



# Multivariate analysis of backcross progeny of *Passiflora* L. (Passifloraceae) for pre-breeding genotype selection

C.A.F. Melo<sup>1</sup>, M.M. Souza<sup>1</sup>, A.G.R. Sousa<sup>1</sup>, A.P. Viana<sup>2</sup> and E.A. Santos<sup>2</sup>

<sup>1</sup>Laboratório de Melhoramento de Plantas, Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz, Ilhéus, BA, Brasil

<sup>2</sup>Centro de Ciências e Tecnologias Agropecuárias, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, RJ, Brasil

Corresponding author: M.M. Souza  
E-mail: souzamagg@yahoo.com.br

Genet. Mol. Res. 14 (4): 15376-15389 (2015)

Received February 3, 2015

Accepted July 23, 2015

Published November 30, 2015

DOI <http://dx.doi.org/10.4238/2015.November.30.15>

**ABSTRACT.** The Ward-MLM procedure was used to evaluate genetic variation in four backcross progenies and in their parents, hybrid F1 HD13 and donor parent *Passiflora subblanceolata*. Sixteen quantitative descriptors and five qualitative characteristics of relevance to ornamental flower production were assessed. Using the pseudo-F and pseudo-T<sup>2</sup> criteria, we identified four groups among these plants in two evaluation periods. In both evaluations, the BC<sub>1</sub> plants showed greater dissimilarity to their recurrent parent, but showed high genetic similarity with the *P. subblanceolata* parent. The first two canonical variables produced by the Ward-MLM procedure accounted for over 90% of the variation in both evaluation periods, enabling the representation of diversity through two-dimensional graphics. Groups II and IV were formed in the first assessment period. Groups I and IV formed in the second period and showed plants with selection potential. We found that it was essential to use both qualitative and quantitative variables for this analysis. Assessments of quantitative descriptors indicate that the

selection of BC<sub>1</sub> plants can be performed in any of the four progenies. Because of the similarities observed for some floral descriptors between BC<sub>1</sub> and the *P. subanceolata* parent, a second generation backcross was not recommended. However, the selection of BC<sub>1</sub> plants for evaluation and direct use as an ornamental cultivar, or as a resource in other breeding programs, can be recommended.

**Key words:** Multivariate analysis; Interspecific hybrid; Plant breeding; Modified location model.

## INTRODUCTION

The genus *Passiflora* L. is of great commercial importance not only in the food industry, but also for ornamental use (Ulmer and MacDougal, 2004). The ornamental potential of the genus is highly valued in Europe and North America and many of the varieties used for this purpose are derived from interspecific hybrids (Abreu et al., 2009). Interspecific hybrids are comparatively easy to generate in this genus because of weak reproductive barriers and weak self-incompatibility; these characteristics are a hallmark of the genus (Schifino-Wittmann and Dall'Agnol, 2002). In general, interspecific hybridization has been the main strategy used for breeding ornamental *Passiflora* varieties and for the improvement of quantitative and qualitative characteristics of the plants and the development of genotypes with unique characteristics for the ornamental plant market (Abreu et al., 2009; Belo, 2010; Santos et al., 2012). Genetic variability is the main advantage of interspecific hybrids, especially in high quality allogamic species (Lippman and Zamir, 2007; Sanghera et al., 2011). However, interspecific hybrids have occasionally been found to show low performance, for example, in *Phaseolus vulgaris* L. (Koinange and Gepts, 1992), *Solanum* L. (Veilleux and Miller, 1998) and *Oryza* L. (Sobrizal and Yoshimura, 2009), preventing their use in genetic improvement programs.

In breeding programs involving interspecific hybridization, it is common to lose important agronomic traits post-hybridization. Therefore, plant breeders regularly employ backcrosses for genome introgression from an elite donor in plant varieties/species with a recurring genotype, with the intention of improving the varieties (Borém and Miranda, 2005; Bueno et al., 2006). Post-hybridization selection is a problem in conducting backcrosses due to the success of selection being directly proportional to the sample number analyzed. Currently, there is little information on the generation of ornamental *Passiflora* cultivars by backcrossing.

Various procedures can be used to characterize germplasm: simple descriptive analysis of quantitative and qualitative variables; estimation of genetic parameters, which reveals the heritability of characteristics and assesses the influence of the environment on the expression of evaluated characteristics; and molecular analyses based on the segregation of polymorphic loci (Santos et al., 2011; Pereira et al., 2012). Multivariate methods are ideal for describing genetic diversity and can be carried out by analysis of quantitative and qualitative descriptors or by analysis of the binary data obtained from molecular information (Sudré et al., 2010; Santos et al., 2011; Pereira et al., 2012). However, few studies have actually used multivariate analyses for this purpose, or have excluded variables that might contribute effectively to the quantification of genetic diversity and be applied for genetic improvement and germplasm conservation (Crossa and Franco, 2004; Gonçalves et al., 2009; Pereira et al., 2012). The Modified Location Model (MLM) was developed in order to analyze all the variables together (Franco et al., 1998). This procedure has

two steps: the first is hierarchical (Ward) grouping (Ward, 1963) of data in a Gower dissimilarity array (Gower, 1971); the second phase is to use the MLM procedure on each sub-population in order to estimate the vector averages of the quantitative variables. The Ward-MLM strategy not only allows the inclusion of all available data for the characterization of germplasm (quantitative, qualitative, and binary), but also defines the number of groups with probabilistic support for each genotype of a specific group (Gonçalves et al., 2009; Cabral et al., 2010; Sudré et al., 2010). The Ward-MLM procedure has been applied to diverse cultures to analyze genetic variation. It provides an indication of the descriptors that give the highest contribution to group formation, and aids in the recommendation of germplasm with potential for genetic improvement (Campos et al., 2013; Santos, 2013). In *Passiflora*, the use of the Ward-MLM technique enabled the effective identification of parental plants in a breeding program aimed at improving resistance to the *Cowpea aphid-borne mosaic virus* (Santos, 2013).

The purpose of the present study was to estimate genetic variation among the backcross progeny of crosses between a recurrent parent (HD13 hybrid) and a parent donor (*Passiflora sublancoolata*). To this end, we used a joint analysis of quantitative and qualitative variables with a multivariate Ward-MLM procedure. This analysis not only measures genetic variation but also identifies promising progeny for use in breeding programs for improved ornamental varieties.

