Molecular and agronomic analysis of intraspecific variability in *Capsicum baccatum* var. *pendulum* accessions

P.S.S. Leite[1], R. Rodrigues[1], R.N.O. Silva[3], S. Pimenta[1], A.M. Medeiros[1], C.S. Bento[1] and L.S.A. Gonçalves[2]

[1]Departamento de Melhoramento de Plantas, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, RJ, Brasil
[2]Departamento de Agronomia, Universidade Estadual de Londrina, Londrina, PR, Brasil
[3]Universidade Federal do Piauí, Campus Amílcar Ferreira Sobral, Floriano, PI, Brasil

Corresponding author: R. Rodrigues
E-mail: rosana@uenf.br

ABSTRACT. *Capsicum baccatum* is one of the most important chili peppers in South America, since this region is considered to be the center of origin and diversity of this species. In Brazil, *C. baccatum* has been widely explored by family farmers and there are different local names for each fruit phenotype, such as *cambuci* and *dedo-de-moça* (lady's finger). Although very popular among farmers and consumers, *C. baccatum* has been less extensively studied than other *Capsicum* species. This study describes the phenotypic and genotypic variability in *C. baccatum* var. *pendulum* accessions. Twenty-nine

accessions from the Universidade Estadual do Norte Fluminense Darcy Ribeiro gene bank, and one commercial genotype ('BRS-Mari') were evaluated for 53 morphoagronomic descriptors (31 qualitative and 22 quantitative traits). In addition, accessions were genotyped using 30 microsatellite primers. Three accessions from the *C. annuum* complex were included in the molecular characterization. Nine of 31 qualitative descriptors were monomorphic, while all quantitative descriptors were highly significant different between accessions (P < 0.01). Using the unweighted pair group method using arithmetic averages, four groups were obtained based on multicategoric variables and five groups were obtained based on quantitative variables. In the genotyping analysis, 12 polymorphic simple sequence repeat primers amplified in *C. baccatum* with dissimilarity between accessions ranging from 0.13 to 0.91, permitting the formation of two distinct groups for Bayesian analysis. These results indicate wide variability among the accessions comparing phenotypic and genotypic data and revealed distinct patterns of dissimilarity between matrices, indicating that both steps are valuable for the characterization of *C. baccatum* var. *pendulum* accessions.

Key words: Chili peppers; Phenotyping; Microsatellite markers; Bayesian analysis

INTRODUCTION

Capsicum baccatum is an important genetic resource in South American countries, since this species is thought to have originated from this region (Eshbaugh, 1970; D'Arcy and Eshbaugh, 1974). In 2012, the world production of chili, including *C. baccatum*, was approximately 34 million tons (FAO, 2012) not including the informal market, which is significant in developing countries. *C. baccatum* var. *pendulum* is mainly cultivated at medium altitudes on the plains of Argentina, Bolivia, Peru, Ecuador, Paraguay, Colombia, and Chile, and in different environments in Brazil (Albrecht et al., 2012), from rain forests to savannah areas. In Brazil, chili peppers belonging to this species are widely produced by family farmers, and the types *cambuci* and *dedo-de-moça* (finger's lady) are commonly grown for fresh consumption or for processing into powder or extracts (Rêgo et al., 2011; Rodrigues et al., 2012; Moulin et al., 2015). The conservation of local varieties of *C. baccatum* in germplasm banks is essential to preserve the diversity of this species and to provide researchers with a range of genetic resources that can be applied in breeding programs. However, accessions must be properly characterized and evaluated if they are to be conserved in gene banks and be made available for use in breeding programs (Gonçalves et al., 2009; Ibiza et al., 2012; Nicolaï et al., 2013).

Gene banks and collections of *Capsicum* germplasm predominantly retain domesticated accessions, with emphasis on complex *C. annuum* species. This is also observed for the characterization and evaluation of germplasm. In studies involving multiple species, 89% of the accessions belong to the *C. annuum* complex, 9.5% correspond to *C. baccatum*, 0.61% to *C. pubescens*, and 0.78% to other species. An exception is the study of Ibiza et al. (2012) on which the species in larger quantities is *C. baccatum* (84 accessions) and a lower number is *C. annuum* (25 accessions). It is likely that this occurred because the accessions

were of Andean origin, which is a region less favorable for the development of *C. annuum*.

