

# Comparison of methods for selection of castor beans lineages

Maria Márcia Pereira Sartori\*, Jackson da Silva, Mauricio Dutra Zanotto

São Paulo State University, Botucatu, Brazil  
\*Corresponding author, email: mmpsartori@fca.unesp.br

## Abstract

The choice of the most appropriate method is determined by the precision desired by the researcher, by the ease of the analysis, as well as by the way of obtaining the data. In order to select lineages of low size and high productivity this study aimed to evaluate different methods of cluster analysis in the representation of genetic divergence, compared to univariate methods. The analyzed variables were grain yield, plant size and oil yield of 24 lineages of castor beans cultivated in the years 2014 and 2015. The Single and Average methods presented similar results in the formation of groups and different from the Complete. Evaluating the purpose of this research the Complete method and principal components analysis, together with the discriminant analysis, were considered the most appropriate methods to evaluate the genetic divergence of the castor bean crop. Lineages 18, 19 and 20 showed average grain yields above 1555 kg.ha<sup>-1</sup>, high oil content (above 46.9%), and low size plants (below 116 cm).

**Keywords:** *Ricinus communis* L., multivariate statistics, genetic divergence, genotypes, plant breeding

## Introduction

The castor bean (*Ricinus communis* L.) is an oleaginous of high socioeconomic value, in Brazil the commercialization of production, in 2015 generated approximately R\$ 67 million (IBGE, 2015). Its products and by-products are used in industry and agriculture, as well as presenting perspectives of use as an energy source in the form of biodiesel (Tabile et al., 2009). According to CONAB (2017) the productivity in the southeastern and central-south regions of Brazil for the 2016/2017 harvest was 1000 kg.ha<sup>-1</sup>.

Almeida et al. (2015) pointed out that the greatest difficulty in the rational exploitation of castor bean in Northeast Brazil is the low

availability of seeds of cultivars adapted, productive, with high oil content and tolerant to pests and diseases.

Considering the need to select improved cultivars, breeding programs have directed the evaluation of the behavior of a group of cultivars related to environmental variations, such as different sites, seasons or years of planting, as well as several technological levels (Cruz et al., 2012). In this context, multivariate predictive methods, such as principal component analysis (PCA) and agglomerative methods, have been of paramount importance, for example, in assessing genetic divergence, making it possible to size diversity and group the material, managementing

Received: 20 July 2017  
Accepted: 23 June 2018

better the use of the groups formed (Cruz et al., 2012; Bertini et al., 2010). Bertan et al. (2006) concluded that the methods of clustering Average linkage method and Tocher method showed similar results in the representation of the groups of most divergent genotypes for wheat.

For Cruz et al. (2012) the analysis of principal components provides a considerable simplification in statistical calculations and interpretation of the results in relation to the other alternative methods, especially when the number of progenitors or lineages evaluated is relatively large. According to the same authors, an advantage of this method is the evaluation of the importance of each character studied on the total variation available among the evaluated genotypes, thus enabling the exclusion of characters that contribute little to the discrimination of the evaluated material.

Multivariate methods were also used by Goodarzi et al. (2011) who carried out the cluster analysis of 12 access of castor beans using the hierarchical algorithm Ward based on Euclidean distance squared. The discriminant analysis was applied to confirm the accuracy of the cluster produced by cluster analysis and to identify patterns of morphological variation used the principal component analysis (PCA) thus assessing similarity.

Thus, the need for simultaneous analysis of multiple measures highlights the importance of the application of multivariate techniques in the genetic improvement of plants and justifies the development of this project that aims to identify the best methodology for the selection of lineages of low size, high oil and high productivity.

### Material and Methods

The present work was developed in the years 2014 and 2015 in the Experimental Farm of Lageado, belonging to the Faculty of Agronomic Sciences (FCA) of Paulista State University (UNESP), Campus de Botucatu-SP. According to Alvares et al. (2014), the climate of the city of Botucatu-SP is of type Cfa (Climate Temperate Mesothermal), the average annual temperature is between 18 to 20 °C, with annual average rainfall is between 1000 to 1300 mm and altitude between 600 and 800 meters of altitude. The soil

was classified as Typic Dystrophic Red-Yellow Latosol with a average texture (Silva et al., 2013).

