Silva dos Santos, Vinicius; Martins Filho, Sebastião; Moysés Alves, Rafael; Deon Vilela de
Resende, Marcos; Fonseca e Silva, Fabyano
Genetic divergence among cupuaçu accessions by multiscale bootstrap resampling

Genetic divergence among cupuaçu accessions by multiscale bootstrap resampling

Vinicius Silva dos Santos ([1]*); Sebastião Martins Filho ([1]); Rafael Moysés Alves ([2]);
Marcos Deon Vilela de Resende ([3]); Fabyano Fonseca e Silva ([4])

([1]) Universidade Federal de Viçosa (UFV), Departamento de Estatística, Av. Peter Henry Rolfs, s/n, 36570-900 Viçosa (MG), Brasil.

([2]) Embrapa Amazônia Oriental, Caixa Postal 48, 66095-100 Belém (PA), Brasil.

([3]) Embrapa Florestas, Caixa Postal 319, 83411-000 Colombo (PR), Brasil.

([4]) UFV, Departamento de Ciência Animal, 36570-900 Viçosa (MG), Brasil.

(*) Corresponding author: 2santosvinicius@gmail.com

Abstract

This study aimed at investigating the genetic divergence of eighteen accessions of cupuaçu trees based on fruit morphometric traits and comparing usual methods of cluster analysis with the proposed multiscale bootstrap resampling methodology. The data were obtained from an experiment conducted in Tomé-Açu city (PA, Brazil), arranged in a completely randomized design with eighteen cupuaçu accessions and 10 repetitions, from 2004 to 2011. Genetic parameters were estimated by restricted maximum likelihood/best linear unbiased prediction (REML/BLUP) methodology. The predicted breeding values were used in the study on genetic divergence through *Unweighted Pair Cluster Method with Arithmetic Mean* (UPGMA) hierarchical clustering and Tocher's optimization method based on standardized Euclidean distance. Clustering consistency and optimal number of clusters in the UPGMA method were verified by the cophenetic correlation coefficient (CCC) and Mojena's criterion, respectively, besides the multiscale bootstrap resampling technique. The use of the clustering UPGMA method in situations with and without multiscale bootstrap resulted in four and five clusters, respectively, while the Tocher's method resulted in seven clusters. The multiscale bootstrap resampling technique proves to be efficient to assess the consistency of clustering in hierarchical methods and, consequently, the optimal number of clusters.

Key words: clustering, UPGMA, tocher, *Theobroma grandiflorum*.

1. INTRODUCTION

Theobroma grandiflorum (Willd. ex Spreng.) Schum. (cupuaçu) is a fruitful tree native to the Amazon region that stoods out in agroindustry, due to the demand for its fruit pulp, which is widely used in the cuisine of Pará state (PA, Brazil) (Venturieri, 2011).

The cupuaçu is an allogamous species, self-incompatible, with hermaphrodite flowers, whose fertilization occurs, beyond the stigma, along of the style. In the Amazonian region, the flowering occurs in period of July to December, the driest period of the year, and the fruits production from August to April, with a major peak from January to March, the local rainy season (Prance & Silva, 1975). Therefore, the specie has a strong interaction with the environment, because is in the dry season which the activity of pollinating insects is more intense, while during the rainy season occurs greater physiological demand by water, for development and maturation of the fruits.

The study of genetic diversity has been fundamental in cupuaçu genetic breeding programs, as it allows assessing and selecting superior genotypes for fruit production and resistance to diseases, such as the witches' broom. This disease is caused by the *Crinipellis (Moniliophthora) perniciosa* fungus and is responsible for great yield losses observed in recent years (Alves et al., 2010). Studies on the genetic diversity of cupuaçu were carried out by Alves et al. (2013), Araújo et al. (2002) and Maia et al. (2011a, b).

Since the cupuaçu is a perennial species with long reproductive cycle, the optimal selection procedure is that involving the estimation of variance components by restricted maximum likelihood (REML) method and the prediction of genotypic values by best linear unbiased prediction (BLUP) (Resende, 2007b).

Several multivariate techniques can be employed in genetic divergence prediction, mainly the principal components analysis, canonical variables and the hierarchical clustering and optimization methods (Cruz et al., 2012). In agglomerative hierarchical methods, two questions are addressed: how to evaluate cluster consistency and how to determine the final optimal number of clusters,

when using hierarchical agglomerative methods (Suzuki & Shimodaira, 2006).

