GMR

Multivariate approach in popcorn genotypes using the Ward-MLM strategy: morpho-agronomic analysis and incidence of *Fusarium* spp

R.N.F. Kurosawa, A.T. do Amaral Junior, F.H.L. Silva, A. dos Santos, M. Vivas, S.H. Kamphorst and G.F. Pena

Laboratório de Genética e Melhoramento Vegetal,
Universidade Estadual do Norte Fluminense Darcy Ribeiro,
Campos dos Goytacazes, RJ, Brasil

Corresponding author: A.T. do Amaral Junior
E-mail: amaraljr@pq.cnpq.br

Genet. Mol. Res. 16 (1): gmr16019528
Received November 7, 2016
Accepted January 13, 2017
Published February 8, 2017
DOI http://dx.doi.org/10.4238/gmr16019528

ABSTRACT. The multivariate analyses are useful tools to estimate the genetic variability between accessions. In the breeding programs, the Ward-Modified Location Model (MLM) multivariate method has been a powerful strategy to quantify variability using quantitative and qualitative variables simultaneously. The present study was proposed in view of the dearth of information about popcorn breeding programs under a multivariate approach using the Ward-MLM methodology. The objective of this study was thus to estimate the genetic diversity among 37 genotypes of popcorn aiming to identify divergent groups associated with morpho-agronomic traits and traits related to resistance to *Fusarium* spp. To this end, 7 qualitative and 17 quantitative variables were analyzed. The experiment was conducted in 2014, at Universidade Estadual do Norte Fluminense, located in Campos dos Goytacazes, RJ, Brazil. The Ward-MLM strategy allowed the identification of

four groups as follows: Group I with 10 genotypes, Group II with 11 genotypes, Group III with 9 genotypes, and Group IV with 7 genotypes. Group IV was distant in relation to the other groups, while groups I, II, and III were near. The crosses between genotypes from the other groups with those of group IV allow an exploitation of heterosis. The Ward-MLM strategy provided an appropriate grouping of genotypes; ear weight, ear diameter, and grain yield were the traits that most contributed to the analysis of genetic diversity.

Key words: *Zea mays* var. everta; Genetic diversity; Canonical variables; Heterosis; Ear infection

INTRODUCTION

Many trial procedures are multivariate, as they involve the evaluation of several traits, or response variables, in all experimental units. Multivariate data analysis methods permit the unified study of different variables under investigation, putting emphasis on associations, similarities, or differences between them and thus losing the least information possible. The first ideas about data analysis employing multiple responses emerged from the contributions of Pearson (1901), Fisher (1928), Hotelling (1933), Wilks (1932), and Bartlett (1939), who initiated the development of analytical procedures to address these evaluations.

In plant breeding, the use of multivariate approaches enables the evaluation of the genetic material on a set of traits that combines multiple pieces of information in a way that it is possible to select the most promising materials, considering the contribution and relative importance of the traits for the total existing variance (Cruz et al., 2014). With the recent advances in computer science resulting from the development of interactive data-processing systems, as well as the generation of fast algorithms, coupled with the increasing popularization of computational techniques, the large demand for processing required in the multivariate data analysis ceased to be a problem.

The use of multivariate analysis techniques provides some advantages in relation to the univariate approach, e.g.: i) the study of direct and indirect effects of traits on a basic variable, as occurs in path analysis; ii) the evaluation of interrelationships between two sets determined by an arbitrary number of traits, considering not only one single dependent variable (example: canonical variables); iii) the combination of multiple pieces of information contained in the experimental unit, such that selection can be performed based on a set of variables that includes several attributes of economic interest, as is the case of selection indices; and iv) the generation of knowledge of the genetic diversity between parents aiming to identify the hybrid combinations with greater heterotic effect and higher heterozygosis, and even to identify duplicates in germplasm banks (Cruz et al., 2014), among others.

In the prediction of genetic diversity, several multivariate methods can be applied. The choice of the most suitable model has been determined according to the precision desired by the researcher, the ease of analysis, and how the data are obtained (Cargnelutti Filho et al., 2008; Bezerra Neto et al., 2010). The cluster analysis is a method that aims at gathering sampling units into groups, through some classification criterion, allowing for homogeneity within group and heterogeneity between groups. In this way, the use of multivariate techniques may identify combinations of greater heterotic effect, raising the possibility of obtaining superior materials.