The experimental design was a randomized block with 24 lineages and four replicates (blocks), being the experimental plot of 8 m<sup>2</sup> of useful area, constituting three lines of 10 meters in length, with spacing between plants and between rows of 1.0 m in two time, 2014 and 2015. The evaluated characteristics were oil content in the grains (TO), in %, measured by nuclear magnetic resonance, being evaluated a sample composed of seed mix of all the plants of the useful area of each experimental plot; grain yield (PG), in kg.ha<sup>-1</sup>, determined by the total grain mass in 8 m<sup>2</sup> corrected to 13% moisture, being evaluated all plants of the useful area of each experimental plot and height of the plant (AP), in cm, being measured from the soil surface to the highest point of the plant, for 8 plant samples of the useful area of each experimental plot.

The data were submitted to analysis of joint variance, complemented with Scott Knott's test, where it was considered significant when  $p < 0.05$ . After this evaluation was used hierarchical grouping and optimization methods to find the most appropriate method for the data set.

Five different agglomerative methods were used:

The Euclidean distance method is a standard mathematical measure of distance (square root of the sum of squared differences).

The nearest neighbor method "Single Linkage Method", which defines the distance between two clusters by the minimum distance between each item with respect to each cluster formed (Lattin et al., 2011).

The most distant neighbor method "Complete linkage method", which defines the distance between two clusters by the maximum distance between each item with respect to each formed cluster (Lattin, et al., 2011).

The unweighted average pairing method "Average linkage method" that defines the distance between two clusters by the average distance between each item with respect to each formed cluster (Lattin et al., 2011).

The Ward method defines the distance between two clusters by a function of binding

criteria: the sum of the square deviations of the points for the centroids, minimizing the sum of squares within the cluster (Lattin, et al., 2011).

It was also used the Tocher optimization method (Cruz et al., 2012), which performs the partitioning of the set of accesses in non-empty and mutually exclusive subgroups, by means of the maximization or minimization of some preestablished measure. In this method, adopts the criterion that the mean dissimilarity measures within each group should be less than the mean distances between any groups, thus evaluating the importance of the selection process in the groupings obtained.

To determine the best methodology, the representations of the dendrograms were evaluated. The discriminant analysis was applied to confirm the accuracy of the grouping produced by the cluster analysis, being Calculated the distance between the centroids of the groups formed by the generalized distance of Mahalanobis (Cargnelutti Filho et al., 2008) was calculated.

Genetic divergence by major components was also analyzed. The software used for all the analyzes was Minitab 16 (Minitab version 14, Minitab Inc., State College, PA, USA), only for the optimization method SAEG software was used.

## Results and Discussion

It can be observed that the variable yield of grains presented significant difference for the source of variation lineage. The meaningfulness of the lineage x time interaction, for the variables height and oil content, reveals the need to evaluate the lineages for more than one year, indicating that the genetic constitutions are divergent for the characters evaluated, thus enabling studies of divergence genetics (Oliveira et al., 2013). It is verified also a high experimental precision of the work, since the coefficients of variation were of low magnitude, varying from 4.79% for the oil content to 17.45% for grain yield, these being considered of optimal and regular experimental precision, respectively (Table 1) (Cargnelutti Filho et al., 2010).

**Table 1.** Joint analysis of variance, average, maximum and minimum values of the variables: plant height (AP), oil content (TO) and grain yield (PG) of 24 lineages grown in the times 2014 and 2015.

	AP (cm)	TO (%)	PG (kg ha <sup>-1</sup> )
QM time (E)	10912.0 <sup>ns</sup>	8.3 <sup>ns</sup>	4380323.1 <sup>**</sup>
QM lineages (L)	3441.3 <sup>ns</sup>	19.9 <sup>ns</sup>	907823.6 <sup>**</sup>
QM E x L	3351.3 <sup>**</sup>	24.2 <sup>**</sup>	107000.4 <sup>ns</sup>
QM resíduo	267.1	4.6	80682.8
CV%	10.4	4.79	17.45

<sup>\*\*</sup>, <sup>ns</sup>: Significant at 1% probability and not significant at 5% probability by the F test, respectively.