Capsicum germplasm has been characterized using phenotypic and molecular markers (Sudré et al., 2010). Molecular markers, especially those based on polymerase chain reaction (PCR), have been widely used due to the minimal influence of environmental conditions or plant development factors (González-Pérez et al., 2014). Several molecular markers have been used to investigate genetic diversity in *Capsicum* spp, including random amplified polymorphic DNA (Costa et al., 2006), amplified fragment length polymorphism (Ibiza et al., 2012; Albrecht et al., 2012; Moses et al., 2014; Baba et al., 2015), and simple sequence repeats (SSR) or microsatellites (Nicolaï et al., 2013; Villela et al., 2014). SSR markers have many advantages over other markers, including their reproducibility, multi-allelic nature, co-dominant inheritance, relative abundance, and wide genome coverage (Varshney et al., 2005). However, phenotypic descriptors are essential during germplasm characterization because they provide important morphological and agronomic information for gene bank conservation or for plant-breeding programs.

A study published by Albrecht et al. (2012) involving 226 accessions of *C. baccatum* from different countries concluded that collecting wild and domesticated germplasm in some geographic regions, including Brazilian areas, must be focused to elucidate issues regarding origin and domestication of *C. baccatum*. In this study, we describe the quantification of intraspecific variability in *C. baccatum* var. *pendulum* of the germplasm collection from Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF) in Brazil using morphoagronomic descriptors and molecular markers.

MATERIAL AND METHODS

Plant material

Thirty genotypes of *C. baccatum* were evaluated, 29 of which are gene bank accessions of from the UENF, and one of which is a commercial genotype (BRS-Mari). The accessions used in this experiment were collected in three Brazilian regions (South, Southeast and Midwest), comprising five Brazilian States (Rio de Janeiro, Minas Gerais, Espírito Santo, Mato Grosso do Sul, and Paraná). The greatest distance between sampling sites in Brazil was 1914 km (Celina, Espírito Santo State and Miranda, Mato Grosso do Sul State) and one accession was donated from Peru. Most of the accessions have been maintained for more than 20 years in the germplasm bank of UENF.

Morphoagronomic characterization

The accessions were planted in polystyrene trays with the commercial substrate Vivatto. After 30 days, plants were individually transplanted to plastic 5-L pots containing a mixture of soil and substrate (2:1 ratio). Plants were grown in a greenhouse under the same light, temperature, and crop management conditions at UENF, Campos dos Goytacazes, Rio de Janeiro, Brazil. The experimental design was completely randomized with five replicates. Fertilizing and other cultural practices were carried out according to standard crop management recommendations (Filgueira, 2012).

C. baccatum var. *pendulum* accessions were characterized using 53 descriptors established by Biodiversity International, 31 of which were qualitative and 22 of which were quantitative. Each descriptor was evaluated at specific stages of plant development according

to the rules of the 'Descriptors for *Capsicum*' (IPGRI, 1995). The qualitative descriptors evaluated were: stem color (SC); stem pubescence (SP); stem shape (SS); nodal anthocyanin (NA); leaf color (LC); leaf shape (LS); leaf density (LD); branching habit (BH); plant growth habit (PGH); corolla color (CC); anther color (AC); filament color (FC); corolla spot color (CSC); number of flowers per axil (NFA); flower position (FP); stigma exsertion (SE); calyx pigmentation (CP); corolla shape (CS); calyx margin (CM); anthocyanin spots or stripes (AS); ripe fruit persistence/pedicel with fruit (RFP); number of locules (NL); neck at base of fruit (NBF); fruit shape at blossom end (FSBE); fruit blossom-end appendage (FBEA); fruit cross-sectional corrugation (FCC); fruit shape at pedicel attachment (FSH); fruit surface (FS); fruit color at intermediate stage (FCIS); fruit color at mature stage (FCMS); and fruit shape (FSE).