Shimodaira (2002) proposed the use of the multiscale bootstrap resampling technique to evaluate these questions. In bootstrap resampling, the initial data size n is used in a sampling process with replacement, producing pseudo-samples with the same size as the original ones, denominated bootstrap samples. In the multiscale bootstrap resampling method, the sample size is reduced, equaled or augmented, considering the size n of the initial sample (Shimodaira, 2002; Suzuki & Shimodaira, 2006).

The algorithm is similar to that of bootstrap resampling. After being obtained, the usual bootstrap probability values (BP) are fitted by a theoretical equation for p-value calculation denominated approximately unbiased (AU). P-value of a cluster is a value between 0 and 1, which indicates how strong the cluster is supported by data. The AU p-value obtained by multiscale bootstrap resampling corrects the selection bias of the p-value obtained by standard bootstrap resampling (Suzuki & Shimodaira, 2006). If this value is lower than a significance level α, for a certain cluster, there are signs that the individuals that compose the cluster are not similar. The multiscale bootstrap resampling technique was used in genetic diversity studies of *Trifolium* (Rizza et al., 2007) and of *Nipponia nippon* (Taniguchi et al., 2013), among others.

This study aimed to analyze the genetic divergence of eighteen cupuaçu accessions through methods of cluster analysis, comparing the outcome obtained by the hierarchical UPGMA agglomerative method with and without the application of the multiscale bootstrap resampling technique and by the Tocher's optimization method.

2. MATERIAL AND METHOD

The experiment was installed in February 1999, at the Embrapa Oriental Amazon region unit, located in Tomé–Açu, city (PA, Brazil), at approximately 01°57'38" and 03°16'37" south latitude, and 47°53'32" and 48°49'15" west longitude, at 45m altitude. Eighteen cupuaçu accessions were evaluated in a completely randomized design with ten repetitions. The data were obtained based on the average of five fruits per plant in the growing seasons, from 2004 to 2011. The following variables were evaluated: fruit length in mm (FL); fruit diameter in mm (FD); fruit weight in g (FW); pulp weight in g (PW); rind weight in g (RW); rind thickness in mm (RT); mean weight of seeds per fruit in g (MSW) and seed number (SN).

The data were initially analyzed by REML/BLUP (restricted maximum likelihood/best linear unbiased prediction) mixed models methodology, with the adoption of the following statistical model (Resende, 2007b):

$$y = Xb + Zg + e \qquad (1)$$

in which y is the data vector, b is the scalar referring to the general average (fixed effect), g is the vector of genotypic effects (assumed as random) and e is the vector of random errors. X and Z are incidence matrices for b and g, respectively. Based on treatments with random effects models, the estimated breeding values must be used to replace phenotypic values (Resende, 2007b). The Selegen-Reml/Blup software, version 2009 (Resende, 2007a), was used to estimate the variance and prediction components of breeding values. Multicollinearity detection was carried out based on the genetic correlation matrix, according to Montgomery and Peck criterion (Cruz & Carneiro, 2006).

For the genetic divergence study, the estimated breeding values were used in the linear mixed model and submitted to standardization, in order to obtain a variable with a mean of zero and standard deviation of 1. The hierarchical method of *Unweighted Pair Cluster Method with Arithmetic Mean* (UPGMA) and the Tocher's optimization method were employed based on the Eucleadian distance matrix.

To evaluate clustering adaptation and consequently the optimal number of clusters in the hierarchical method, the results of the analyses were compared without (situation 1) and with the multiscale bootstrap resampling methodology (situation 2). In the usual analysis (situation 1), clustering consistency was verified by the cophenetic correlation coefficient (CCC), in which the obtained value was evaluated by Mantel's test at 5% significance level and with 1000 permutations. The separation of clusters through the cut-off point in the dendrogram was defined based on Mojena's criterion, with the value k=1.25 (Milligan & Cooper 1985). These analyses were performed with the use of the Genes software, version 2012.30.1 (Cruz, 2013).

In the multiscale bootstrap method (situation 2), samples were produced with replacement of the original sample X_n by several values in which r_k, for k = 1, 2, 3..., 10, was fixed in values: $r_1 = 0.5$; $r_2 = 0.6$; $r_3 = 0.7$;...; $r_{10} = 1.4$. The probability values AU and BP were obtained through the function implemented in the "pvclust" package (Suzuki & Shimodaira, 2014) of the free R software (R Development Core Team, Vienna, AT). For each scale value r_k, B=10000 replications were produced and AU probability values were compared at 5% significance level. The Tocher's optimization method was also employed aiming at comparing the results obtained in situations 1 and 2 of the UPGMA method. The method adopts the criterion that the mean intra-cluster distance must be lower than the mean inter-cluster distance (Cruz & Carneiro, 2006).