The highest variation was found in the first year of cultivation for plant height, presenting average values of 58.2 to 225.6 cm, being that the lineages 18, 19 and 20 presented smaller height of plant, already in the second year there was no difference between the lineages for this evaluation (Table 2).

For the oil content, the variations between the lineages were similar for the two times of cultivation, varying between 40 and 50%, it can be observed that the lineages 8, 11, 12, 18, 19 and 20 presented values higher than 46 % For the two time of cultivation, (Table 2). It is noteworthy that the performances of these lineages were similar regardless of the time of cultivation.

The yields did not differentiate between

the lineages, for each time reported, however, for the average of two time, lineages 1, 3, 10, 11, 19 and 22 had the highest yields, being higher than 1880 kg ha<sup>-1</sup>, Well above the region average of 1000 kg.ha<sup>-1</sup>. It is also observed that the second time of cultivation provided the best PG, attributing to this fact the best climatic conditions (Table 2).

One of the requirements used by farmers to choose cultivars is high productivity (França et al., 2013; Souza et al., 2010; Zuchi et al., 2010), so is elementary an selection based on this characteristic, ally to this, the present work also took into account plants of low size and high oil content, with this it was possible to select only the lineage 19, being this the only one that presented these three characteristics, by the commonly

**Table 2.** Average of plant height (AP), oil content (TO) and grain yield (PG) of 24 lineages grown in two times 2014 and 2015.

Lineages	AP (cm)		TO (%)		PG (kg ha <sup>-1</sup> )		Average
	Time		Time		Time		
	1	2	1	2	1	2	
1	200.7 bA	163.0 aB	48.3 aA	42.7 cB	1670.5	2107.63	1889.0 a
2	116.0 dB	165.3 aA	45.6 bA	45.2 bA	1293.8	1562.25	1428.0 b
3	189.5 bA	165.6 aB	43.3 bA	45.2 bA	2026.8	2292.67	2159.7 a
4	111.8 dB	163.6 aA	46.3 aA	46.0 bA	1666.8	1690.88	1678.8 b
5	155.8 cA	170.5 aA	47.6 aA	46.6 bA	1435.8	1821.22	1628.4 b
6	154.7 cA	167.0 aA	45.0 bA	45.6 bA	1703.3	1309.62	1506.4 b
7	155.7 cA	173.6 aA	45.0 bA	47.5 bA	1432.0	1842.82	1637.4 b
8	172.5 cA	171.6 aA	46.1 aA	46.9 bA	1350.3	1614.14	1482.1 b
9	165.9 cA	161.6 aA	49.1 aA	43.9 cB	1475.8	1944.98	1710.3 b
10	152.0 cA	151.3 aA	44.7 b A	41.1 cB	2039.3	2345.08	2192.1 a
11	156.0 cA	171.2 aA	50.2 aA	46.9 bB	1880.0	2611.64	2245.8 a
12	156.2 cA	169.0 aA	48.3 aA	46.2 bA	1440.3	2052.23	1746.2 b
13	176.4 bA	163.1 aA	43.8 bA	44.4 bA	1351.3	1825.98	1588.6 b
14	165.7 cA	168.1 aA	46.8 aA	45.9 bA	1544.0	1804.31	1674.1 b
15	181.3 bA	168.2 aA	41.9 bB	50.1 aA	1394.8	1478.41	1436.5 b
16	141.7 cB	166.6 aA	45.9 aA	45.5 bA	1175.0	1240.16	1207.5 c
17	152.2 cA	164.6 aA	48.1 aA	44.9 bB	948.8	1057.15	1002.9 c
18	68.9 eB	162.3 aA	46.4 aA	49.3 aA	1408.5	1702.40	1555.4 b
19	58.2 eB	163.4 aA	48.2 aA	50.1 aA	1997.5	2455.86	2226.6 a
20	67.9 eB	158.6 aA	46.1 aA	47.5 bA	1530.0	1823.74	1676.8 b
21	159.6 cA	172.0 aA	43.7 bB	47.1 bA	1532.8	1864.11	1698.4 b
22	146.0 cA	146.2 aA	44.2 bA	41.8 cA	2182.5	2618.74	2400.6 a
23	129.1 cA	147.7 aA	48.3 aA	40.1 cB	1323.8	1880.98	1602.3 b
24	225.6 aA	147.3 aB	43.1 bA	45.4 bA	1441.3	1547.38	1494.3 b
Average	148.4	163.4	46.1	45.7	1551.8 B	1853.9 A	1702.8