The grouping method proposed by Ward (1963), called minimum variance, is aimed at clustering R and S groups that minimize the sum of squares within the groups, i.e., the sum of squared errors (Ferreira, 2008). Ward's method has been preferred, in some cases, because of the graphic effect generated by the dendrogram that allows the visualization of well-defined groups (Romesburg, 1984). The Modified Location Model (MLM) procedure was proposed by Franco et al. (1998) as a substitute for the Location Model (LM) procedure, proposed by Lawrence and Krzanowski (1996). It classifies n individuals when quantitative variables p and qualitative variables q are obtained in an environment, assuming that m levels of variable W and p-multinormal variables for each subpopulation are independent.

Although the combined analysis of the quantitative and qualitative variables is a potential indicator of the existing variability in germplasm banks, few studies have adopted this strategy. This is likely due to the lack of knowledge of the statistical techniques that allow for this approach, the lack of free software that analyzes these data together, as well as the tendency of researchers to place greater importance on those variables directly related to traits measured in breeding programs (Gonçalves et al., 2008).

Among the traits evaluated in popcorn breeding programs, to meet the interest of farmers, agroindustry, and consumers, grain yield and popping expansion have been considered those of highest importance (Pereira and Amaral Júnior, 2001). In addressing the great damage caused by *Fusarium* in corn, the development of resistant cultivars signifies intense work, involving a priori the identification of sources of resistance with subsequent elucidation of the type of genetic control. In this regard, germplasm banks have an indispensable importance concerning the inclusion of genetic variability, especially in that they provide bases for plant-breeding programs in the generation of superior genotypes.

The Ward-MLM strategy has been employed to measure genetic diversity in fruit trees and other agricultural crops such as *Brassica rapa* (Padilla et al., 2005), tomato (Gonçalves et al., 2009), common bean (Cabral et al., 2010), green bean (Barbé et al., 2010), banana (Pereira et al., 2012), guava (Campos et al., 2013), passion fruit (Silva et al., 2014), papaya (Nunes da Luz et al., 2014), and coffee (Rodrigues et al., 2016). There are no reports in the literature on popcorn breeding programs under a multivariate approach using the Ward-MLM methodology. Therefore, the present study is a pioneering proposal that may generate interesting results for the advancements of studies of popcorn breeding.

This study aimed to estimate the genetic diversity among 37 genotypes of popcorn by the Ward-MLM method to identify divergent groups that comprise traits of interest and a higher heterotic effect in subsequent crosses.

MATERIAL AND METHODS

Thirty-seven genotypes belonging to the Popcorn Germplasm Bank of Universidade Estadual do Norte Fluminense (UENF) were used in the experiment (Table 1).

The experiment was conducted in 2014, at the Experimental Unit of UENF, located at the Antônio Sarlo State College of Agriculture (21º45'S, 41°20'W, 11 m in altitude), in the city of Campos dos Goytacazes, RJ, Brazil. The period from planting to harvest was between May and September. During this period, the main meteorological variables were recorded daily in the Database of the National Institute of Meteorology (Instituto Nacional de Meteorologia) and the monthly average was calculated and represented graphically (Figure 1). The soil of the experimental area is classified as Dystrophic Yellow Argisol (Ultisol) (EMBRAPA, 2006).

Table 1. Description of popcorn genotypes from the Germplasm Bank of the State University of Northern Rio de Janeiro.

Genotypes	Type	Population of origin	Climatic adaptation	Developed by
L88	Line	Viçosa: UFV	Temperate/Tropical	UENF
L70	Line	BRS Angela	Tropical	UENF
L65	Line	BRS Angela	Tropical	UENF
L80	Line	Viçosa: UFV	Temperate/Tropical	UENF
L51	Line	Beija-flor	Temperate/Tropical	UENF
L77	Line	Viçosa: UFV	Temperate/Tropical	UENF
L76	Line	Beija-flor	Temperate/Tropical	UENF
L75	Line	Viçosa: UFV	Temperate/Tropical	UENF
L66	Line	BRS Angela	Tropical	UENF
L53	Line	Beija-flor	Temperate/Tropical	UENF
L52	Line	Beija-flor	Temperate/Tropical	UENF
L55	Line	Beija-flor	Temperate/Tropical	UENF
L54	Line	Beija-flor	Temperate/Tropical	UENF
L61	Line	BRS Angela	Tropical	UENF
L63	Line	BRS Angela	Tropical	UENF
L59	Line	Beija-flor	Temperate/Tropical	UENF
L71	Line	BRS Angela	Tropical	UENF
P1	Line	Zaeli Hybrid	Temperate/Tropical	UEM
P2	Line	CMS-42 Compound	Temperate/Tropical	UEM
P3	Line	CMS-42 Compound	Temperate/Tropical	UEM
P4	Line	South American races	Temperate/Tropical	UEM
P5	Line	Zaeli Hybrid	Temperate/Tropical	UEM
P6	Line	Zaeli Hybrid	Temperate/Tropical	UEM
P7	Line	Zaeli Hybrid	Temperate/Tropical	UEM
P8	Line	IAC112 Hybrid	Temperate/Tropical	UEM
P9	Line	IAC112 Hybrid	Temperate/Tropical	UEM
P10	Line	IAC112 Hybrid	Temperate/Tropical	UEM
BOYA 462	Race	Bolivia	Temperate	CIMMYT
URUG 298 Roxo	Race	Uruguay	Temperate	CIMMYT
URUG 298 AM	Race	Uruguay	Temperate	CIMMYT
BOZM 260	Race	Bolivia	Temperate	CIMMYT
ARZM 07- 49	Race	Argentina	Temperate	CIMMYT
CHZM 13-0134	Race	Chile	Temperate	CIMMYT
ARZM 13- 050	Race	Argentina	Temperate	CIMMYT
ARZM 05-083	Race	Argentina	Temperate	CIMMYT
PARA 172	Race	Paraguay	Temperate	CIMMYT
IAC 125	Triple-cross hybrid	Lines	Temperate/Tropical	IAC