In addition, 22 quantitative descriptors were used: plant height (PH); plant canopy width (PCW); stem diameter (SD); leaf length (LL); leaf width (LW); petiole length (PL); corolla length (CL); width of the petal (WP); anther length (AL); filament length (FIL); fruit length (FL); fruit width (FW); fruit pedicel length (FPL); fruit wall thickness (FWT); number of fruits per plant (NFP); days to flowering (DFL); days to fruiting (DFR); fresh fruit mass (FFM); dry fruit mass (DFM); fruit dry matter content (%; FDMC); 100-seed mass (HSM); and soluble solids content (°Brix; SSC).

Molecular characterization

For molecular characterization, three complex *C. annuum* accessions (UENF 1491, UENF 1719, and UENF 1795, identified as *C. frutescens*, *C. chinense*, and *C. annuum*, respectively) were added as outgroups.

Leaves were sampled at the same stage of development. All plant material sampled was placed into envelopes, identified, immersed in liquid N_2, and transported for storage in an ultrafreezer at -82°C. Subsequently, 300 mg macerated tissue in liquid N_2 medium was transferred to microtubes (2.0 mL) and properly identified.

DNA extraction was carried out according to the procedure described by Doyle and Doyle (1987), with modifications as suggested by Daher et al. (2002). After extraction, DNA integrity was checked on 1.0% agarose gels using a high molecular weight DNA ladder (Invitrogen, USA). To measure the volume of extracted DNA, the samples were subjected to quantification in a NanoDrop 2000c, at a 260-nm absorbance. Based on the result, samples were diluted to working concentrations of 5 ng/µL.

A total of 30 microsatellite primers developed and optimized for mapping *C. annuum* were tested (Minamiyama et al., 2006). Each reaction contained the following: 1X PCR buffer (500 mM KCl, 100 mM Tris-HCl, pH 8.8), 0.2 mM dNTPs, 2.0 mM $MgCl_2$, 0.4 µM each primer, 0.6 U Taq polymerase, 10 ng genomic DNA. Samples were made up to a final volume of 13 µL using ultrapure water. PCR was performed in a thermocycler (model Veriti 384; Biosystems, USA) as follows: 4 min at 94°C followed by 35 cycles [94°C for 1 min, 55°-63°C for 1 min (depending on the primer), and 72°C for 3 min], and a final extension at 72°C for 7 min.

The amplified fragments were separated by high-resolution agarose gel 4% Metaphor and stained with a mixture of 6X Blue Juice (10X 0.4 mL TAE, 0.5 M 0.2 mL 10% SDS, 0.2 mL bromophenol blue, 7.0 mL glycerol, 1.7 mL sterile water) with 5X GelRed (10,000X 1 µL GelRed, 0.5 mL DMSO, 2 mL ultrapure water) at a ratio of 1:1. To estimate the size of the fragments, a 100-bp DNA Ladder (Invitrogen) was used. The gels were photographed using a Minibis Pro (Bio Imaging System, USA).

Statistical analysis

Quantitative morphological data were submitted to analysis of variance and Pearson's correlation. To study the genetic diversity among accessions, quantitative and qualitative morphological data were used to determine the generalized Mahalanobis distances and coincidence, respectively, and were then grouped by hierarchical unweighted pair group method using arithmetic averages (UPGMA). The relative importance of quantitative variables was analyzed using the method described by Singh (1981). For molecular data, we calculated the observed and expected heterozygosities (H_O and H_E, respectively) and polymorphic information content (PIC). Distance was measured using the shared allele and UPGMA methods (10,000 bootstrap resamplings). Analysis of variance, correlation, relative importance, and genetic distances were determined using the Genes software (Cruz, 2013), while groups, bootstrap (for molecular data), and graphics were obtained in the R software (http://www.r-project.org) using the Agricolae, Dendextend, and Corrplot packages.

To verify possible clustering between accessions and the molecular data, a Bayesian analysis was performed using the Structure 2.3.4 software (Pritchard et al., 2000), according to the method described by Evanno et al. (2005), with a burn-in of 10,000 repetitions and Markov Chain Monte Carlo (MCMC) simulations of 100,000 repetitions. K values ranging from 1 to 32 were tested, with 17 independent interactions for each grouping. Probable K clusters were determined using Structure Harvester (Earl and vonHoldt 2012).