3. RESULTS AND DISCUSSION

Estimates of genetic parameters are presented in table 1. In general, it was observed low to medium genetic variability among accessions, within each evaluated trait, according to

Table 1. Variance component estimates for variables fruit length (FL, mm), fruit diameter (FD, mm), fruit weight (FW, g), rind thickness (RT, g), rind weight (RW, g), mean weight of seeds per fruit (MSW, g), seed number (SN) and pulp weight (PW, g) of 18 experimental cupuaçu clones

Estimates[1/]	FL(mm)	FD(mm)	FW(g)	RT(mm)	RW(g)	MSW(g)	SN	PW(g)
Vg	163.96	50.45	19206.30	0.58	6573.55	203.56	3.19	5472.06
Ve	305.51	38.40	47294.10	0.86	12684.40	1162.62	26.11	10058.70
Vf	469.47	88.85	66500.40	1.45	19258.00	1366.18	29.30	15530.70
h²g	0.35	0.57	0.29	0.40	0.34	0.15	0.11	0.35
CVgi (%)	6.44	6.27	10.86	10.04	13.93	8.16	6.34	15.39
CVe (%)	8.79	5.47	17.04	12.2	19.34	19.5	18.16	20.87

[1/]Vg: genotypic variance. Ve: residual variance; Vf: individual phenotypic variance; h²g: broad-sense individual heritability coefficient, i.e., total genotypic effects; CVgi (%): genotypic variation coefficient; CVe (%): residual coefficient of variation.

the coefficients of genetic variation (CVgi%). This coefficient ranged from 6% to 15%. This low genetic variability may be associated with the origin of the accessions, since all of them come from the same commercial orchard, in the municipality of Tomé-Açu (PA, Brazil).

According to Laviola et al. (2011), CVgi% values lower than 10% indicate low genetic variability, whereas values above 20% indicate considerable genetic variability. The obtained values were lower than those reported by Maia et al. (2011b) in cupuaçu progenies for the same variables under study. In studies on progenies, Alves & Resende (2008) reported coefficients of genetic variation ranging from 27% to 88% at progeny level, and from 38% to 123% at individual level, which evidences excellent selection possibilities in the studied population. Whereas for the trait witches'-broom resistance, Alves et al. (2009) obtained CVgi% values ranging from 5% to 41%. In studies with clones carried out by Maia et al. (2011a), the coefficients of genetic variation ranged from 5% to 22% for the same traits evaluated in this study.

The estimates of the residual coefficient of variation (CVe%) were considered of low magnitude for all evaluated traits (Table 1), which demonstrates good experimental precision. The CVgi/CVe ratio was higher than the unit only for the character fruit diameter. Thus, this trait might be used without precautions in the cupuaçu breeding program (Laviola et al., 2011).

Low genetic variability associated with high experimental precision led to moderate and high individual broad-sense heritability values, for the traits fruit length (35%), fruit weight (29%), rind thickness (40%), pulp weight (35%) and fruit diameter (57%), respectively, indicating considerable genetic gain in the selection process using such descriptors. It is worth emphasizing that the magnitude of deviations did not lead to null heritability estimates, which makes them more reliable (Maia et al., 2011b).

The estimated individual broad-sense heritability was considered of low magnitude for the characters mean weight of seeds per fruit (15%) and number of seeds (11%), which demonstrates that these traits are more influenced by the

environment. Similar results were found by Maia et al. (2011a) in cupuaçu clones. In studies with progenies, Maia et al. (2011b) described mean heritability values ranging between 25% and 67%, and individual narrow-sense heritability values ranging between 16% and 64% for the variables of characterization and fruit production of cupuaçu. Alves & Resende (2008) reported individual narrow-sense heritability values, in one growing season, ranging from 25 to 54%.

Regarding the components of the predicted average (individual BLUPs), it is observed in table 2 that the access 378 was the most prominent, with higher predicted genotypic values than the overall average for all traits. This access can be selected in order to produce seeds, where the butter will be extracted for cupulate manufacturing, since that, figured in the first and third position in the ranking for the variables weight and number of seeds, respectively. It is interesting, also, for pulping, because occupied the third position in this classification. Access 371 had the worst performance, with the lowest genotypic value for the characters fruit diameter, fruit weight and pulp weight.