Averages followed by equal letters, uppercase in the row and lowercase in the column, do not differ by the Scott Knott test, 5% probability.

used method (univariate).

However, other methodologies may be used for selection, such as, the methods of grouping, where they are extremely important when selection is dependent on more than one choice factor, such as, for example, productivity, plant size and oil content.

The "Euclidean Distance, Single Linkage" grouping method allowed dividing of the 24 lineages into 4 distinct groups (Figure 1), with the following characteristics: Group I (red) with 16 lineages (1, 2, 4, 5, 7, 8, 9, 12, 13, 14, 15, 18, 20, 21,23 and 24) and intermediate grain yield between 1428 and 1889 Kg.ha<sup>-1</sup>, oil content between 44 and 47 and plant height between 113 and 186 cm; Group II (blue) formed by a single lineage (6) and with average yield of 1500 kg.ha<sup>-1</sup>, oil content of 45 and height of 160 cm; Group III (green) with 5 lineages (3, 10, 11, 19 and 22) and higher average productivity above 2150 kg.ha<sup>-1</sup>, oil content between 42 and 49 and plant height between 110 and 177 cm; Group IV (orange) with 2 lineages (16 and 17)

and productivity between 1000 and 1200 kg.ha<sup>-1</sup>, oil content between 45 and 46 and plant height between 154 and 158 cm.

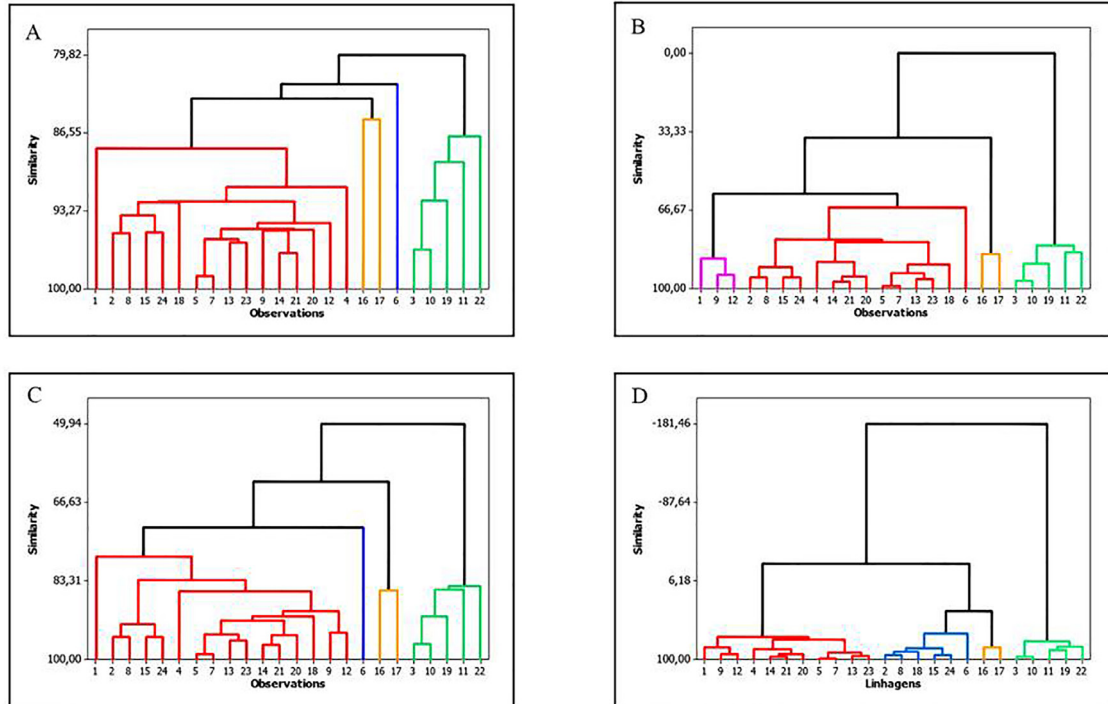
Milani et al. (2009), evaluating the genetic divergence of castor bean genotypes also found similar results with four classification groups.