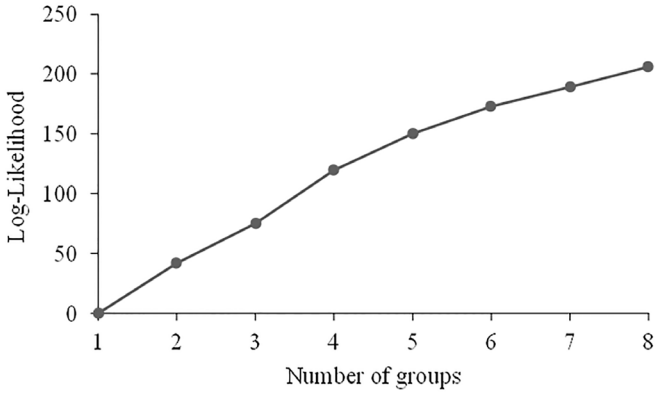

Figure 1. Graph of the log-likelihood function for the number of groups formed by the Ward-MLM strategy in popcorn genotypes.

The genotypes were arranged in a randomized block design with four replications, totaling 148 experimental plots consisting of 3-m rows with 16 plants/row. The experimental area was prepared mechanically by harrowing with later plowing and base fertilization with the application of a 4-14-8 NPK formulation, following technical recommendations. Cultivation practices included thinning the seedlings 15 days after emergence, pest control, manual weeding, regular irrigation with maintenance of soil under field capacity, and top-dressing 30 and 45 days after emergence with a 20-0-20 NPK formulation containing 300 and 200 kg/ha urea, respectively.

Seven qualitative and 17 quantitative variables were assessed. Variables were characterized based on the list of minimum descriptors for the corn crop, as established by the Ministry of Agriculture, Livestock, and Supply (Ministério da Agricultura, Pecuária e Abastecimento) (Table 2). For the evaluation of qualitative variables, traits during the following plant stages were considered: stage R1 "mid-anthesis stage": goose-necking degree or stalk curvature (GND), stigma color given by anthocyanin (STCA), and anther color given by anthocyanin (ACA); stage R3, "grain milk stage": sheath pigmentation given by anthocyanin (SPA) and root color given by anthocyanin (RCA); stage R6 "harvest stage": "grain milk" stover color (SC2); and, finally, post-harvest stage: pericarp color.

The following quantitative variables were evaluated: days to female flowering, obtained by quantifying the period between planting and the release of the style-stigmas of at least 50% of the plants in the row, recorded every two days; plant height (PH), in m, corresponding to the distance between the insertion of the sheath of the flag leaf and the point of insertion of the stalk in the soil; stalk diameter (SD), in cm; proportion of lodged plants, obtained by counting the plants that showed a slope angle greater than 45° relative to the vertical, at the harvesting; proportion of broken plants (PBP), obtained by counting the plants that showed the stalk broken below the upper leaf in each plot, at harvesting; number of ears (NE), the total number of ears harvested in each plot; number of diseased ears (NDE), obtained by counting the diseased ears per plot; height of insertion of the first ear, in cm, determined as the distance between the insertion of the first ear and the insertion of the stalk in the soil; ear diameter, in cm, measured in the mid-point, in all ears, collected at random in the plot, using a digital caliper; ear weight, in g, obtained by weighing on a precision scale (considering two decimal places) with two replications, for which 100 grains were randomly collected from distinct plants from each plot; grain yield (GY), in kg/ha, determined by weighing the grains in each plot; 100-grain weight (P100G), in g, obtained by weighing, on a scale with precision of two decimal places, 100 grains collected at random from distinct plants from each plot; percentage of rotten kernels (RK), according to the scale of the International Maize and Wheat Improvement Center (CIMMYT) (1985); popping expansion (PE), determined by heating 30 g seeds in Kraft paper bags for 3 min in a microwave oven, in two replications per experimental unit. Afterwards, the expanded-popcorn volume was quantified in a 2000-mL beaker, and the PE was expressed in mL/g; length of the first leaf below the first ear; and number of secondary branches in the tassel. The quantitative traits were obtained by randomly taking six competitive plants per experimental unit.