In addition to access 378, accessions 382, 387, 415, 402, 372 and 425 also showed superiority as to fruit weight, with genotypic values larger than the overall average. These accessions were also important as to the pulp weight, which indicates a strong positive correlation between these variables, which was confirmed in additional analyzes with the value of genetic correlation of 0.89. Accessions 382 and 415 were those who had larger genotypic value for the variable pulp weight. However, these same accessions showed lower overall average genotypic values for the traits weight and number of seeds, to be recommended for when the focus for improvement is directed specifically to increase of pulp yield.

The multicollinearity detection based on the genetic correlation matrix revealed condition number (NC) equal to 289.87. According to the Montgomery and Peck criterion (Cruz & Carneiro, 2006), NC lower than 100 indicates weak multicollinearity, between 100 and 1000, moderate to low-severe multicollinearity and higher than 1000, severe multicollinearity. If it is observed multicollinearity

Table 2. Prediction of the components of average (BLUP individual) of 18 cupuaçu clones for variables: fruit length (FL), fruit diameter (FD), fruit weight (FW), rind thickness (RT), rind weight (RW), mean weight of seeds per fruit (MSW), seed number (SN) and pulp weight (PW). The values in parentheses represent the rank of clones in descending order of predicted genetic values for each variable

Clone*	FL(mm)	FD(mm)	FW(g)	RT(mm)	RW(g)	MSW(g)	SN	PW(g)
378 (2.4)	225.5 (1)	116.8 (5)	1485.3 (1)	8.0 (4)	721.9 (1)	201.3 (1)	29.5 (3)	550.6 (3)
402 (4.9)	216.2 (2)	111.0 (12)	1355.6 (5)	8.4 (3)	656.3 (5)	187.6 (2)	30.5 (2)	486.0 (8)
387 (5.4)	196.3 (10)	125.6 (1)	1446.1 (3)	9.0 (1)	705.0 (2)	177.2 (8)	27.6 (13)	517.6 (5)
382 (5.5)	205.3 (5)	124.4 (3)	1461.7 (2)	7.7 (9)	663.8 (3)	173.2 (9)	27.7 (12)	605.6 (1)
372 (6.9)	201.5 (8)	113.2 (7)	1342.1 (6)	7.0 (14)	558.2 (11)	186.5 (4)	30.5 (1)	548.4 (4)
425 (7.5)	187.8 (15)	121.9 (4)	1330.2 (7)	7.4 (11)	595.1 (7)	181.3 (5)	28.7 (5)	517.6 (6)
375 (7.9)	207.2 (4)	111.6 (9)	1262.5 (9)	7.5 (10)	562.5 (9)	180.6 (6)	28.4 (7)	485.5 (9)
415 (8.3)	191.8 (13)	124.6 (2)	1414.7 (4)	7.8 (7)	657.8 (4)	162.7 (16)	25.3 (18)	583.4 (2)
406 (10.4)	203.8 (7)	112.1 (8)	1255.8 (10)	7.9 (6)	606.9 (6)	163.6 (15)	26.0 (17)	444.7 (14)
366 (11.3)	194.5 (11)	111.1 (11)	1254.8 (11)	7.0 (13)	538.8 (14)	167.4 (13)	28.0 (10)	495.4 (7)
412 (11.3)	193.8 (12)	111.5 (10)	1275.7 (8)	8.0 (5)	561.4 (10)	161.8 (17)	27.5 (15)	448.3 (13)
367 (11.4)	180.9 (18)	113.5 (6)	1172.8 (14)	6.7 (17)	478.3 (18)	186.6 (3)	28.9 (4)	474.1 (11)
374 (11.6)	204.6 (6)	108.5 (15)	1197.7 (13)	7.3 (12)	542.7 (13)	171.6 (11)	27.9 (11)	452.2 (12)
363 (11.9)	210.2 (3)	105.8 (17)	1216.1 (12)	6.9 (15)	531.7 (15)	177.9 (7)	27.4 (16)	474.5 (10)
377 (12.6)	183.5 (17)	108.2 (16)	1128.1 (17)	8.9 (2)	579.4 (8)	157.5 (18)	28.4 (6)	368.0 (17)
405 (13.4)	190.6 (14)	108.5 (14)	1154.5 (15)	7.8 (8)	554.5 (12)	166.9 (14)	27.6 (14)	407.7 (16)
407 (14.3)	187.1 (16)	109.7 (13)	1137.3 (16)	6.6 (18)	482.2 (17)	173.2 (10)	28.1 (9)	432.7 (15)
371 (14.4)	200.4 (9)	101.3 (18)	1078.8 (18)	6.9 (16)	483.4 (16)	170.8 (12)	28.4 (8)	358.2 (18)
General average	198.9459	113.2916	1276.105	7.6152	582.2213	174.8682	28.1432	480.585