The groups formed by the "Euclidean Distance, Average Linkage" method were identical to those formed by the "Euclidean Distance, Single Linkage" method (Figure 1), although they presented different similarities, however, they were distinct from the grouping method "Euclidean Distance, Complete Linkage" (Figure 1).

For the "Euclidean Distance, Complete Linkage" method the division of the 24 lineages was in 4 distinct groups (Figure 1), group I and II formed by the other methods were divided into two other groups I and II, as described below: Group I (red) with 14 lineages (2, 4, 5, 6, 7, 8, 13, 14, 15, 18, 20, 21,23 and 24) with productivity between 1428 and 1698 kg.ha<sup>-1</sup>, oil

content between 44.2 and 47.9 and plant height between 113 and 186 cm; Group II (pink) with 3 lineages (1, 9 and 12) and productivity between 1710 and 1889 kg.ha<sup>-1</sup>, oil content from 45.5 to 47.3 and height from 162 to 181.9 cm; Group III (green) with 5 lineages (3, 10, 11, 19 and 22) and higher average productivity above 2150 kg.ha<sup>-1</sup>,

oil content between 42 and 49 and plant height between 110 and 177 cm; Group IV (orange) with 2 lineages (16 and 17) and average productivity between 1000 and 1200 kg.ha<sup>-1</sup>, oil content between 45 and 46 and plant height between 154 and 158 cm.



**Figure 1.** Grouping of 24 lineages of castor bean according to the variables of productivity, oil content and plant height, by Single (A), Complete (B), Average (C) and Ward (D).

It can be observed that the grouping method “Euclidean Distance, Complete Linkage” presents advantages over previous ones, since it determined groups with higher and lower productivity maintaining the variability of the variables in the groups formed.

In the “Euclidean Distance, Ward Linkage” method the division of the 24 lineages was in 4 distinct groups (Figure 1), group I and II formed by the other methods were divided into two other groups I and II, as described below: Group I (red) with 11 lineages (1, 4, 5, 7, 9, 12, 13, 14, 20, 21, and 23) with productivity between 1589 and 1889 kg.ha<sup>-1</sup>, oil content between 44.2 and 47.3 and height of the plant between 113 and 182 cm; Group II (blue) with 6 lineages (2, 6, 8, 15, 18 and 24) with productivity between 1428 and 1555 kg.ha<sup>-1</sup>, oil content between 44.3 and 47.9 and plant height between 115 and 186 cm; Group III (green) with 5 lineages (3, 10, 11, 19 and

22) and higher productivity above 2150 kg.ha<sup>-1</sup>, oil content between 42 and 49 and plant height between 110 and 177 cm; Group IV (orange) with 2 lineages (16 and 17) and average productivity between 1000 and 1200 kg.ha<sup>-1</sup>, oil content between 45 and 46 and plant height between 154 and 158 cm.

For the Tocher optimization method we can observe the formation of 11 groups (Table 3), much higher than the work done by Bezerra Neto et al. (2010), where wanting to quantify the genetic diversity of eleven genotypes of castor bean and Cavalcante et al. (2008) evaluating the genetic divergence between accessions of castor bean, using the same test, observed the formation of only three and four groups, respectively, which may be related to the high index of similarity among the varieties, besides the low number of genotypes used.