To determine the incidence of *Fusarium* spp in the ear (IFE), the harvested experimental plots had their ears evaluated separately, and those showing symptoms of grain infection were identified. Initially, the ears whose symptoms were most recurrent among the plots were sent to the Plant Pathology Unit of the Laboratory of Entomology and Plant Pathology at UENF, for the identification of the causal agent. The analysis was performed according to the filter

paper method (Blotter test). Subsamples were obtained by collecting 16 grains located in the periphery of the lesions of each symptomatic ear. Subsamples were kept at room temperature ±25°C for seven days, until the appearance of colonies of fungi. Later, the colonies were photographed and compared with the photos of the ears symptoms, aiming to correlate the causal agent and its respective visual symptom. Once the visual symptom was attributed to *Fusarium* spp, the remaining ears of the plots obtained in the field were subjected to analysis of incidence by threshing the corn and subsequently separating 100-grain samples, which represented the studied samples. Values are expressed in percentage terms.

The qualitative and quantitative variables were analyzed simultaneously using the Ward-MLM procedure to compose the groups of progenies via the CLUSTER and IML procedures of the SAS software (SAS Institute, 2009). Subsequently, Gower's algorithm (Gower, 1971) was used to obtain the distance matrix to use Ward's clustering method. The Gower index utilizes qualitative and quantitative data to generate a single dissimilarity index that ranges from 0 to 1. Dissimilarity was given by:

$$S_{ij} = \frac{\sum_{k=1}^{p} W_{ijk} . S_{ijk}}{\sum_{k=1}^{p} W_{ijk}}$$

where i and j represent the individuals to be compared with respect to the trait k; p = total number of traits; and S_{ij} = contribution of variable k to total distance. If one variable is qualitative, S_{ijk} assumes the value 1, when the agreement is positive or negative for the trait k between individuals i and j. On the other hand, when the variable is quantitative, then:

$$S_{ij} = \left| \frac{Y_{ik} - Y_{jk}}{R_k} \right|$$

where R_k = amplitude of variation of variable k, whose values range between 0 and 1. The W_{ijk} value was used to define the contributions of individuals S_{ijk}. Thus, when the value of variable k is absent in one or more individuals, $W_{ijk} = 0$; otherwise, it is equal to 1.

RESULTS AND DISCUSSION

The Ward-MLM procedure, using quantitative and qualitative variables simultaneously, was efficient to discriminate between the 37 genotypes. By the pseudo-F and pseudo-t^2 criteria, the ideal number of groups was established as four. The risk profile, associated with the likelihood ratio test, showed that the highest increase in the likelihood function occurred in group four, with an increase of 44.73 (Figure 1).

The likelihood function analysis can be used if the relationship between n observations and p variables is higher than 5 and n > 50. In any case, the ratio likelihood or the growth of the risk is a useful guide to define the number of groups. Therefore, the method can define more precise group-formation criteria, resulting in less subjective adhesion groups. According to Gonçalves et al. (2009), the number of groups can vary according to species, number of accessions, and number and type of descriptor.

Considering the qualitative traits evaluated at stage R1 - GND, STCA, and ACA -, GND had 51 and 46% of the evaluated genotypes with absent stalk curvature and slightly curved stalk, respectively (Table 2).

Table 2. Variables and number of genotypes per group of qualitative traits in each of the four groups (GI, GII, GIII, and GIV), formed by the Ward-MLM strategy.