*Average rank.

in moderate to high and severe levels, then the highest correlated variables should be excluded. In the present study, the variables indicated for exclusion were fruit weight (FW) and rind weight (RW). Therewith, NC decreased to 94.02, which indicates weak multicollinearity. This allowed appropriate clustering.

In studies on beans, Cargnelutti et al. (2009) concluded that the presence of multicollinearity changed the clustering standards, and it was necessary to deal with the effects. The authors also mention that there are other forms of dealing with multicollinearity, such as using the Mahalanobis distance as dissimilarity measure. However, the function implemented in the "pvclust" package does not encompass this distance. Thus, it is necessary to carry out a more detailed study on its hierarchical clustering use by multiscale bootstrap resampling.

The estimated breeding values of cupuaçu accessions were used in the genetic divergence analysis through UPGMA hierarchical clustering methods, without (situation 1) and with (situation 2) multiscale bootstrap resampling and Tocher's optimization method. According to the matrix of Euclidean distances, the most divergent accessions were 377 and 378, whose distance was 6.289. But accessions 405 and 412 showed the smallest distance (0.973).

In situation 1, the formation of four clusters was observed according to the cut-off point based on Mojena's criterion, whose distance was 3.51, according to figure 1a. Cluster I was composed by the largest number of accessions, twelve, representing 66.67% of the total of 18 accessions studied.

Cluster II was formed by accessions 378 and 402 and cluster III by accessions 382, 387 and 415. The cluster IV was formed only by access 377. CCC was 0.76, which is significant (p<0.01) by Mantel's test and indicates that the clustering reflected the original distances (Nunes et al., 2011).

In situation 2, the UPGMA hierarchical method with multiscale bootstrap resampling allowed the formation of five clusters, according to figure 1b. For a cluster with p-value AU> 0.95, the hypothesis that "clustering does not exist" is rejected at 0.05 significance level. These clusters were the same formed in the previous situation, except for accessions 378 and 402 in which the null hypothesis was not rejected. In other words, considering 5% of significance level, the clustering did not exist, differently from what occurred in the analysis without resampling, in which the same accessions formed cluster II.

Thus, it is possible to affirm that the use of the multiscale bootstrap resampling technique in UPGMA clustering allowed a larger divergence among the accessions. So, the number of formed clusters was higher, compared to the performance of the usual method. The multiscale bootstrap resampling procedure confirms stability in the clusters when the number of samples bootstraps increases, thereby decreasing the sampling error. Suzuki & Shimodaira (2006) recommend to use nboot = 10000 for smaller errors.