**Table 3.** Formation of the clusters by the Tocher test..

Group	Number of lineages	Selected lineages
1	12	2; 4; 5; 6; 7; 8; 9; 12; 13; 14; 16; 21
2	2	18; 20
3	2	10; 22
4	1	1
5	1	3
6	1	11
7	1	15
8	1	17
9	1	19
10	1	23
11	1	24

By the discriminant analysis it can be verified that all the analyzed groups presented / displayed 100% of the groups formed of correct form. The formation of 3 distinct response groups ( $p < 0.05$ ) was verified, for the "Euclidean Distance, Single Linkage" and "Average Linkage" grouping methods, similar results were found by Bahia et al. (2008) who also evaluated the genetic divergence between castor bean varieties. In these cases the lineage 6 was grouped with the lines described in group I whose lineages are 1, 2, 4, 5, 7, 8, 9, 12, 13, 14, 15, 18, 20, 21, 23 and 24, this occurred because the method did not allow group formation with a single lineage. For the group found by the grouping method "Euclidean Distance, Complete Linkage" and "Ward Linkage" one can verify the distinction of 4 groups, as found by the method.

It can be observed that only the groups with higher and lower productivity, were formed

by the same lineages by any hierarchical agglomerative method used, presenting the longest distance from Mahalanobis (Table 4), which is interesting for the indication of possible crosses between genetically divergent progenitors, leading to the production of hybrid hybrids, however, the lineages with yield of grains between 1400 and 1900, were allocated in different groups depending on the method used, which is not interesting, because it can lead to mistaken decision making (Machado et al., 2013; Rodrigues et al., 2010). The "Euclidean Distance", "Ward Linkage" method, presented a smaller distance between groups 2 and 4, proving that these groups present greater similarity and greater distance between 2 and 3, showing the lowest similarity, which may represent a better result when the selection is in function to a single variable, for example grain yield (Cargnelutti Filho & Guadagnin, 2011).

**Table 4.** Squared Distance Between Groups from Euclidean Distance, Single Linkage e Average Linkage (A), Complete (B), Ward (C) e PCA (D).

A			B			
Groups	II	III	Groups	II	III	IV
I	19.0988	14.2823	I	10.111	52.269	29.242
II		63.9506	II		29.507	58.230
			III			156.594

C				D			
Groups	II	III	IV	Groups	II	III	IV
I	9.989	50.392	58.179	I	41.550	75.301	113.969
II		95.188	27.201	II		9.242	40.072
III			214.628	III			20.771

However, if we consider the selection in function also of the plant with smaller height and higher oil content, it can be emphasized that none of the previously evaluated methods allows this selection.

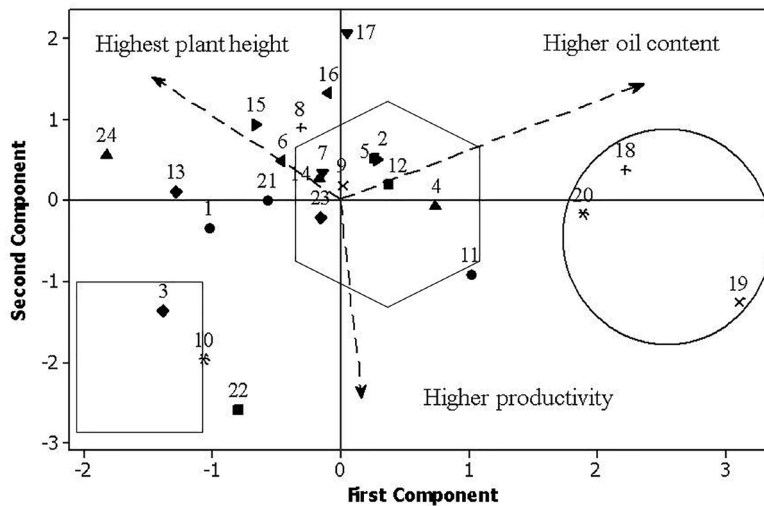
The Single Linkage method should be avoided in ecological studies, since in by bringing together object to the "closest" element of the already formed group, the intermediate objects between the groups are rapidly agglomerated

to these, then taking place a chaining of objects which makes it difficult to separate the groups (Diniz et al., 2012).

Because the Average Linkage and Complete Linkage methods resemble to Single Linkage, it is believed that these methods should also be avoided, especially when using more than one type of variable, being, therefore, more suitable when used for the same type of variable obtained only at different times.