RESULTS

Morphoagronomic characterization

Among the 31 qualitative morphological descriptors investigated, 22 were polymorphic and nine were monomorphic (SC, CP, SS, LC, CC, NFA, AC, FC, and CS) (Table 1).

Most of the accessions (90%) had stems with sparse pubescence, and the remaining 10% was classified as intermediate, according to diagram proposed by IPGRI (1995). Anthocyanin was absent in the nodes of 66.7% of accessions. The evaluated accessions possessed different leaf shapes, with 60.0% being oval, 23.3% lanceolate, and 16.7% deltoid. Leaf density ranged from dense (93.3%) to intermediate (6.7%), while branch density was classified as dense, intermediate, and scarce in 66.7, 30.0 and 3.3% accessions, respectively. Most accessions had pending flowers at an intermediate position (96.3%) with yellowish green corolla stains (93.3%), and the remaining accessions had an erect habit with green corolla spots. Stigma and anther were located on the same level in 86.7% of the accessions.

For most accessions, the traits dentate calyx margin (76.7%), orange fruits at the intermediate stage (86.7%), and red color at ripeness (90%) were observed. The accessions UENF 1573, UENF 1633, and UENF 1738 were the only ones to display anthocyanin spots.

Among the evaluated fruit, almost all formats proposed by IPGRI on a list of descriptors for *Capsicum* were observed: 50% of accessions produced elongated fruits, 26.7% produced triangular, 13.3% produced campanulate, 6.7% produced nearly round, and 3.3% produced bell-shaped (campanulate with indentations) fruit. Regarding the number of locules, fruit with two, three, and four locules at 40, 56.7, and 3.3% were observed among the evaluated accessions, respectively. An intermediate persistence of fruit and pedicel was also noted (Figure 1).

Table 1. Mean squares, coefficients of variation (CV), and average values for 22 morphoagronomic variables in 30 *Capsicum baccatum* var. *pendulum* genotypes (Campos dos Goytacazes, Brazil).

| Variables	Mean squares		CV (%)	Average
	Genotypes	Error		
PH	1591.43**	59.24	9.29	82.8
PCW	485.56**	28.26	6.43	82.7
SD	8.60**	3.43	11.39	16.2
LL	59.43**	2.52	13.51	11.8
LW	3.01**	0.25	11.82	4.2
PL	4.50**	0.19	13.76	3.2
CL	0.34**	0.38	9.61	2.0
WP	0.014**	0.002	9.19	0.5
AL	0.0015**	0.0004	9.09	0.2
FIL	0.0043**	0.0009	8.02	0.4
DFL	213.35**	12.25	10.91	32.1
DFR	177.47**	21.48	5.78	80.1
FL	3821.99**	13.68	6.60	56.0
FW	556.27**	4.82	8.50	25.8
FPL	709.93**	14.23	7.57	49.8
FWT	1.49**	0.055	9.70	2.42
NFP	741.15**	35.62	28.47	21.0
FFM	63.54**	2.87	23.82	7.1
DFM	212.76**	10.94	24.35	13.6
FDMC	26.84**	2.37	8.53	18.0
HSM	0.038**	0.002	8.66	0.5
SSC	10.35**	0.58	8.89	8.6

PH: plant height (cm), PCW: plant canopy width (cm), SD: stem diameter (cm), LL: leaf length (cm), LW: leaf width (cm), PL: petiole length (cm), CL: corolla length (mm), WP: width of the petal (mm), AL: anther length (mm), FIL: filament length (mm), DFL: days to flowering, DFR: days to fruiting, FL: fruit length (mm), FW: fruit width (mm), FPL: fruit pedicel length (mm), FWT: fruit wall thickness (mm), NFP: number of fruits per plant, FFM: fresh fruit mass in gram (g), DFM: dry fruit mass in gram (g), FDMC: fruit dry matter content (%), HSM: 100-seed mass (g), and SSC: soluble solids content (ºBrix).