Based on Tocher's method, eighteen cupuaçu accessions were separated into seven clusters, according to table 3. Cluster I was composed by eight accessions, all of them
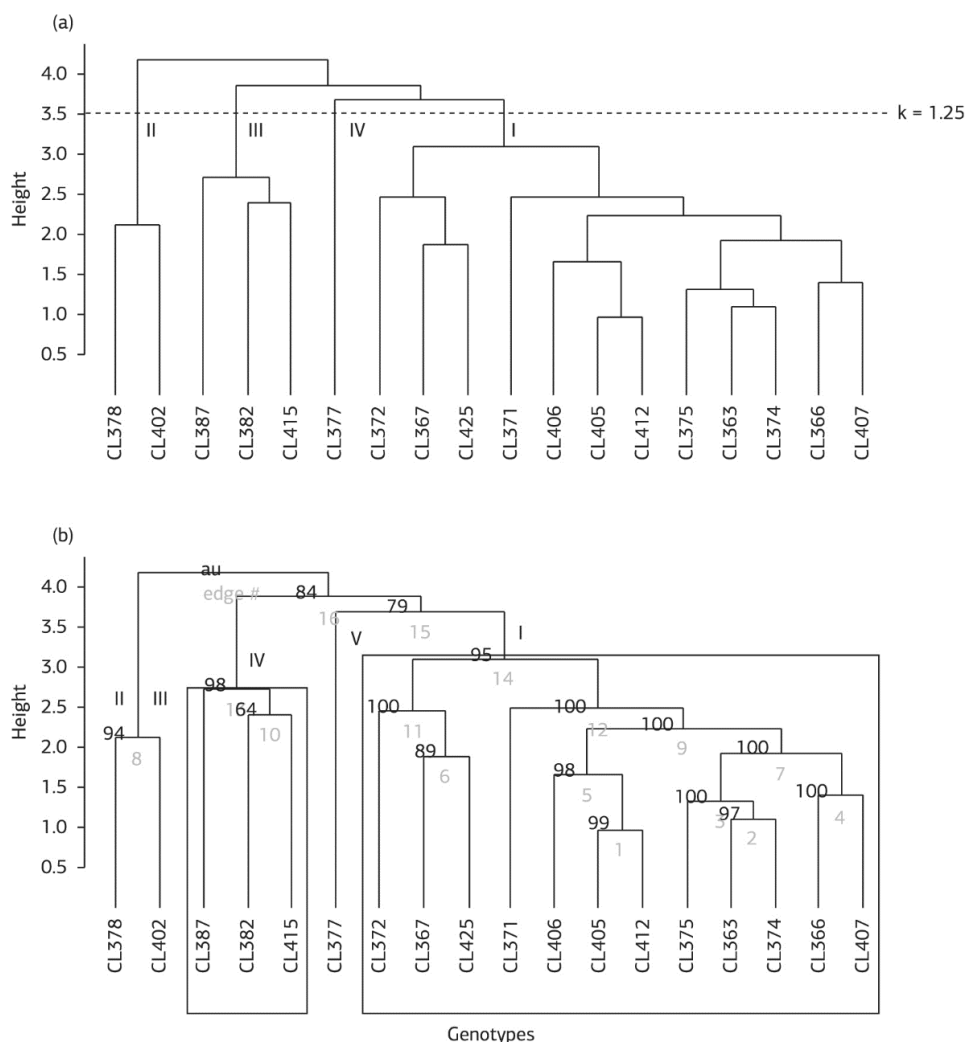
Figure 1. Dendrograms for eighteen cupuaçu clones using UPGMA method based on the Euclidean dissimilarity measure. Clusters formed by Mojena's criterion (a) and by p-value (AU) of multiscale bootstrap resampling (b); clusters with p-values (AU> 0.95) are identified by rectangles. Tomé-Açu city (PA, Brazil).

Table 3. Euclidean distance clustering (Tocher) for 18 experimental clones

Cluster	Genotypes
1	363, 366, 374, 375, 405, 406, 407, 412
2	367, 372, 425
3	378, 402
4	382, 415
5	387
6	377
7	371

belong to cluster I obtained by the UPGMA method in situations 1 and 2. Cluster II was formed by accessions 367, 372 and 425. Cluster III, formed by accessions 378 and 402, was the equal to cluster II, which was obtained by the UPGMA method in situation 1. Cluster IV was formed

by accessions 382 and 415, and cluster V was composed by access 387 alone. The combination of these two last clusters forms exactly the clusters III and IV, obtained by the UPGMA method, in situations 1 and 2, respectively.

It is worth highlighting that the combination of clusters I and II and cluster VII (access 371), obtained by Tocher's method form exactly the cluster I obtained by the UPGMA method in situations 1 and 2. Of the seven clusters formed by Tocher's method, three consisted of only one access, being that the access 377 formed isolated cluster in the three evaluated methods. According to Vasconcelos et al. (2007), in Tocher's method, the most dissimilar genotypes tend to form isolated clusters with only one genotype.

Although the numbers of clusters formed by the UPGMA method in situations 1 and 2 and by Tocher's method are not the same, there was no significant difference among the clustering methods used. This similarity in the discrimination

of accessions was also observed by Zuin et al. (2009) in cassava, by Nunes et al. (2011) in melon and by Guedes et al. (2013) in coffee.

