sementes de juçara. *Revista Árvore*, 33(4),635-642. DOI: 10.1590/S0100-67622009000400006

Mello, R., Magalhaes, A. L. R., Breda, F. C., & Regazzi, A. J. (2008). Modelos para ajuste da produção de gases em silagens de girassol e milho. *Pesquisa Agropecuária Brasileira* , 43(2), 261-269. DOI: 10.1590/S0100-204X2008000200016.

Mulamba, N. N., & Mock, J. J. (1978). Improvement of yield potential of the ETO blanco maize (*Zea mays* L.) population by breeding for plant traits. *Egyptian Journal of Genetics and Cytology*, 7(1), 40-51.

Paine, C. E. T., Marthews, T. R., Vogt, D. R., Purves, D., Rees, M., Hector, A., & Turnbull, L. A. (2012). How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. *Method in Ecology and Evolution*, 3(1), 245-256. DOI: 10.1111/j.2041-210X.2011.00155.x

Penfield, S., & MacGregor, D. R. (2017). Effects of environmental variation during seed production on seed dormancy and germination. *Journal of Experimental Botany*, 68(4), 819-825. DOI: 10.1093/jxb/erw436

Piepho, H. P., Möhring, J., Melchinger, A. E., & Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*, 161(1), 209-228. DOI: 10.1007/s10681-007-9449-8

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., & Joppa, L. N. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, 344(6187), e1246752. DOI: 10.1126/science.1246752

Resende, M. D. V. (2007). *Matemática e estatística na análise de experimentos e no melhoramento genético*. Colombo, PR: Embrapa Florestas.

Rix, K. D., Gracie, A. J., Potts, B. M., Brown, P. H., Spurr, C. J., & Gore, P. L. (2012). Paternal and maternal effects on the response of seed germination to high temperatures in *Eucalyptus globulus*. *Annals of Forest Science* , 69(6), 673-679. DOI: 10.1007/s13595-012-0190-7

Sarmento, J. L. R., Regazzi, A. J., Souza, W. H., Torres, R. A., Breda, F. C., & Menezes, G. R. O. (2006). Estudo da curva de crescimento de ovinos Santa Inês. *Revista Brasileira de Zootecnia*, 35(2), 435-442. DOI: 10.1590/S1516-35982006000200014

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461-464.

Silva-Cardoso, I. M. A.; & Souza, A. M. (2017). The palm tree *Syagrus oleracea* Mart. (Becc.). A review. *Scientia Horticulturae*, 225, 65-73. DOI: 10.1016/j.scienta.2017.06.054

Tavares, A. R., Ramos, D. P., Aguiar, F. F. A., & Kanashiro, S. (2008). Jussara palm seed germination under different shade levels. *Horticultura Brasileira*, 26(4), 492-494. DOI: 10.1590/S0102-05362008000400013

Team R. (2014). *A language and environment for statistical computing*. Vienna, AU: R Foundation for Statistical Computing. Retrieved on Feb. 20, 2014 from http://www.R-project.org/.

Teixeira, M. T., Vieira, H. D., Partelli, F. L., & Silva, R. F. (2011). Despolpamento, armazenamento e temperatura na germinação de sementes de palmeira real australiana. *Pesquisa Agropecuária Tropical*, *41*(3), 378-384. DOI: 10.5216/pat.v41i3.9617

Tian, Y., Zhang, Z., & Sun, K. (2016). Qualitative Analysis of a Pest Management Gompertz Model with Interval State Feedback Impulsive Control. *Discrete Dynamics in Nature and Society*, *2016*, 1-10. DOI: 10.1155/2016/4294595

Tjørve, C. M. C., & Tjørve, E. (2017). The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family. *PloS ONE* , *12*(6), 1-17. DOI: 10.1371/journal.pone.0178691

Vasconcelos, E. S., Reis, M. S., Sediyama, T., & Cruz, C. D. (2012). Estimativas de parâmetros genéticos da qualidade fisiológica de sementes de genótipos de soja produzidas em diferentes regiões de Minas Gerais. *Semina: Ciências Agrárias*, *33*(1), 65-76. DOI: 10.5433/1679-0359.2012v33n1p65

Yokomizo, G. K., & Farias Neto, J. T. (2003). Caracterização fenotípica e genotípica de progênies de pupunheira para palmito. *Pesquisa Agropecuária Brasileira* , *38*(1), 67-72. DOI: 10.1590/S0100-204X2003000100009

Yono, B., Syaichurrozi, I., & Sumardiono, S. (2014). Kinetic model of biogas yield production from vinasse at various initial pH: Comparison between modified Gompertz model and first order kinetic model. *Research Journal of Applied Sciences, Engineering and Technology*, 7(13), 2798-2805. DOI: 10.19026/rjaset.7.602

Author notes

*

Author for correspondence. E-mail: carolinabernardes84@yahoo.com.br