Figure 1. Phenotypic variability of fruit from *Capsicum baccatum* var. *pendulum* accessions deposited in the germplasm bank of Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes, Brazil.

Among the evaluated accessions, 46.7% had truncate fruit shape at pedicel attachment, 43.3% had obtuse fruit, and 10% had heart-shaped. Only the accessions UENF 1628 and UENF 1629 had a 'neck' at the base of the fruit, and most were pointed with the apex at the fruit tip. Cross-sectional analysis of the fruit revealed that 53.3% had an intermediate corrugation, 40.0% were slightly corrugated, and 6.7% were corrugated. In addition, the majority of fruit was smooth (73.3%) and the others were part-wrinkled.

Multicategorical data were submitted to multivariate analysis, and hierarchical clustering was performed by the UPGMA method, according to coincidence analysis of genetic distances, which resulted in four clusters (Figure 2). The first group (Group I) was formed by six accessions (UENF 1492, UENF 1637, UENF 1714, UENF 1738, UENF 1573, and UENF 1633), the second group (Group II) contained the largest number of accessions (21), the third group (Group III) had only one accession (UENF 1643b), and the fourth group (Group IV) had two accessions (UENF 1417 and UENF 1616). Most of the fruit in Groups I and II were orange and red, representing intermediate-to-ripe stages, respectively, while Groups III and IV had fruit that was orange-yellow to orange in color. Another important feature of these groups is the shape of the fruit. The majority of fruit from Group I was triangular, while fruit from Groups III and IV were elongated and almost round, respectively. Group II clustered accessions with a predominance of elongated fruits (52.38%), followed by triangular and campanulate (both types with 19.05%), and bell-shaped (9.52%) fruits.
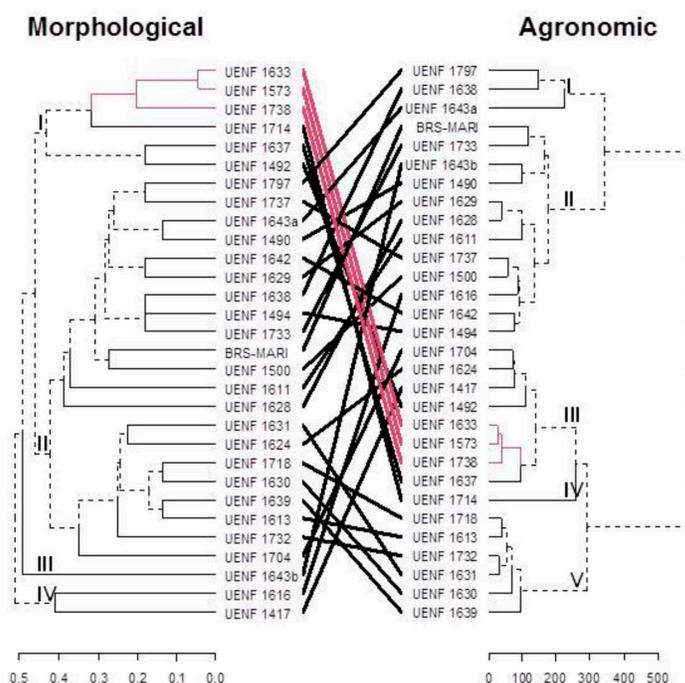
Figure 2. Dendrogram of genetic dissimilarity between 30 genotypes of *Capsicum baccatum* var. *pendulum* obtained by unweighted pair group method using arithmetic averages (UPGMA) based on the dissimilarity matrix of morphological and agronomic descriptors. Campos dos Goytacazes, Brazil.

Significant differences were detected for all quantitative morphoagronomic traits, indicating that there was wide variation among accessions (Table 2). The earliest accessions were UENF 1417, UENF 1490, UENF 1573, and UENF 1738, taking approximately 24 and 71 days for flowering and fruiting, respectively. The UENF 1500 accession presented the latest flowering and fruiting, taking 45 and 89 days after transplanting, respectively.

Table 2. Genetic diversity of 30 *Capsicum baccatum* var. *pendulum* accessions based on 12 microsatellite markers (Campos dos Goytacazes, Brazil).