## MATERIAL AND METHODS

### Plant material and cultivation conditions

Four BC<sub>1</sub> progenies (HD17 to HD20), five cloned plants of each recurrent parent (HD13-101, HD13-133, HD13-134, and HD13-143), and a parent donor (*P. sublancoolata*) were evaluated using the multivariate approach. Reciprocal crosses were performed (e.g., *P. sublancoolata* x HD13 and HD13 x *P. sublancoolata*). Additionally, the species *Passiflora foetida*, that represents the male genitor of the hybrids HD13 was cloned and included in the analysis (Santos et al., 2012). The backcross progenies and their parents are listed in Table 1. The experiment was conducted in the experimental unit on the campus of Universidade Estadual de Santa Cruz, Ilhéus, Bahia (longitude 39°10'W, latitude 14°39'S, at an altitude of 78 m) during the period November 2011 to July 2013. The plants were each grown in a 25-L vessel containing vegetable soil as the substrate, with a vertical trellis system using 2-m fence posts and a wire mesh 1.5 m above the ground; the plants were distributed in a random block design of three blocks and four installments, with approximately 1 m between plants and 0.7 m between each row. Plants were managed with weekly pruning and were fertilized monthly using 650 mg CH<sub>4</sub>N<sub>2</sub>O, 68 mg B, 34 mg Cu, 2 mg (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>, 100 mg Zn, and 76 mg Mn.

### Quantitative and qualitative descriptors

Morphometric data were obtained in two consecutive years (2012-2013), representing two evaluation periods. The descriptors selected for morphological assessment were those most relevant to *Passiflora* ornamental plants (MAPA, 2008). Measurements were made using a digital caliper. Eleven quantitative floral descriptors were measured: flower diameter (FD) from points at both sides of the flower; corona diameter (DCO) at the two ends of the corona; petal width, using the larger petals; petal length (CPE) from base to apex; sepal width using the larger sepals;

sepal length (CSE) from base to apex; width of bract, using the larger bracts; length of the bract from base to apex; size of the first coronal filaments (TF1) from the insertion of the flower into the receptacle to the apex of the major filament; the size of the second coronal filaments from the insertion of the flower receptacle into the longest apex; length of the filament and the flower stalk from the insertion of the flower stem to the top of flower box. Five vegetative descriptors were evaluated: leaf length and leaf width; leaf area; main branch diameter from the first internode; and the main branch internode length (CEN). All measurements were made in mm.

Additionally, five qualitative characteristics were evaluated according to Munsell's Plant Tissue Color chart (Munsell, 1977), which uses a letter and number code based on primary and secondary colors and other variables such as saturation and luminescence (contrast). The qualitative characteristics evaluated were as follows: presence or absence of pattern bands on the petals and sepals (PB); the predominant petal color (COLP); predominant sepal color (COLS); and color of the first and second rows of coronal filaments.

**Table 1.** Crosses used to produce the *Passiflora* BC<sub>1</sub> progenies analyzed in the present study.

Progeny	Genitor donor	Recurrent parent	First trial period	Second trial period
HD17*	<i>Passiflora subblanceolata</i>	HD13-101	86	83
HD18	<i>P. subblanceolata</i>	HD13-133	77	80
HD19	<i>P. subblanceolata</i>	HD13-134	77	76
HD20	<i>P. subblanceolata</i>	HD13-143	78	81
Total			318	320

\*HD designation refers to the backcross progeny followed by the identifier plant number within the progeny. Number of genotypes analyzed for each progeny type in the first and second assessment periods.

## Data analysis

Each quantitative variable was measured three times per plant. For qualitative variables, the color codes obtained from Munsell's Color Chart (1977) were transformed to a numerical code and displayed as qualitative data.

The analysis of genetic variation was performed by simultaneously using phenotypic data from both qualitative and quantitative variables and by applying the Multivariate Ward-MLM procedure using the SAS software (SAS institute Inc. 2002, Carry, NC, USA). The distance matrix was obtained using the Gower logarithmic function (Gower, 1971), which defines the optimal number of groups based on pseudo-F and pseudo-T<sup>2</sup> criteria that are combined with the likelihood profile associated with the likelihood ratio test. The use of quantitative and qualitative data from the Gower index (1971) for obtaining the 0 to 1 variant dissimilarity index 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 being compared and *k* the characteristics; *p* = total number of features, and *S<sub>ij</sub>* = the contribution of *k* to the total distance. If the variable is qualitative, *S<sub>ijk</sub>* assumes the value 1 whether the correlation is positive or negative for the characteristic *k* between individuals *i* and *j*. In the case of a quantitative characteristic:

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

where  $R_k$  = variation amplitude of the variable K, having values 0 and 1. The value of  $W_{ik}$  was used to define the contributions of the individuals  $S_{ik}$ . Thus, when the value of the variable k is missing in one or both individuals,  $W_{jk} = 0$ ; if present, the value equals 1.

## RESULTS

MLM defines the group number through the logarithmic function of Log-Likelihood probability following the criteria of pseudo-F and pseudo- $T^2$  tests. The ideal number of groups established during both assessment periods was four (groups I, II, III, and IV). The greatest verisimilitude function increase occurred in group IV, with increments of 241,161 and 91,951 in the first and second characterization periods, respectively (Figure 1; Tables 2 and 3).

In the first trial period, group I comprised 25 genotypes: five clones of each HD13 parent (HD13-101, HD13-133, HD13-134, and HD13-143) and five clones of the male *P. foetida* parent (Tables 4 and 5). Group I included individuals with lower floral characteristic values for mean FD (59.24 mm) and for their related floral part lengths, namely, 27.42 mm for the petal (CPE) and 26.82 mm for the sepal (CSE) (Table 4). In the same group, genotypes with larger average values (31.28 mm) DCO were found. For the five vegetative characteristics evaluated in the first period, Group I contained genotypes with the largest LF with an average of 54.65 mm; Group I genotypes also had the lowest internode lengths with an average of 67.36 mm (Table 4). In the same evaluation period, Group II, contained 42 genotypes, including five plants of the recurrent *P. subblanceolata* parental clones and 37 BC<sub>1</sub> plants, representing the progenies of HD17, HD18, and HD19 (Table 5). Groups III and IV in the first assessment period comprised 136 and 149 BC<sub>1</sub> genotypes, respectively, with both groups containing plants from the progenies of HD17 to HD20 (Table 5).

**Table 2.** Number of groups formed based on the logarithmic function of probability (Log-Likelihood) and its increment in backcross hybrid progenies of *Passiflora* that were characterized in the first trial period.