The number of clusters formed in the three methodologies demonstrates the existence of genetic variability among the evaluated accessions, considering all variables simultaneously. Maia et al. (2011a), while evaluating eight cupuaçu clones, obtained four clusters by Tocher's method, based on mean and square Euclidean distances. Araújo et al. (2002), while evaluating 27 cupuaçu clones, obtained five clusters by Tocher's method, based on the generalized Mahalanobis' distance. However, in the evaluation of 36 cupuaçu progenies, Maia et al. (2011b) obtained smaller clusters, two and three, by the Tocher's method based on Euclidean & Mahalanobis square distances, respectively. It is worth emphasizing that the clusters formed in Araújo et al. (2002) were based on phenotypic values of clones, whereas estimated breeding values were used in the studies of Maia et al. (2011a, b).

The cupuaçu is a typical allogamous species. This ability leads to larger genetic dissimilarity among accessions (Araújo et al., 2002). This dissimilarity provides larger probability of obtaining different materials, some with favorable agronomic characteristics. On the other hand, the cross between different materials can promote larger heterosis. In this context, in order to obtain superior genotypes for production of pulp, would be recommended the cross between the access 378 with any access of the clusters 4 and 5 of the Tocher's method or clusters III and IV of the UPGMA method in situations 1 and 2 respectively. To improve the seed yield, crosses would be the most favorable among the access 378 and the accessions 367, 372 and 425 belonging to the cluster 2 of the Tocher's method or cluster I of the UPGMA method in both situations studied.

4. CONCLUSION

In cupuaçu accessions, although originated from the same locality, significant genetic variability is detected among them, with estimates of the individual broad-sense heritability of moderate magnitude for production traits. The UPGMA hierarchical clustering method in the two studied situations and Tocher's method are concordants in the formation of the clusters. The UPGMA method formed four and five clusters based on Mojena's criterion and on the multiscale bootstrap resampling technique, respectively, and the Tocher's method formed seven clusters. The multiscale bootstrap resampling technique can be used in studies of clustering consistency of hierarchical methods and, consequently in the determination of the optimal number of clusters.

REFERENCES

Alves, R. M., & Resende, M. D. V. (2008). Avaliação genética de indivíduos e progênies de cupuaçuzeiro no estado do Pará e estimativas de parâmetros genéticos. Revista Brasileira de Fruticultura, 30, 696-701. http://dx.doi.org/10.1590/S0100-29452008000300023.

Alves, R. M., Resende, M. D. V., Bandeira, B. S., Pinheiro, T. M., & Farias, D. C. R. (2010). Avaliação e seleção de progênies de cupuaçuzeiro (Theobroma grandiflorum), em Belém, Pará. Revista Brasileira de Fruticultura, 32, 204-212. http://dx.doi.org/10.1590/S0100-29452010005000010.

Alves, R. M., Resende, M. D. V., Bandeira, B. S., Pinheiro, T. M., & Farias, D. C. R. (2009). Evolução da vassoura-de-bruxa e avaliação da resistência em progênies de cupuaçuzeiro. Revista Brasileira de Fruticultura, 31, 1022-1032. http://dx.doi.org/10.1590/S0100-29452009000400015.

Alves, R. M., Silva, C. R. S., Silva, M. S. C., Silva, D. C. S., & Sebbenn, A. M. (2013). Diversidade genética em coleções amazônicas de germoplasma de cupuaçuzeiro [Theobroma grandiflorum (Willd. ex Spreng.) Schum.]. Revista Brasileira de Fruticultura, 35, 818-828. http://dx.doi.org/10.1590/S0100-29452013000300019.

Araújo, D. G., Carvalho, S. P., & Alves, R. M. (2002). Divergência genética entre clones de cupuaçuzeiro (Theobroma grandiflorum Willd. ex Spreng. Schum.). Ciência e Agrotecnologia, 26, 13-21.

Cargnelutti, A., Fo., Storck, L., & Ribeiro, N. D. (2009). Agrupamento de cultivares de feijão em presença e em ausência de multicolinearidade. Ciência Rural, 39, 2409-2418. http://dx.doi.org/10.1590/S0103-84782009000900005.

Cruz, C. D. (2013). GENES - a software package for analysis in experimental statistics and quantitative genetics. Acta Scientiarum, 35, 271-276.

Cruz, C. D., Regazzi, A. J., & Carneiro, P. C. S. (2012). Modelos biométricos aplicados ao melhoramento genético (4th ed.). Viçosa: UFV. 514 p.

Cruz, C. D., & Carneiro, P. C. S. (2006). Modelos biométricos aplicados ao melhoramento genético (2nd ed.). Viçosa: UFV. 585 p.