Variables	GI (10)	GII (11)	GIII (9)	GIV (7)
Goose-necking degree or stalk curvature				
Absent	3	6	4	6
Slightly recurved	7	5	4	1
Strongly recurved	-	-	1	-
Stigma color given by anthocyanin				
Low	7	6	5	6
Medium	2	4	3	1
Strong	1	1	1	-
Anther color given by anthocyanin				
Low	6	5	7	3
Medium	4	6	2	4
Strong	-	-	-	-
Sheath pigmentation given by anthocyanin				
Absent	10	4	9	7
Medium	-	4	-	-
Strong	-	3	-	-
Root color given by anthocyanin				
Absent	3	-	4	2
Medium	7	6	4	5
Strong	-	5	1	-
Stover color				
Purple	-	5	-	-
Brown	10	6	9	7
Pericarp color				
Yellow	-	-	-	1
Orangish-yellow	8	2	4	6
White		2	3	-
Variegated	2	-	2	7

For the traits related to anthocyanin pigmentation (Table 2), there was a predominance of weak pigmentation for both STCA and ACA, in 65 and 57% of the evaluated genotypes, respectively. Smith and Smith (1989) observed that pigmentation with anthocyanin in the anthers and stigma is an excellent trait to distinguish between genotypes, due to its repeatability in several years and places, and it is valid to use it as a phenotypic descriptor of corn, depending on the genotypes assessed.

For the qualitative traits evaluated in stage R3 - SPA and RCA -, groups I, III, and IV showed 100% of their genotypes lacking SPA (Table 2), corresponding to 81% of the total evaluated. However, for RCA, medium pigmentation prevailed, and this attribute was present in 59% of the evaluated genotypes.

For the trait evaluated in stage R6 - SC2 -, the brown color of ear stover predominated, corresponding to 86% of the evaluated genotypes, with groups I, III, and IV showing 100% of their genotypes with this type of color. Pericarp thickness, evaluated post-harvest, had the orange-yellow color predominating, corresponding to 54% of the total, while the variegated color was attributed to 30% of the evaluated genotypes.

Based on the quantitative descriptors, a large phenotypic variation was detected among the evaluated genotypes (Table 3).

Table 3. Mean values for the quantitative variables of each of the four groups formed by the Ward-MLM method and coefficients of quantitative variables pertaining to the first two canonical variables.

Variables	GI (10)	GII (11)	GIII (9)	GIV (7)	CAN1	CAN2
DFF	65.05	67.27	66.97	69.53	-0.27	0.26
PH	1.60	1.79	1.27	1.41	0.55	0.49
SD	16.67	18.15	14.06	14.05	0.72	0.39
NPL	2.25	2.86	3.25	2.60	-0.01	-0.01
NBP	1.15	1.88	0.86	1.10	0.22	0.30
LFE	75.1	76.05	64.37	67.40	0.53	0.30
STB	12.68	19.49	11.62	13.88	0.41	0.74
NE	18.45	16.27	13.05	7.5	0.77	-0.23
EW	0.75	0.84	0.35	0.12	0.80	0.10
ED	31.89	33.43	27.71	22.54	0.87	0.03
HIE	0.92	1.14	0.72	0.73	0.68	0.49
NDE	2.4	1.75	1.22	1.78	0.11	0.01
IFE	18.72	15.63	12.66	33.62	-0.43	0.25
100GW	10.87	11.94	11.53	8.06	0.56	-0.14
PRK	1.853	1.07	1.72	5.68	-0.62	0.17
GY	1965.75	2,222.89	883.61	315.08	0.83	0.14
PE	27.31	15.50	26.60	26.78	0.05	-0.88

DFF = days to female flowering; PH = plant height; SD = stalk diameter; NLP = number of lodged plants; NBP = number of broken plants; LFE = length of the first leaf below the first ear; STB = number of secondary tassel branches; NE = number of ears; EW = ear weight; ED = ear diameter; HIE = height of insertion of the first ear; NDE = number of diseased ears; IFE = incidence of *Fusarium* in the ear: 100GW = 100-grain weight; PRK = percentage of rotten kernels; GY = grain yield; PE = popping expansion.

The greatest contributions to diversity analysis came from the variables ED, GY, EW, NE, SD, ear height, and RK. This finding demonstrates the importance of these traits for studies of genetic diversity and, consequently, for the choice of parents to be used in crosses that optimize the process of development of popcorn cultivars (Table 3).

Of the evaluated quantitative variables, the number of lodged plants (NLP) was the variable that least contributed to the diversity analysis (Table 3). This can be verified by the correlation between NLP and the first canonical variable (CAN1). NLP and number of broken plants are traits of importance in popcorn, and these variables should not be missed, since they are correlated with stalk thickness, which implies greater yields during mechanized harvesting.