Number of groups	Log-Likelihood	Increment
1	-14,942.61	0.000
2	-14,932.73	9.882
3	-14,908.28	34.335
4	-14,694.11	248.496*
5	-14,658.79	283.825
6	-14,619.91	322.704
7	-14,593.60	349.007
8	-14,450.55	492.056
9	-14,569.47	373.138

\*Greater increment in the formation of the four groups given by the logarithmic function.

In the second period of assessment, group I consisted of 95 BC<sub>1</sub> genotypes (Tables 6 and 7), whose main quantitative characteristics were having the highest FD with an average of 78.43 mm (Table 6). On the other hand, group II contained 64 genotypes, among them recurrent parental clones (HD13) and the male parent *P. foetida*. Additionally, the group also contained 34 BC<sub>1</sub> plants (Table 7). Group II, from the second trial period contained plants with the lowest values

**Table 3.** Number of groups formed based on the logarithmic function of probability (Log-Likelihood) and its increment in backcross hybrid progenies of *Passiflora* characterized in the second trial period.

Number of groups	Log-Likelihood	Increment
1	-14,879.46	0.000
2	-14,869.12	10.345
3	-14,841.76	37.705
4	-14,749.81	129.656*
5	-14,679.79	199.676
6	-14,647.35	232.113
7	-14,603.08	276.384
8	-14,565.67	313.790
9	-14,476.68	402.786

\*Greater increment in the formation of the four groups given by the logarithmic function.

**Table 4.** Averages of the sixteen quantitative variables for each of the four groups formed by Ward-MLM, and the first two canonical variables (CAN) in the first trial period in backcross hybrid progenies of *Passiflora*.

Trait	Group				CAN	
	I (25)	II (42)	III (136)	IV (149)	CAN1	CAN2
DF (mm)	59.24	80.70	80.56	78.16	-0.69914	-0.10028
DCO (mm)	31.28	19.40	19.22	20.04	0.70246	-0.01344
LPE (mm)	9.92	10.02	9.70	9.42	0.06059	-0.35250
CPE (mm)	27.42	36.21	36.37	35.10	-0.64048	-0.12246
LSE (mm)	9.83	10.25	10.33	10.24	-0.12630	-0.00636
CSE (mm)	26.82	36.57	36.68	35.46	-0.66318	-0.09749
LBR (mm)	25.77	26.37	26.92	26.17	-0.05116	-0.06749
CBR (mm)	28.37	31.76	32.52	33.42	-0.24138	0.25224
TF1 (mm)	11.88	8.59	9.58	9.51	0.51817	0.18844
TF2 (mm)	12.97	11.08	12.07	12.22	0.23123	0.37762
CPED (mm)	42.55	64.28	63.27	61.86	-0.42175	-0.03515
CF (mm)	74.27	77.95	80.08	81.29	-0.14776	0.20368
LF (mm)	54.65	45.19	48.86	48.08	0.29642	0.11042
AF (mm <sup>3</sup> )	24.39	26.76	23.48	21.28	0.02548	-0.39350
DRP (mm)	6.652	6.08	6.19	6.21	0.09798	0.03157
CEN (mm)	67.36	91.19	88.31	89.12	-0.46368	0.03076

DF = Flower diameter; DCO = diameter of corona; LPE = width of the petal; CPE = petal length; LSE = sepal width; CSE = sepal length; LBR = bract width; CBR = bract length; TF1 = size of the first coronal filaments; TF2 = size of the second coronal filaments; CPED = floral peduncle length; CF = length of the leaf; LF = leaf width; AF = leaf area; DRP = main branch diameter; CEN = internode length.

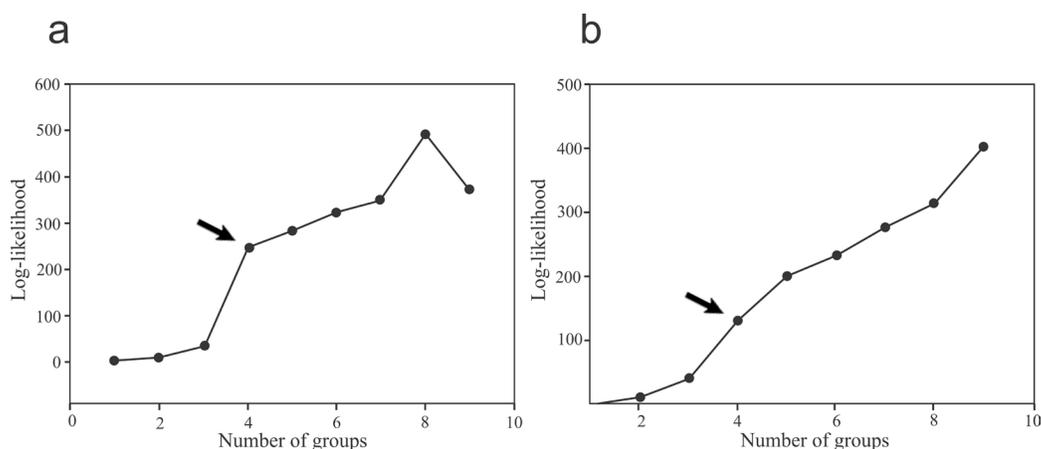
for FD, CPE, CSE, and CEN, with averages of 70.58, 31.44, 31.54, and 25.90 mm, respectively (Table 6). In group II, with the exception of CEN, all vegetative descriptors had higher values among the groups in the second evaluation period (Table 6). Groups III and IV, in the second period, comprised 135 and 55 genotypes, respectively; Group III contained only BC<sub>1</sub> plants, while Group IV contained 5 clones of the parental donor *P. subblanceolata* and 5 more BC<sub>1</sub> genotypes of all progenies, except HD20 (Table 7).

In the first period, a greater distance was found between group I (composed primarily of recurrent parents) and other groups by the Ward-MLM procedure (Table 8). On the other hand, the genotypes in groups II, III, and IV showed similar average values, particularly for floral characteristics. Thus, quantitative variables displayed a low spread between these groups. A greater spread was observed between groups I and II (41.79), while a lower spread was observed between groups III and IV (1.04); the latter two groups were represented only by BC<sub>1</sub> plants (Table 8).

**Table 5.** Groups formed by Ward-MLM for the twenty characterized variables from the first evaluation period in backcross hybrid progenies of *Passiflora*.