Guedes, J. M., Vilela, D. J. M., Rezende, J. C., Silva, F. L., Botelho, C. E., & Carvalho, S. P. (2013). Divergência genética entre cafeeiros do germoplasma Maragogipe. Bragantia, 72, 127-132. http://dx.doi.org/10.1590/S0006-87052013000200003.

Laviola, B. G., Rosado, T. B., Bhering, L. L., Kobayashi, A. K., & Resende, M. D. V. (2011). Genetic parameters and variability in physic nut accessions during early developmental stages. Pesquisa Agropecuaria Brasileira, 45, 1117-1123.

Maia, M. C. C., Resende, M. D. V., Oliveira, L. C., Álvares, V. S., Maciel, V. T., & Lima, A. C. (2011a). Seleção de clones experimentais de cupuaçu para características agroindustriais via modelos mistos. Revista Agro@mbiente On-line, 5, 35-43.

Maia, M. C. C., Resende, M. D. V., Oliveira, L. C., Alves, R. M., Silva, J. L., Fo., Rocha, M. M., Cavalcante, J. J. V., & Roncatto, G. (2011b). Análise genética de famílias de meios-irmãos de cupuaçuzeiro. Pesquisa Florestal Brasileira, 31, 123-130. http://dx.doi.org/10.4336/2011. pfb.31.66.123.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. Psychometrika, 50, 159-179. http://dx.doi.org/10.1007/BF02294245.

Nunes, G. H. S., Costa, J. H., Fo., Silva, D. J. H., Carneiro, P. C. S., & Dantas, M. S. M. (2011). Divergência genética entre linhagens de melão pele de Sapo. Revista Ciência Agronômica, 42, 765-773. http://dx.doi.org/10.1590/S1806-66902011000300024.

Prance, G. T., & Silva, M. F. (1975). Árvores de Manaus. Manaus: INPA. 312 p.

Resende, M. D. V. (2007a). Selegen–Reml/Blup: Sistema Estatístico e Seleção Genética Computadorizada via Modelos Lineares Mistos. Colombo: Embrapa Florestas. 360 p.

Resende, M. D. V. (2007b). Matemática e estatística na análise de experimentos e no melhoramento genético. Colombo: Embrapa Florestas. 561 p.

Rizza, M. D., Real, D., Reyno, R., Porro, V., Burgueño, J., Errico, E., & Quesenberry, K. H. (2007). Genetic diversity and DNA content of three South American and three Eurasiatic Trifolium species. Genetics and Molecular Biology, 30, 1118-1124. http://dx.doi.org/10.1590/S1415-47572007000600015.

Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. Systematic Biology, 51, 492-508. http://dx.doi.org/10.1080/10635150290069913. PMid:12079646

Suzuki, R., & Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics, 22, 1540-1542. http://dx.doi.org/10.1093/bioinformatics/btl117. PMid:16595560

Suzuki, R., & Shimodaira, H. (2014). Pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling. R package version 1.3-0. Vienna: The R Foundation. Recuperado em 20 de agosto de 2014, de http:// www.R-project.org

Taniguchi, Y., Matsuda, H., Yamada, T., Sugiyama, T., Homma, K., Kaneko, Y., Yamagishi, S., & Iwasaki, H. (2013). Genome-wide SNP and STR discovery in the Japanese crested ibis and genetic diversity among founders of the Japanese population. PLoS ONE, 8, e72781. http://dx.doi.org/10.1371/journal.pone.0072781. PMid:23991150

Vasconcelos, E. S., Cruz, C. D., Bhering, L. L., & Resende, M. F. R., Jr. (2007). Método alternativo para análise de agrupamento. Pesquisa Agropecuaria Brasileira, 42, 1421-1428. http://dx.doi.org/10.1590/S0100-204X2007001000008.

Venturieri, G. A. (2011). Flowering levels, harvest season and yields of cupuassu (Theobroma grandiflorum). Acta Amazonica, 41, 143-152. http://dx.doi.org/10.1590/S0044-59672011000100017.

Zuin, G. C., Vidigal, P. S., Fo., Kvitschal, M. V., Gonçalves-Vidigal, M. C., & Coimbra, G. K. (2009). Divergência genética entre acessos de mandioca-de-mesa coletados no município de Cianorte, região Noroeste do Estado do Paraná. Semina: Ciências Agrárias, 30, 21-30. http://dx.doi.org/10.5433/1679-0359.2009v30n1p21.