However, the principal components analysis (Figure 2) allows selection of lineages with average grain yield over 1555 kg.ha<sup>-1</sup>, high oil content (above 46.9%) and low plants (below

116 cm) (Figura 2), being these the lineages 18, 19 and 20, as well as to verify the lineages with high average productivity of grain (above 2200 kg.ha<sup>-1</sup>), more that present low oil content and higher height (lineages 3, 10 and 22) and lineages with intermediate results, however, with greater variability (lineages 2, 4, 5, 9, 11 and 12), with grain yield between 1428 and 2245 kg.ha<sup>-1</sup>, oil content between 45.5 and 48.6% and height between 137 and 163 cm. Thus, it can be confirmed that this analysis is adequate both when using more than one type of variable, and also when using the same type of variable obtained at different times.



**Figure 2.** Scatter plot of principal component analysis showing the position of the 24 castor bean lineages. In the circumference are grouped the lineages with average productivity of grain, high oil content and plants of low size.

## Conclusions

The univariate analysis allowed the selection of only the lineage 19, being that one that met the three characteristics (low size, high oil content and productivity).

The clustering methods presented different results, being able to be the most adequate when using a single characteristic.

The principal component analysis method was the only one that allowed the efficient selection of castor bean lineages, taking into account the high yield of grains, high oil content and plants of low size, being these lineages 18, 19 and 20.

## Acknowledgments

To the National Council of Scientific and Technological Development Brazil (CNPq), for

the financial aid (process n° 310794/2015-3) and for the scholarship (process n° 134541/2016-2).

## References

- Almeida, B.M.S., Santos, R.L.C., Fluminhan Júnior, A. 2015. Melhoramento genético de mamona (*Ricinus communis* L.) visando à produção de variedades adaptadas ao cultivo mecanizado. *Fórum Ambiental da Alta Paulista* 11: 267-282.
- Alvares, C.A., Stape, J.L., Sentelhas, P.C., Gonçalves, J.L.M., Sparovek, G. 2014. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift* 22: 711-728.
- Bahia, H.F., Silva, S.A., Fernandez, L.G., Ledo, C.A.S., Moreira, R.F.C. 2008. Divergência genética entre cinco cultivares de mamoneira. *Pesquisa Agropecuária Brasileira* 43: 357-362.
- Bertan, I., Carvalho, F.I.F., Oliveira, A.C., Vieira, E.A., Hartwig I., Silva, J.A.G., Shimidt, D.A.M., Valério, I.P., Busato, C.C., Ribeiro, G. 2006.