Group	Genitor and BC <sub>1</sub> genotype
I	<i>Passiflora foetida</i> , *HD13-101, HD13-133, HD13-134, HD13-143.
II	<i>P. subanceolata</i> , HD17-152, HD17-165, HD17-172, HD17-176, HD17-183, HD18-107, HD18-120, HD18-132, HD18-147, HD18-149, HD18-157, HD18-158, HD18-160, HD18-173, HD18-176, HD18-180, HD18-185, HD18-186, HD18-188, HD18-190, HD19-117, HD19-125, HD19-127, HD19-128, HD19-144, HD19-146, HD19-147, HD19-154, HD19-155, HD19-162, HD19-164, HD19-166, HD19-170, HD19-174, HD19-182, HD19-186, HD19-188.
III	HD17-101, HD17-102, HD17-105, HD17-108, HD17-109, HD17-113, HD17-114, HD17-116, HD17-120, HD17-130, HD17-135, HD17-138, HD17-139, HD17-140, HD17-142, HD17-143, HD17-145, HD17-146, HD17-149, HD17-154, HD17-163, HD17-170, HD17-175, HD17-180, HD17-186, HD17-188, HD18-101, HD18-103, HD18-106, HD18-113, HD18-114, HD18-115, HD18-118, HD18-119, HD18-121, HD18-122, HD18-123, HD18-125, HD18-128, HD18-129, HD18-130, HD18-134, HD18-135, HD18-136, HD18-138, HD18-140, HD18-141, HD18-142, HD18-143, HD18-144, HD18-148, HD18-170, HD18-175, HD18-183, HD19-102, HD19-107, HD19-108, HD19-109, HD19-110, HD19-111, HD19-115, HD19-119, HD19-120, HD19-121, HD19-122, HD19-123, HD19-131, HD19-132, HD19-133, HD19-134, HD19-137, HD19-138, HD19-139, HD19-142, HD19-149, HD19-150, HD19-151, HD19-152, HD19-153, HD19-157, HD19-158, HD19-159, HD19-163, HD19-169, HD19-173, HD19-175, HD19-178, HD19-181, HD19-183, HD19-185, HD19-187, HD20-101, HD20-108, HD20-115, HD20-117, HD20-119, HD20-122, HD20-123, HD20-127, HD20-130, HD20-131, HD20-134, HD20-135, HD20-139, HD20-143, HD20-145, HD20-147, HD20-149, HD20-150, HD20-152, HD20-153, HD20-155, HD20-156, HD20-158, HD20-159, HD20-160, HD20-161, HD20-162, HD20-163, HD20-165, HD20-166, HD20-167, HD20-169, HD20-173, HD20-174, HD20-176, HD20-177, HD20-178, HD20-180, HD20-181, HD20-183, HD20-184, HD20-185, HD20-186, HD20-189, HD20-190.
IV	HD17-103, HD17-104, HD17-106, HD17-107, HD17-110, HD17-111, HD17-112, HD17-115, HD17-117, HD17-118, HD17-119, HD17-121, HD17-122, HD17-124, HD17-125, HD17-126, HD17-127, HD17-128, HD17-129, HD17-131, HD17-133, HD17-134, HD17-136, HD17-137, HD17-141, HD17-144, HD17-147, HD17-148, HD17-150, HD17-151, HD17-153, HD17-155, HD17-156, HD17-157, HD17-158, HD17-159, HD17-160, HD17-160, HD17-162, HD17-164, HD17-166, HD17-167, HD17-168, HD17-169, HD17-171, HD17-173, HD17-174, HD17-177, HD17-178, HD17-179, HD17-181, HD17-184, HD17-185, HD17-189, HD17-190, HD18-102, HD18-104, HD18-105, HD18-108, HD18-109, HD18-110, HD18-111, HD18-112, HD18-116, HD18-137, HD18-150, HD18-151, HD18-152, HD18-153, HD18-155, HD18-156, HD18-159, HD18-162, HD18-163, HD18-164, HD18-165, HD18-166, HD18-167, HD18-168, HD18-169, HD18-171, HD18-172, HD18-174, HD18-177, HD18-179, HD18-181, HD18-182, HD18-187, HD18-189, HD19-104, HD19-105, HD19-106, HD19-112, HD19-113, HD19-114, HD19-118, HD19-124, HD19-126, HD19-130, HD19-135, HD19-136, HD19-140, HD19-143, HD19-160, HD19-165, HD19-171, HD19-176, HD19-177, HD19-179, HD19-180, HD19-189, HD19-190, HD20-148, HD20-151, HD20-154, HD20-157, HD20-164, HD20-168, HD20-170, HD20-171, HD20-175, HD20-102, HD20-103, HD20-104, HD20-105, HD20-106, HD20-107, HD20-109, HD20-110, HD20-111, HD20-113, HD20-116, HD20-118, HD20-120, HD20-121, HD20-124, HD20-125, HD20-126, HD20-129, HD20-132, HD20-133, HD20-136, HD20-138, HD20-140, HD20-141, HD20-144, HD20-179, HD20-187, HD20-188.

\*HD designation refers to the backcross progeny number followed by the identifier number of the plant within the progeny.



**Figure 1.** Logarithmic function probability (Log-Likelihood) with the formation of four groups in the population of backcrossed hybrid progenies of *Passiflora* characterized in the first assessment period (a) and the second trial period (b).

**Table 6.** Averages of the sixteen quantitative variables for each of the four groups formed by Ward-MLM, and the first two canonical variables (CAN) in the second assessment period in backcross hybrid progenies of *Passiflora*.

Trait	Group				CAN	
	I (97)	II (64)	III (134)	IV (55)	CAN1	CAN2
DF (mm)	78.43	70.58	76.11	77.35	-0.379431	-0.086512
DCO (mm)	18.22	25.90	18.34	17.23	0.821974	-0.036404
LPE (mm)	9.51	9.28	9.37	9.64	-0.102255	0.213038
CPE (mm)	34.91	31.44	33.95	34.72	-0.373320	-0.012488
LSE (mm)	10.12	9.85	9.88	10.35	-0.114143	0.370778
CSE (mm)	35.14	31.54	34.29	35.04	-0.384406	-0.008576
LBR (mm)	25.76	26.74	24.57	26.67	0.134558	0.427424
CBR (mm)	31.53	32.55	30.24	32.32	0.122195	0.341477
TF1 (mm)	9.33	10.92	9.14	8.54	0.703475	-0.416527
TF2 (mm)	12.08	12.76	11.88	11.49	0.367070	-0.347788
CPED (mm)	55.56	57.20	54.45	58.37	0.048971	0.310549
CF (mm)	71.77	80.14	70.56	70.84	0.340017	0.063893
LF (mm)	46.84	55.33	46.10	46.34	0.441397	0.103428
AF (mm <sup>2</sup> )	21.09	26.09	19.97	20.88	0.300443	0.126913
DRP (mm)	6.87	7.92	7.03	6.75	0.298511	-0.061742
CEN (mm)	80.48	72.09	79.46	81.49	-0.338538	0.057463

FD = Flower diameter; DCO = diameter of corona; LPE = width of the petal; CPE = petal length; LSE = sepal width; CSE = sepal length; LBR = bract width; CBR = bract length; TF1 = size of the first coronal filaments; TF2 = size of the second coronal filaments; CPED = floral peduncle length; CF = length of the leaf; LF = leaf width; AF = leaf area; DRP = main branch diameter; CEN = internode length.