Considering the evaluated traits, those that most contributed to the analysis of genetic diversity, based on the first canonical variable, were EW, ED, and GY (Table 3). GY ranged from 315 to 2222 kg/ha for groups IV and II, respectively (Table 3). However, the group of highest yields obtained the lowest mean for PE, which is explained by the existence of negative correlations between GY and PE (Dofing et al., 1991; Daros et al., 2004; Amaral Júnior et al., 2010). Because PE is the most important trait for popcorn and has a typically additive heritability (Larish and Brewbaker, 1999; Pereira and Amaral Júnior, 2001) interpopulation breeding methods aimed at the exploitation of heterosis are not applicable to these groups. Thus, the best breeding strategy would be the use of intrapopulation methods with populations originating mainly from group I, which comprises good yield and PE.

Considering the set of traits related to the vegetative stages, groups I and IV stood out once again (Table 3) for having the closest patterns of modern common-corn genotypes. However, other traits of interest are also found in the other groups, such as IFE, RK, and NDE. The percentage of diseased ears is an important trait in popcorn because of the need for them to be free of toxins, which negatively affect the PE.

Groups I and II include the greatest PHs, and consequently the greatest ear heights (Table 3), since these two variables are correlated (Cabral et al., 2016). Taller plants tend to be

more prone to stalk breaking, especially in regions with a high incidence of strong winds; also, in the evaluation of yield, the ears of broken plants are harvested and processed, generating the sample that will be analyzed for PE. In this way, one can argue that a decrease in the PE of a given family tends to be influenced by damage to the grains caused by the breaking of plants.

The largest distances were observed between groups II and IV (198.98) and between groups I and IV (152.30), whereas the shortest distance was observed between groups I and III (30.23) (Table 4).

Table 4. Distance between groups formed by the Ward-MLM procedure, proposed by Franco et al. (1998).

Groups	I	II	III
II	31.66		
III	30.23	69.49	
IV	152.30	198.98	85.70

The crosses between genotypes from the other groups with those of group IV would be interesting when aiming to exploit heterosis in the production of transgressive individuals for the traits of interest. The importance of the present study lies in the fact that popcorn cultivars are not distributed into complementary heterotic groups for hybrid production, as occurs with common corn, in the groups named 'DENT' and 'FLINT'.

In breeding programs, it is recommended to use parents with the largest divergence possible so as to maximize the possibility of occurrence of superior segregants in advanced generations and expand the genetic base. However, the choice of genotypes should be made also considering their behaviors per se. Thus, for crosses, divergent genotypes that have superior performance for the main traits of agronomic importance and that meet the goals of the breeding programs for which they are being developed should be used. Therefore, the breeding potential of cultivars is determined by the high mean values for the most important agronomic traits and the position of the cultivars in different groups obtained by the estimates of genetic divergence.

The cross between genotypes belonging to groups I and II may result in the generation of progenies with high similarity. However, depending on the strategy of the breeding program, this type of cross, considered convergent, may facilitate the work of breeders in the selection of superior lines in less time, since both groups have higher mean values for important agronomic traits, such as production potential.

The formation of groups can observed in the graphic representation of the first two canonical variables (CAN1 and CAN2), which explained 94.06% of the observed variation, allowing a clear understanding of the genetic variability among the genotypes evaluated by a scatter graph (Figure 2).

This high value indicates that the graphic representation of the first two canonical variables is suitable for examining the relationship between groups and individuals within a group. The analysis of canonical variables allows the discard of those traits that contributed little to the genetic variability between the evaluated cultivars, which saves on time, labor, and financial resources in future studies (Cruz et al., 2014). The graphic analysis of the first two canonical variables shows the distancing of groups I and IV in relation to groups II and III.

It is also noteworthy that group I includes a set of desirable traits for popcorn, making it a great option for the extraction of populations and lines. The genotypes belonging to group

I are mostly from CIMMYT and have already been mentioned as a reservoir of genes for the generation of new lines and for increasing the variability of popcorn germplasm banks (de Carvalho et al., 2013).

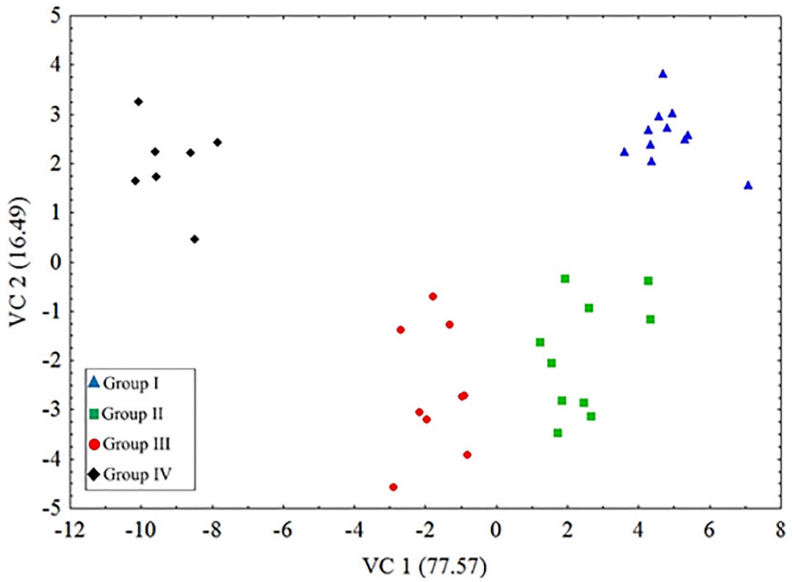

Figure 2. Graph of the first two canonical variables for the four groups formed by the Ward-MLM analysis.