Locus	A	N_A	H_E	H_O	PIC
CAMS-90	0.516	3	0.529	0.000	0.419
CAMS-142	0.515	3	0.617	0.000	0.548
CAMS-153	0.467	4	0.656	0.933	0.592
CAMS-177	0.464	3	0.610	0.071	0.528
CAMS-190	0.452	3	0.633	0.000	0.557
CAMS-191	0.442	4	0.662	1.000	0.601
CAMS-327	0.567	2	0.491	0.000	0.371
CAMS-376	0.455	4	0.666	0.394	0.605
CAMS-405	0.935	3	0.123	0.000	0.119
CAMS-420	0.719	3	0.442	0.031	0.398
CAMS-424	0.500	3	0.594	0.000	0.511
CAMS-808	0.635	3	0.508	0.500	0.435
Mean	0.516	3	0.544	0.244	0.474

A: more frequent allele, N_A: number of alleles by polymorphic locus, H_E: expected heterozygosity, H_O: observed heterozygosity, and PIC: polymorphism information content.

Accession UENF 1638 produced the longest fruit (122.4 mm) and fruit with high values of FDMC (21.2%) and SSC (10.6 ºBrix), respectively. This accession also had the highest values for petal width, early flowering (26 days), and late fruiting (88 days); therefore, this genotype requires longer time to develop fruits. UENF 1797 had the highest fruit fresh matter per plant and 100-seed mass, and produced fruits with the second longest fruit length (111.9 mm). This accession was clustered with others that had the best mean values for dry fruit mass and soluble solids, and also possessed wider leafs, longer petals and anther, early flowering, and intermediary fruiting (27 and 76 days, respectively). Correlation analysis revealed the highest values for the following combinations: DFM x FFM (r = 0.93), FFM x NFP (r = -0.86), and DFM x NFP (r = -0.85) (Figure 3). NFP showed a highly significant negative correlation by way of the *t*-test for most of the variables (LL, CL, WP, AC, FIL, DFR, FPL, HSM, SSC, FFM, and DFM). Conversely, FFM and DFM showed positive correlations for LL, CL, WP, AC, FIL, DFR, FPL, HSM, and SSC. FDMC conflicted with FFM and DFM, only for AC, FIL, DFR, FPL, and FWT. The FWT showed a positive correlation with DFR and PCW, but a negative correlation with PH, FDMC, HSM, and SSC. Furthermore, SSC showed a positive correlation with PH, LL, LW, PL, CL, WP, FIL, FL, and FPL. The correlation between petal width and fruit length was 0.73, and that between petal width and total number of fruit per plant was -0.59.

Five clusters were formed by UPGMA hierarchical clustering, based on the Mahalanobis distance (Figure 2). Group I had higher values for CL, WP, FL, FPL, FFM, and SSC, and lower values for DFL. Similarly, Group II had high values for FL and SSC, while those for the other variables were intermediate. Interestingly, accessions in Group III had the lowest values for most variables (PH, PCW, LW, PL, CL, WP, FIL, DFR, FL, FDMC, and SSC), except for FWT.

Only UENF 1714 was placed in Group IV, due to its high values for DFL, DFR, NFP, and FDMC, and low values for FL, DFR, FWT, FFM, and DFM. UENF 1613, UENF 1630, UENF 1631, UENF 1639, UENF 1714, UENF 1718, and UENF 1732 were allocated to Group V, and had the greatest values for LL, PL, DFR, and DFM.

Using the method described by Singh (1981) to assess the relative importance of quantitative traits, the variable with the most explanatory power for all accessions was fruit length (44.74%). Fruit diameter (15.56%), peduncle length (8.61%), and soluble solids