**Table 7.** Groups formed by Ward-MLM for the twenty characterized variables from the first evaluation period in backcross hybrid progenies of *Passiflora*.

Group	Genitor and BC, genotype
I	*HD17-101, HD17-102, HD17-108, HD17-109, HD17-113, HD17-114, HD17-120, HD17-130, HD17-135, HD17-138, HD17-139, HD17-140, HD17-142, HD17-143, HD17-145, HD17-146, HD17-149, HD17-154, HD17-163, HD17-169, HD17-170, HD17-175, HD18-101, HD18-114, HD18-115, HD18-117, HD18-119, HD18-123, HD18-124, HD18-126, HD18-128, HD18-130, HD18-134, HD18-138, HD18-140, HD18-141, HD18-142, HD18-144, HD18-149, HD18-170, HD18-175, HD18-177, HD18-173, HD19-107, HD19-108, HD19-109, HD19-110, HD19-115, HD19-119, HD19-121, HD19-123, HD19-129, HD19-131, HD19-132, HD19-133, HD19-134, HD19-137, HD19-138, HD19-139, HD19-142, HD19-158, HD19-159, HD19-164, HD19-185, HD19-187, HD20-101, HD20-108, HD20-115, HD20-117, HD20-119, HD20-122, HD20-123, HD20-127, HD20-130, HD20-131, HD20-131, HD20-135, HD20-139, HD20-143, HD20-145, HD20-147, HD20-152, HD20-153, HD20-155, HD20-159, HD20-164, HD20-166, HD20-169, HD20-172, HD20-173, HD20-180, HD20-181, HD20-182, HD20-183, HD20-184, HD20-185, HD20-189, HD20-190.
II	<i>Passiflora foetida</i> , HD13-101, HD13-133, HD13-134, HD13-143, HD17-106, HD17-118, HD17-122, HD17-123, HD17-126, HD17-132, HD17-133, HD17-134, HD17-144, HD17-149, HD17-155, HD17-156, HD17-157, HD17-158, HD17-159, HD17-160, HD17-161, HD17-162, HD17-164, HD17-166, HD17-167, HD17-168, HD17-171, HD17-173, HD17-176, HD17-177, HD17-181, HD17-185, HD17-189, HD17-190, HD18-166, HD18-167, HD18-105, HD19-149, HD19-151, HD19-152, HD20-116, HD20-150, HD20-156.
III	HD17-103, HD17-104, HD17-112, HD17-115, HD17-117, HD17-125, HD17-127, HD17-136, HD17-141, HD17-153, HD17-174, HD17-178, HD17-180, HD17-182, HD17-188, HD18-102, HD18-103, HD18-104, HD18-106, HD18-108, HD18-109, HD18-110, HD18-111, HD18-112, HD18-113, HD18-118, HD18-120, HD18-121, HD18-122, HD18-127, HD18-133, HD18-135, HD18-136, HD18-139, HD18-143, HD18-145, HD18-146, HD18-147, HD18-148, HD18-149, HD18-151, HD18-152, HD18-153, HD18-154, HD18-158, HD18-161, HD18-162, HD18-165, HD18-168, HD18-169, HD18-172, HD18-173, HD18-174, HD18-178, HD18-181, HD18-182, HD18-188, HD18-189, HD19-101, HD19-102, HD19-105, HD19-106, HD19-111, HD19-112, HD19-113, HD19-114, HD19-118, HD19-122, HD19-135, HD19-136, HD19-140, HD19-143, HD19-144, HD19-150, HD19-153, HD19-157, HD19-160, HD19-161, HD19-163, HD19-168, HD19-172, HD19-173, HD19-177, HD19-178, HD19-179, HD19-180, HD19-181, HD19-183, HD19-186, HD19-189, HD20-102, HD20-103, HD20-104, HD20-105, HD20-106, HD20-107, HD20-109, HD20-110, HD20-111, HD20-112, HD20-113, HD20-114, HD20-120, HD20-124, HD20-125, HD20-128, HD20-129, HD20-132, HD20-133, HD20-136, HD20-137, HD20-138, HD20-140, HD20-141, HD20-142, HD20-144, HD20-146, HD20-148, HD20-151, HD20-154, HD20-157, HD20-158, HD20-160, HD20-163, HD20-165, HD20-168, HD20-170, HD20-175, HD20-176, HD20-177, HD20-178, HD20-186, HD20-187, HD20-188.
IV	<i>P. subanceolata</i> , HD17-107, HD17-110, HD17-111, HD17-119, HD17-124, HD17-128, HD17-129, HD17-147, HD17-148, HD17-151, HD17-152, HD17-165, HD17-172, HD17-183, HD17-184, HD17-186, HD18-107, HD18-132, HD18-155, HD18-159, HD18-160, HD18-163, HD18-164, HD18-176, HD18-179, HD18-180, HD18-185, HD18-186, HD18-189, HD19-117, HD19-124, HD19-125, HD19-126, HD19-127, HD19-128, HD19-130, HD19-146, HD19-147, HD19-154, HD19-155, HD19-162, HD19-165, HD19-166, HD19-174, HD19-176, HD19-182, HD19-184, HD19-188, HD20-118, HD20-121.

\*HD designation refers to the backcross progeny number followed by the identifier number of the plant within the progeny.

**Table 8.** Distance between the groups formed by Ward-MLM in the first trial period in backcross hybrid progenies of *Passiflora*.

Group	I	II	III	IV
I	0	41.7918	34.7626	34.0211
II	0	0	2.1668	3.7221
III	0	0	0	1.0433

In the second assessment period, group II, represented by recurrent parents and some BC<sub>1</sub> genotypes, presented a greater distance from the other groups (I, III, and IV). The distance was greater between groups II and III (7.78) and lowest between groups III and IV (0.83) (Table 9). It was observed that Group I of the first assessment period and Group II of the second assessment period (which were largely composed of recurrent parents) had greater genetic dissimilarity to the parental donor *P. subanceolata* and the BC<sub>1</sub> progenies.