- Comparação de métodos de agrupamento na representação da distância morfológica entre genótipos de trigo. *Revista Brasileira de Agrociência* 12: 279-286.
- Bertini, C.H.M., Pinheiro, E.A.R., Nóbrega, G.N., Duarte, J.M.L. 2010. Desempenho agrônômico e divergência genética de genótipos de coentro. *Revista Ciência Agronômica* 41: 409-416.
- Bezerra Neto, F.V., Leal, N.R., Gonçalves, L.S.A., Rêgo Filho, L.M., Amaral Júnior, A. T. 2010. Descritores quantitativos na estimativa da divergência genética entre genótipos de mamoneira utilizando análises multivariadas. *Revista Ciência Agronômica* 41: 294-299.
- Cargnelutti Filho, A., Guadagnin, J. P. 2011. Consistência do padrão de agrupamento de cultivares de milho. *Ciência Rural* 41: 1503-1508.
- Cargnelutti Filho, A., Lopes, S.J., Brum, B., Silveira, T.R., Toebe, M., Storck, L. 2010. Tamanho de amostra de caracteres em híbridos de mamoneira. *Ciência Rural* 40: 1-8.
- Cargnelutti Filho, A., Ribeiro, N.D., Reis, R.C.P., Souza, J.R., Jost, E. 2008. Comparação de métodos de agrupamento para o estudo da divergência genética em cultivares de feijão. *Ciência Rural* 38: 2138-2145.
- Cavalcante, M., Paixão, S.L., Ferreira, P.V., Madalena, J.A.S., Costa, J.G. 2008. Divergência genética entre acessos de mamona em dez municípios de Alagoas. *Caatinga* 21: 111-115.
- CONAB. Companhia Nacional de Abastecimento. 2017. <http://www.conab.gov.br/conteudos.php?a=1252&t=2> <Access in 20 May. 2017>
- Cruz, C.D., Regazzi, A.J., Carneiro, P.C.S. 2012. *Modelos Biométricos Aplicados ao Melhoramento Genético*. Editora UFV, Viçosa, Brazil. 514 p.
- Diniz, R.B.N., Soares, V.G., Cabral, L.A.F. 2012. Uso de técnicas de mineração de dados na identificação de áreas hidrologicamente homogêneas no Estado da Paraíba. *Revista Brasileira de Recursos Hídricos* 17: 65-75.
- França, R.G., Afférrri, F.S., Jácome, A.N., Araújo, O.R.B., Naoe, L.K., Pelúzio, J.M. 2013. Comportamento de cultivares de mamona (*Ricinus communis*) sob diferentes densidades populacionais, no estado do Tocantins. *Revista de Ciências Agrárias* 36: 317-323.
- Goodarzi, F., Darvishzadeh, R., Hassani, A., Hassanzaeh, A. 2011. Study on genetic variation in Iranian castor bean (*Ricinus communis* L.) accessions using multivariate statistical techniques. *Journal of Medicinal Plants Research* 5: 5254-5261.
- IBGE. Instituto Brasileiro de Geografia e Estatística. 2015. <http://www.ibge.gov.br/estadosat/> <Access in 20 May. 2017>
- Lattin, J., Carroll, J.D., Green, P.E. 2011. *Análise de dados multivariados*. Cengage Learning, São Paulo, Brazil. 475 p.
- Machado, E.L., Silva, S.A., Santos, A.S., Bastos, L.A., Pestana, C.N., Santos, K.S., Ferreira, C.F., Diamantino, M.S.A.S. 2013. Dissimilaridade genética entre cultivares de mamoneira por meio de marcadores RAPD. *Pesquisa Agropecuária Brasileira* 48: 342-345.
- Milani, M., Dantas, F.V., Martins, W.F.S. 2009. Divergência genética em mamoneira por caracteres morfológicos e moleculares. *Revista Brasileira de Oleaginosas e Fibras* 13: 61-71.
- Oliveira, R.S., Silva, S.A., Brasileiro, B.P., Medeiros, E.P., Anjos, E.V.A. 2013. Genetic divergence on castor bean using the ward-mlm strategy. *Revista Ciência Agronômica* 44: 564-570.
- Rodrigues, H.C.A., Carvalho, S.P., Carvalho, A.A., Santos, C.E.M., Carvalho Filho, J.L.S. 2010. Correlações genotípicas, fenotípicas e ambientais entre caracteres de mamoneira. *Ciência e Agrotecnologia* 34: 1390-1395.
- Silva, C.C., Coelho, R.M., Oliveira, S.R.M., Adami, S.F. 2013. Mapeamento pedológico digital da folha Botucatu (SF-22-Z-B-VI-3): treinamento de dados em mapa tradicional e validação de campo. *Revista Brasileira de Ciência do Solo* 37: 846-857.
- Souza, N.C., Mota, S.B., Bezerra, F.M.L., Aquino, B.F., Santos, A.B. 2010. Produtividade da mamona irrigada com esgoto doméstico tratado, *Revista Brasileira de Engenharia Agrícola e Ambiental* 14: 478-484.
- Tabile, R.A., Lopes, A., Dabdoub, M.J., Camara, F.T., Furlani, C.E.A., Silva, R.P. 2009. Biodiesel de mamona no diesel interior e metropolitano em trator agrícola. *Engenharia Agrícola* 29: 412-423.
- Zuchi, J., Bevilaqua, G.A.P., Zanuncio, J.C., Peske, S.T., Silva, S.D.A., Sediyaama, C.S. 2010. Características agrônômicas de cultivares de mamona em função do local de cultivo e da época de semeadura no Rio Grande do Sul. *Ciência Rural* 40: 501-506.