(4.27%) followed this trend (Figure 4). The other variables provided little contribution to dissimilarity, as they promoted relative contributions of small magnitude.
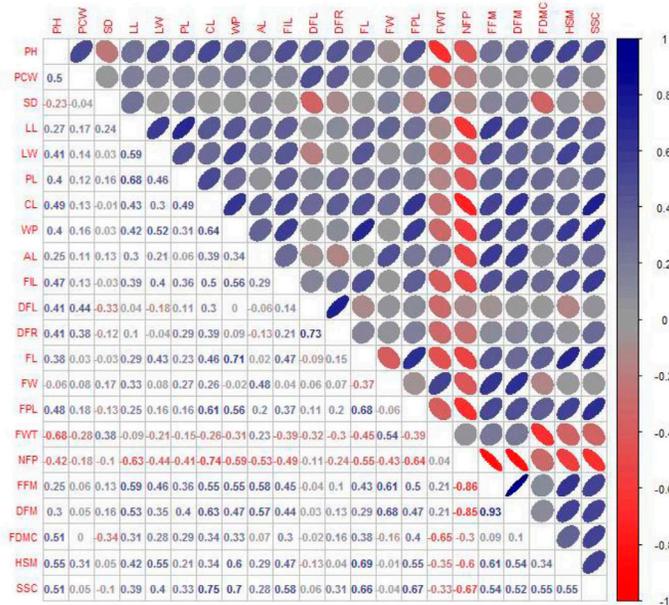
Figure 3. Estimates of Pearson correlation coefficients between 22 agronomic variables of 30 genotypes of *Capsicum baccatum* var. *pendulum*. PH: plant height (cm), PCW: plant canopy width (cm), SD: stem diameter (cm), LL: leaf length (cm), LW: leaf width (cm), PL: petiole length (cm), CL: corolla length (mm), WP: width of the petal (mm), AL: anther length (mm), FIL: filament length (mm), DFL: days to flowering, DFR: days to fruiting, FL: fruit length (mm), FW: fruit width (mm), FPL: fruit pedicel length (mm), FWT: fruit wall thickness (mm), NFP: number of fruits per plant, FFM: fresh fruit mass in gram (g), DFM: dry fruit mass in gram (g), FDMC: fruit dry matter content (%), HSM: 100-seed mass (g), and SSC: soluble solids content (°Brix). Campos dos Goytacazes, Brazil.
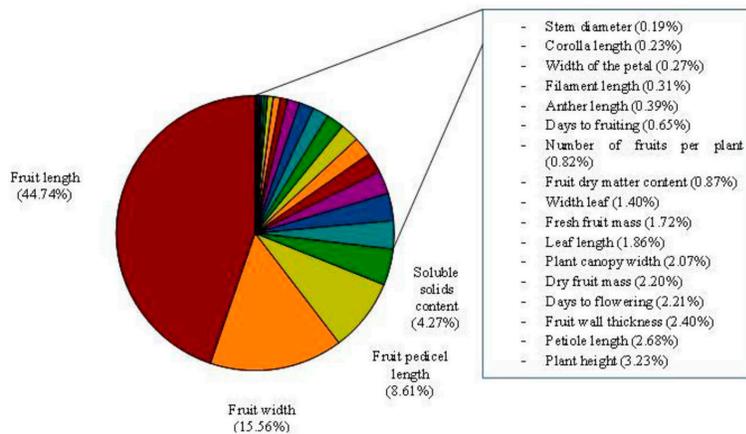
Figure 4. Relative contribution of traits of 22 quantitative morphological descriptors of *Capsicum baccatum* var. *pendulum* for phenotypic divergence by Singh's method (Singh, 1981). Campos dos Goytacazes, Brazil.

Molecular data

Twelve of the 30 microsatellite markers developed for *C. annuum* amplified samples of *C. baccatum* var. *pendulum*. These markers were considered polymorphic with 40% transferability. Analysis of polymorphic loci revealed a total of 41 alleles with an average of three alleles per locus ranging from two (CAMS-327) to four (CAMS-153, CAMS-191, and CAMS-376).

The H_E varied from 0.12 to 0.6, averaging 0.54, while the H_O ranged from 0.00 to 1.00, averaging 0.24 (Table 2), and six loci were homozygous (50%).

The PIC is an index used to estimate how informative each primer is for evaluated accessions. In our study, this ratio ranged from 0.11 to 0.60, averaging 0.47. This index was proposed by Botstein et al. (1980) to measure the extent of polymorphism and markers are considered highly informative when PIC > 0.5, mildly informative at 0.25 < PIC < 0.5, and low informative at PIC < 0.25.