**Table 9.** Distance between the groups formed by Ward-MLM in the second trial period in backcross hybrid progenies of *Passiflora*.

Group	I	II	III	IV
I	0	7.59108	0.23661	0.95039
II	0	0	7.78512	9.82152
III	0	0	0	0.83979

The inclusion of qualitative variables such as the presence of floral PB and floral part coloration was crucial in the multivariate analysis by Ward-MLM for a greater accuracy in the formation and distribution of genotypes among the four groups during both assessment periods. Clustering error was found in preliminary data analysis that did not include qualitative data, revealing a large group number with few differences in increments between them based on the pseudo-F and pseudo-T<sup>2</sup> criteria (data not shown).

In the first assessment period, group I consisted of recurrent parents (HD13) and *P. foetida* and presented few variation for the qualitative characters. The other groups in the same trial period contained genotypes with wide variations in floral colors as is evident in the description of COLP and COLS in Table 10.

Approximately 64.4% of genotypes had colored petals and 71.1% had colored sepals in group II in the first trial period. The colors were predominantly classified as medium-magenta (5RP 4/12). On the other hand, more than 69% of the genotypes in group III showed a predominant violet flower color (5RP 6/8). A high percentage of genotypes in group IV in the first period (58.4%) had predominantly dark violet colored sepals (5RP 5/10). Only 7.7% of genotypes of groups II, III, and IV from the first trial period were identified as having partial floral band patterning of the petals and sepals. Qualitative polymorphisms were not observed for the predominant color of the second row of coronal filaments; a similar conclusion was reached for samples without coloring (white). Overall, in the second trial period, a broad gradient of coloring was observed for petals and sepals among all groups (Table 11).

Most of the genotypes in group II (62.5%) had predominantly white flowers. However, group III was made up of 52% of genotypes with predominantly dark violet (5RP 5/10) petals and 88.8% with colored sepals.

The first two canonical variables obtained by the Ward-MLM procedure were sufficient to

**Table 10.** Absolute frequency of the four qualitative variables in each of the four groups formed by Ward-MLM in the first assessment period in backcross *Passiflora* hybrid progenies.

Trait	Group			
	I (25)	II (42)	III (136)	IV (149)
Flower banding				
Absent	20	35	129	141
Presence	5	7	7	8
Petal coloration				
(3)5RP 3/8	-	-	3	-
(4) 5RP 4/12	-	29	-	9
(5) 5RP 5/10	5	1	34	40
(6) 5RP 6/8	-	4	94	24
(7) 5RP 8/6	5	5	-	9
(8) White	15	3	5	67
Sepal coloration				
(3) 5RP 3/8	-	-	3	-
(4) 5RP 4/12	-	30	1	9
(5) 5RP 5/10	10	7	28	87
(6) 5RP 6/8	-	1	90	5
(7) 5RP 8/6	5	2	-	1
(8) White	10	2	14	47
First corona filament coloration				
(3) 5RP 3/8	-	17	70	27
(5) 5RP 5/10	5	-	-	-
(6) 5RP 6/8	15	-	-	-
(8) White	5	25	66	122

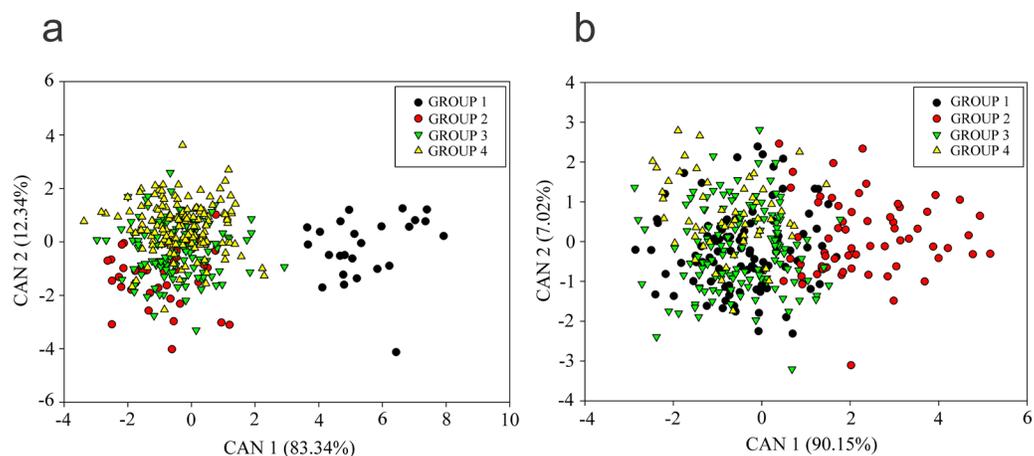
Color intensities and hues were classified using Munsell's color code (1977): (3/8 5RP) dark magenta; (5RP 4/12) medium-magenta; (5/10 5RP) dark violet; (6/8 5RP) medium violet; (8/6 5RP) light violet.

**Table 11.** Absolute frequency of the four qualitative variables in each of the four groups formed by Ward-MLM in the second assessment period in backcross hybrid progenies of *Passiflora*.

Trait	Group			
	I (97)	II (64)	III (134)	IV (55)
Flower banding				
Absent	91	54	125	52
Presence	6	10	9	3
Petal coloration				
(3) 5RP 3/8	3	-	-	-
(4) 5RP 4/12	-	1	-	36
(5) 5RP 5/10	3	5	70	1
(6) 5RP 6/8	85	-	21	3
(7) 5RP 8/6	1	11	22	-
(8) White	5	47	21	15
Sepal coloration				
(3) 5RP 3/8	3	-	-	-
(4) 5RP 4/12	1	1	-	37
(5) 5RP 5/10	2	11	119	-
(6) 5RP 6/8	84	1	-	4
(7) 5RP 8/6	2	11	1	1
(8) White	5	40	14	13
First corona filament coloration				
(3) 5RP 3/8	38	12	32	26
(5) 5RP 5/10	-	5	-	-
(6) 5RP 6/8	-	15	-	-
(8) White	59	32	102	29

Color intensities and hues were classified using Munsell's color code (1977): (3/8 5RP) dark magenta; (5RP 4/12) medium-magenta; (5/10 5RP) dark violet; (6/8 5RP) medium violet; (8/6 5RP) light violet.

explain 95.68 and 97.17% of the total variation, respectively, for the first and second assessment periods. Thus, the visual representation of the genetic variation in the two-dimensional graphics in Figure 2 provides an accurate description of the relationship between the groups.