The genetic divergence observed among the popcorn genotypes indicates the variability contained in the germplasm bank of the UENF. In addition, future studies may be conducted to obtain hybrid combinations that associate high production and resistance to *Fusarium* spp in the ear.

CONCLUSIONS

The Ward-MLM and canonical variables procedures were useful tools to detect genetic diversity and to group popcorn accessions by using qualitative and quantitative variables simultaneously.

This study brings useful information that allows the breeder to plan the direction of popcorn breeding programs under development.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors thank to FAPER and CNPq for their financial support to this experiment, and CAPES, for the Master's degree scholarship granted to the first author.

REFERENCES

Amaral Júnior AT, Freitas Júnior SP, Rangel RM, Pena GF, et al. (2010). Improvement of a popcorn population using selection indexes from a fourth cycle of recurrent selection program carried out in two different environments. *Genet. Mol. Res.* 9: 340-347.

Barbé TC, Amaral Júnior AT, Gonçalves LSA, Rodrigues R, et al. (2010). Association between advanced generations and genealogy in inbred lines of snap bean by the Ward-Modified Location Model. *Euphytica* 173: 337-343. http://dx.doi.org/10.1007/s10681-009-0089-z

Bartlett MS (1939). A note on tests of significance in multivariate analysis. *Math. Proc. Camb. Philos. Soc.* 35: 180-185. http://dx.doi.org/10.1017/S0305004100020880

Bezerra Neto FV, Leal NR, Gonçalves LSA, Rêgo Filho LM, et al. (2010). Descritores quantitativos na estimativa da divergência genética entre genótipos de mamoneira utilizando análises multivariadas. *Rev. Cienc. Agron.* 41: 294-299. http://dx.doi.org/10.1590/S1806-66902010000200018

Cabral PDS, Soares TCB, Gonçalves LSA, Amaral Júnior ATD, et al. (2010). Quantification of the diversity among common bean accessions using Ward-MLM strategy. *Pesqui. Agropecu. Bras.* 45: 1124-1132. http://dx.doi.org/10.1590/S0100-204X2010001000011

Cabral PDS, Amaral Júnior AT, Freitas ILJ, Ribeiro RM, et al. (2016). Relação de causa e efeito de caracteres quantitativos sobre a capacidade de expansão do grão em milho-pipoca. *Rev. Cienc. Agron.* 47: 108-117.

Campos BM, Viana AP, Quintal SSR, Gonçalves LSA, et al. (2013). Quantificação da divergência genética entre acessos de goiabeira por meio da estratégia Ward-MLM. *Rev. Bras. Frutic*. 35: 087-094.

Cargnelutti Filho A, Ribeiro ND, Reis RCPD, Souza JRD, et al. (2008). Comparação de métodos de agrupamento para o estudo da divergência genética em cultivares de feijão. *Cienc. Rural* 38: 2138-2145. http://dx.doi.org/10.1590/S0103-84782008000800008

Cruz CD, Carneiro PCS and Regazzi AJ (2014). Modelos biométricos aplicados ao melhoramento genético. 3rd edn. Ed. UFV, Viçosa.

Daros M, Amaral Júnior AT, Pereira MG, Santos FS, et al. (2004). Correlações entre caracteres agronômicos em dois ciclos de seleção recorrente em milho-pipoca. *Cienc. Rural* 34: 1389-1394.

de Carvalho MS, Mangolin CA, Scapim CA, da Silva TA, et al. (2013). A collection of popcorn as a reservoir of genes for the generation of lineages. *Mol. Biotechnol.* 53: 300-307. http://dx.doi.org/10.1007/s12033-012-9527-6

Dofing SM, ĎCroz-Mason N and Thomas-Compton MA (1991). Inheritance of expansion volume and yield in two popcorn X dent corn crosses. *Crop Sci.* 31: 715-718.

EMBRAPA (2006). Centro Nacional de Pesquisa do Solo. Sistema brasileiro de classificação de solos. 2nd edn. Embrapa Solos, Rio de Janeiro.