**Figure 2.** First two canonical variables (CAN) for the four groups formed by analyzing Ward-MLM in backcrossed hybrid progenies of *Passiflora*. **a.** Groups formed in the first trial period, where group I contains recurring HD13 and given by the HD13 male parent hybrids, *P. foetida*; group II contains donor parent, *P. sublancoolata*, and BC<sub>1</sub> plants, groups III and IV have only BC<sub>1</sub> plants. **b.** Groups formed in the second trial period, in which group I contains BC<sub>1</sub> plants; group II contains recurrent parent HD13, the male parent hybrids HD13, *P. foetida*, and BC<sub>1</sub> plants; group III have BC<sub>1</sub> plants; group IV contains the *P. sublancoolata* parent donor, and BC<sub>1</sub> plants.

In the two-dimensional graph obtained by analysis of the data from the first trial, it was apparent that the efficiency of the backcross method in which the BC<sub>1</sub> genotypes were contained in groups II, III, and IV were similar to those of the parental donors of *P. sublancoolata*, but showed greater distance from group I which had the recurrent parents (HD13) and the progenitor *P. foetida* (Figure 2a). The characters that contributed most to the genetic divergence observed in the first assessment period were, hierarchically, DCO, FD, CSE, and CPE, with values of 0.70246, -0.69914, -0.66318, and -0.64048, respectively. In the second evaluation period, greater genetic divergence between the genotypes was observed in the two-dimensional graphs. However, despite not being isolated from group II (recurrent parents and *P. foetida*), the majority of genotypes were distributed far from the other groups (Figure 2b). The DCO descriptors, TF1, LF, and CSE contributed most to the groups generated in the second trial period with 0.82197, 0.70347, -0.44139, -0.38440, respectively.

## DISCUSSION

The simultaneous analysis of 16 quantitative and 4 qualitative variables using a multivariate approach with Ward-MLM methodology was highly efficient in differentiating plants within BC<sub>1</sub> progenies and their recurrent parents (HD13) in two assessment periods. This combined approach has similarly been exploited to quantify genetic variability in other plant species, such as *P. vulgaris* L. (Cabral et al., 2010), *Solanum lycopersicum* (Gonçalves et al., 2009), *Musa* sp (Pestana et al., 2011; Pereira et al., 2012), *Jatropha curcas* (Brasileiro et al., 2013), and *Passiflora* (Santos, 2013).

In this study, four groups were defined in both assessment periods using pseudo-F and pseudo-T<sup>2</sup> criteria and by analysis of the verisimilitude function. Such a strategy ideally defines groups without subjectivity and eliminates the personal establishment of group numbers formed in intra- and inter-population analyses (Campos et al., 2013; Santos, 2013). The optimal number of defined groups varies with factors such as species, cultivar, number of accessions, and the amount and type of descriptors evaluated (Gonçalves et al., 2009).

We found greater dissimilarity between recurrent parents and BC<sub>1</sub> genotypes in both evaluation periods, and also demonstrated that the characters examined were efficient for the assessment of genetic diversity in the BC<sub>1</sub> plants and their respective parents. In a previous study of the HD13 progeny and the parental species, *P. sublancoolata* and *P. foetida*, used both scatter charts and the dendrograms generated by the UPGMA algorithm. This study showed the separation between the parents and the F<sub>1</sub> population and noted greater genetic similarity between male *P. foetida* parents and the segregated progeny (Santos et al., 2011).

Low dissimilarity between groups I and II, the segregated hybrid population, and the genitor *P. setacea* was established by the Ward-MLM approach, indicating similar values for most of the characteristics evaluated between the two groups (Santos, 2013). In the present study, the BC<sub>1</sub> population showed higher genetic dissimilarity with the recurrent parent and the male progenitor *P. foetida*, and showed greater similarity with the *P. sublancoolata* donor genitor. This conclusion was verified by the averages of the FD, DCO, CSE, and TF1 descriptor values among groups II, III, and IV from the first assessment period and among groups I, III, and IV from the second. In one HD13 hybrid population with *P. sublancoolata* and *P. foetida*, the FD and DCO variables contributed greatly to the formation of the groups (Santos et al., 2011). In hybrids between *P. setacea* and *P. edulis*, the lengths of the androgynophorous and floral peduncle gave a greater contribution within floral traits for genetic variation in interspecific F<sub>1</sub> hybrids and their female genitors (Santos, 2013).

Generally, the accumulation of progeny from the same parents within a specific group was not observed. Nevertheless, HD20 progeny plants were present in group II and were analyzed using the data obtained in the first trial period. The difference between the HD20 progeny with a *P. sublancoolata* parent donor may have contributed to the exclusion of representatives from progeny in which the genitor donor was found.

The choice of the 2D graph was supported by the high percentages obtained by the first two canonical variables, which accounted for much of the variation within and between groups defined by the pseudo-F and pseudo-T<sup>2</sup> criteria. Two-dimensional graphs have been widely used in studies employing the Ward-MLM procedure as the first two canonical variables and are able to explain much of the genetic variation between groups (Gonçalves et al., 2009; Oliveira et al., 2013; Santos, 2013). However, if the first two canonical variables do not explain a large part of the variation, then inclusion of the third canonical variable may be required and result in a three-dimensional representation of the data set. For example, Campos et al. (2013) reported that the first two canonical variables accounted for only 61.79% of the variation in 138 guava genotypes (*Psidium guajava* L.). In this case, a third canonical variable included and raised accountability to 81.30% of the total variation and enabled the representation of the data in a three-dimensional graph. In the present study, the two-dimensional graph formed from the second assessment period showed higher dispersion among the groups/genotypes, indicating that environmental differences contributed to the dispersion of the genotypes, but did not considerably influence the number of groups or the distribution of genotypes within them.

The few observed differences between average values of the *P. sublancoolata* donor

parent groups and backcross genotypes was consistent with the low dissimilarity between these groups; this indicated success in the recovery of quantitative characteristics by the backcrossing method. Thus, selection of any progeny based on quantitative parameters represented an increase in the average values of the main features for floral ornamentation, such as FD, CPE, and CSE in relation to the HD13 recurrent parent. However, for the selection of plants for ornamental flowers, the joint use of the quantitative and qualitative characteristics is recommended. In this way, one can select BC<sub>1</sub> genotypes from groups II and IV of the first assessment period and from groups I and IV of the second characterization period. These groups, in addition to presenting plants with high FD values, also had genotypes with large coloring gradients in their petals and sepals; these are important characteristics for the selection of plants aimed at ornamental breeding. The Ward-MLM strategy has been used to identify sour passion fruit plants with superior agronomic characteristics to their parents; thus, in crossbreeding, it is possible to predict effective gains in productivity using the increase in total soluble solid content and the decrease in the thickness of the shell (Santos, 2013).