Ferreira DF (2008). Estatística multivariada. 1st edn. Ed. UFLA, Lavras.

Fisher RA (1928). The general sampling distribution of the multiple correlation coefficient. *Proc. R. Soc. Lond.* 121: 654-673. http://dx.doi.org/10.1098/rspa.1928.0224

Franco J, Crossa J, Villaseñor J, Taba S, et al. (1998). Classifying genetic resources by categorical and continuous variables. *Crop Sci.* 38: 1688-1696. http://dx.doi.org/10.2135/cropsci1998.0011183X003800060045x

Gonçalves LSA, Rodrigues R, Amaral Jr AT, Karasawa M, et al. (2008). Comparison of multivariate statistical algorithms to cluster tomato heirloom accessions. *Genet. Mol. Res.* 7: 1289-1297. http://dx.doi.org/10.4238/vol7-4gmr526

Gonçalves LSA, Rodrigues R, do Amaral Júnior AT, Karasawa M, et al. (2009). Heirloom tomato gene bank: assessing genetic divergence based on morphological, agronomic and molecular data using a Ward-modified location model. *Genet. Mol. Res.* 8: 364-374. http://dx.doi.org/10.4238/vol8-1gmr549

Gower JC (1971). A general coefficient of similarity and some of its properties. *Biometrics* 27: 857-871. http://dx.doi.org/10.2307/2528823

Hotelling H (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24: 417-441. http://dx.doi.org/10.1037/h0071325

International Maize and Wheat Improvement Center (CIMMYT) (1985). International Progeny Testing Trials. Available at [http://repository.cimmyt.org/xmlui/bitstream/handle /10883/697/13201.pdf]. Accessed August 3, 2016.

Larish LLB and Brewbaker JL (1999). Diallel analyses of temperate and tropical popcorns. *Maydica* 44: 279-284.

Lawrence CJ and Krzanowski WJ (1996). Mixture separation for mixed-model data. *Stat. Comput.* 6: 85-92. http://dx.doi.org/10.1007/BF00161577

Nunes da Luz L, Cavalcanti dos Santos R, Melo Filho PDA, Gonçalves A, et al. (2014). Combined selection and multivariate analysis in early generations of intraspecific progenies of peanuts. *Chil. J. Agric. Res.* 74: 16-22.

Padilla G, Cartea ME, Rodríguez VM and Ordás A. (2005). Genetic diversity in a germplasm collection of *Brassica rapa*

subsp rapa L. from northwestern Spain. *Euphytica* 145: 171-180.
Pearson K (1901). On lines and planes of closest fit to systems of point in space. *Philos. Mag.* 2: 559-572. http://dx.doi.org/10.1080/14786440109462720
Pereira MG and Amaral Júnior AT (2001). Estimation of genetic components in popcorn based on the nested design. *Crop Breed. Appl. Biotechnol.* 1: 3-10. http://dx.doi.org/10.13082/1984-7033.v01n01a01
Pereira VM, Borges CV, Brandão LP, Oliveira LS, et al. (2012). Genetic diversity between improved banana diploids using canonical variables and the Ward-MLM method. *Pesqui. Agropecu. Bras.* 47: 1480-1488. http://dx.doi.org/10.1590/S0100-204X2012001000010
Rodrigues WP, Vieira HD, Teodoro PE, Partelli FL, et al. (2016). Assessment of genetic divergence among coffee genotypes by Ward-MLM procedure in association with mixed models. *Genet. Mol. Res.* 15: 1-7. http://dx.doi.org/10.4238/gmr.15027889
Romesburg CH (1984). Cluster Analysis for Researchers. Lifetime Learning Publications, Belmont.
SAS Institute (2009) SAS/STAT: User's guide. Version 9.2. SAS Institute, Cary.
Silva FHL, Viana AP, Ferreira RT, Freitas JCO, et al. (2014). Measurement of genetic diversity in progenies of sour passion fruit by Ward-MLM methodology: a strategy for heterotic group formation. *Rev. Cienc. Agron* 38: 1234-1239.
Smith JSC and Smith OS (1989). The description and assessment of distance between inbred lines of maize. 2: The utility of morphological-biochemical-and genetic descriptors and a scheme for the testing of distinctiveness between inbred lines. *Maydica* 34: 151-161.
Ward JH, Junior. (1963). Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58: 236-244. http://dx.doi.org/10.1080/01621459.1963.10500845
Wilks SS (1932). Certain generalizations in the analysis of variance. *Biometrika* 24: 471-494. http://dx.doi.org/10.1093/biomet/24.3-4.471