The Ward-MLM analysis showed a low level of dissimilarity between the *P. sublancoolata* genitor donor and BC<sub>1</sub> hybrids; one can predict that little change in average values for floral descriptors would be seen after a second generation of backcrossing. In this case, it would be advisable to obtain RC<sub>2</sub> progenies, with a view to reducing heterosis and increasing the similarities between the quantitative and qualitative descriptor values of the likely RC<sub>2</sub> progeny within the *P. sublancoolata* species. Further backcrossed generations using *P. sublancoolata* as a donor genitor should result in progenies with low genetic diversity and a strong resemblance to the average values of the quantitative descriptors. Therefore, these plants present a range of backcrossed genotypes with potential for development of new ornamental varieties, or they can be subjected to other methods of genetic improvement, such as hybridization with a source of genetic variation for increasing the length of the flowering period.

### Conflicts of interest

The authors declare no conflict of interest.

### ACKNOWLEDGMENTS

Research supported by UESC (Universidade Estadual de Santa Cruz), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and FAPESB (Fundação de Amparo à Pesquisa do Estado da Bahia). We would like to thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for the scholarships granted to the first author and CNPq for the scholarship awarded to the second author.

### REFERENCES

- Abreu PP, Souza MM, Santos EA, Pires MV, et al. (2009). Passion flower hybrids and their use in the ornamental plant market: perspectives for sustainable development with emphasis on Brazil. *Euphytica* 166: 307-315.
- Belo GO (2010). Análise morfológica e genética em progênie híbrida F1 do cruzamento *Passiflora gardneri* MAST x *Passiflora gibertii* N.E. BROW. Master's thesis, UESC, Ilhéus.
- Borém A and Miranda GV (2005). *Melhoramento de Plantas*. 4th edn. UFV, Viçosa.
- Bueno LCS, Mendes ANG and Carvalho SP (2006). *Melhoramento de plantas: princípios e procedimentos*. UFLA, Lavras.

- Brasileiro BP, Silva SA, Souza DR, Santos PA, et al. (2013). Genetic diversity and selection gain in the physic nut (*Jatropha curcas*). *Genet. Mol. Res.* 12: 2341-2350.
- Cabral PDS, Soares TCB, Gonçalves LSA, Amaral Júnior AT, et al. (2010). Quantification of the diversity among common bean accessions using Ward-MLM strategy. *Pesq. Agropec. Bras.* 45: 1124-1132.
- 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. Frut.* 35: 571-578.
- Crossa J and Franco J (2004). Statistical methods for classifying genotypes. *Euphytica* 137: 19-37.
- Franco J, Crossa J, Villasenõr J, Taba S, et al. (1998). Classifying genetic resources by categorical and continuous variables. *Crop Sci.* 38: 1688-1696.
- Gonçalves LSA, Rodrigues R, Amaral Júnior AT and Karasawa M (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
- Gower JC (1971). A general coefficient of similarity and some of its properties. *Biometrics* 27: 857-874.
- Koinange EMK and Gepts P (1992). Hybrid weakness in wild *Phaseolus vulgaris* L. *J. Hered.* 83: 135-139.
- Lippman ZB and Zamir D (2007). Heterosis: revisiting the magic. *Trends Genet.* 23: 60-66.
- MAPA (2008). Instruções para execução dos ensaios de distinguibilidade, homogeneidade e estabilidade de cultivares de *Passiflora*. Ministério da Agricultura, Pecuária e Abastecimento. Available at [http://www.agricultura.gov.br].
- Munsell AH (1977). Munsell color chart for plant tissue. Munsell Color Company, Baltimore.
- Oliveira RS, Silva SA, Brasileiro BP, Medeiros EP, et al. (2013). Genetic divergence on castor bean using the Ward-MLM strategy. *Rev. Cienc. Agron.* 44: 564-570.
- 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. *Pesq. Agropec. Bras.* 47: 1480-1488.
- Pestanana RKN, Amorim EP, Ferreira CF, Amorim VBO, et al. (2011). Agronomic and molecular characterization of gamma ray induced banana (*Musa* sp.) mutants using a multivariate statistical algorithm. *Euphytica* 178: 151-158.
- Sanghera GS, Wani SH, Hussain W, Shafi W, et al. (2011) The magic of heterosis: new tools and complexities. *Nat. Sci.* 9: 43-53.
- Santos EA (2013). Melhoramento do maracujazeiro-azedo (*Passiflora edulis* Sims) visando à resistência ao *Cowpea aphid-borne mosaic virus*. Doctoral thesis, UENF, Campos dos Goytacazes.
- Santos EA, Souza MM, Viana AP, Almeida AAF, et al. (2011). Multivariate analysis of morphological characteristics of two species of passion flower with ornamental potential and of hybrids between them. *Genet. Mol. Res.* 10: 2457-2471.
- Santos EA, Souza MM, Abreu PP, Conceição LDHCS, et al. (2012). Confirmation and characterization of interespecific hybrids of *Passiflora* L. (Passifloraceae) for ornamental use. *Euphytica* 184: 389-399.
- Sobrizal S and Yoshimura A (2009). Hybrid weakness found in the backcross progeny of an inter-specific cross of *Oryza sativa*/O. *glumaepatula*. *Atom Indonesia* 35: 49-56.
- Schifino-Wittmann MT and Dall'Agnol M (2002). Auto-incompatibilidade em plantas. *Cienc. Rural* 32: 1083-1090.
- Sudré CP, Gonçalves LSA, Rodrigues R, Amaral Júnior AT, et al. (2010). Genetic variability in domesticated *Capsicum* spp as assessed by morphological and agronomic data in mixed statistical analysis. *Genet. Mol. Res.* 9: 283-294.
- Ulmer T and MacDougal JM (2004). *Passiflora* - Passion flowers of the world. Timber Press, Oregon.
- Veilleux RE and Miller AR (1998). Hybrid breakdown in the F<sub>1</sub>, between *Solanum chacoense* and *S. phureja* and gene transfer for leptine biosynthesis. *J. Am. Soc. Hortic. Sci.* 123: 854-858.
- Ward JH (1963). Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58: 236-244.