Dissimilarity among the 30 *C. baccatum* accessions ranged from 0.13 (UENF 1638 x UENF1643a and UENF 1573 x UENF 1738) to 0.91 (UENF 1611 x UENF 1719), averaging 0.58. Clustering based on weighted genetic distance identified two groups (Figure 5).
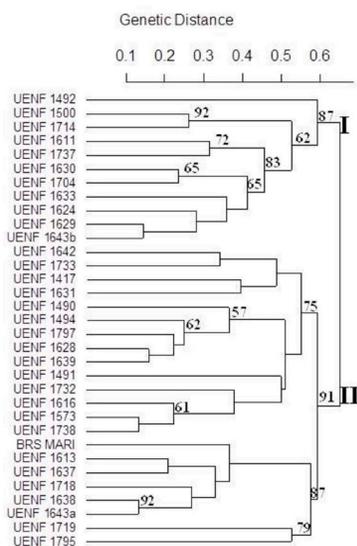

Figure 5. Dendrogram of genetic dissimilarity among 30 genotypes of *Capsicum baccatum* var. *pendulum* by UPGMA method. Campos dos Goytacazes, Brazil.

Eleven accessions were clustered in Group I, in which high intragroup variability was found, ranging from 25.28 to 89.51, 14.37 to 42.14, 2.26 to 12.61, and 6.18 to 9.92 for length, diameter, fresh mass, and soluble solids content, respectively. The 19 accessions of *C. baccatum* clustered in Group II had high morphological and agronomic variability. The values for length, diameter, fresh mass, and soluble solids ranged from 23.67 to 122.41, 12.73 to 46.47, 2.68 to 13.77, and 5.94 to 11.33, respectively, indicating high intragroup variability.

Bayesian analysis was used as an alternative method to classify accessions based on molecular data. This method offers advantages over others, mainly based on genetic distance it can produce highly accurate assignments using modest numbers of loci. The Bayesian analysis

constitutes a robust statistical procedure, with numerous application possibilities. This strategy can be applied when there is heterogeneity of variances, unbalanced data and absence of the normality assumption, conditions that are restrictive for many methodologies (Pritchard et al., 2000). In this context, the highest ΔK value was achieved when two clusters were formed (Figure 6). The groups formed by the Bayesian analysis were consistent with those formed by UPGMA, with the exception of UENF 1492, which was allocated to Bayesian Group II (56% similarity), while in UPGMA this accession was allocated to Group I.
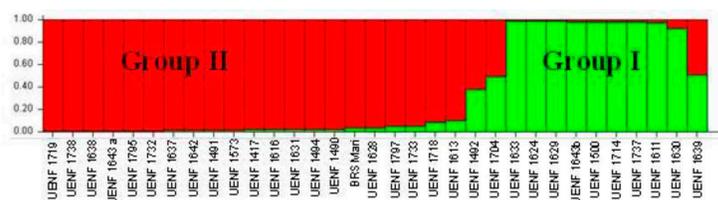
Figure 6. Bar plot structure estimates for 33 accessions of *Capsicum baccatum* var. *pendulum* by 12 microsatellite markers. The accessions are represented on the horizontal line and each genetic group is represented by a distinct color. Campos dos Goytacazes, Brazil.

DISCUSSION

Morphoagronomic characterization

Among nine monomorphic descriptors, three (CC, AC, and FC) are used in the identification of *C. baccatum* species and one (NFA) is used for botanical variety. *C. baccatum* var. *pendulum* is characterized by white corolla with spotted petals, yellow anther and fillet, and one flower per node. Other features of this species include clear green leaves and angled stems; however, these traits are not able to distinguish species. None of the other descriptors, including starry corolla, absence of calyx pigmentation, and greenish stems, are characteristic of the species. Thus, it was possible to detect wide phenotypic variability among accessions, especially for leaf shape, growth habit, flower position, and fruit characteristics.

The observed values for branching plants are consistent with their growth habit, and 60% of the accessions had intermediate or prostrate habit, since erect plants have fewer branches. These characteristics are important for plant spacing and for the adoption of crop practices, such as staking and input use.

In general, the stigma in domesticated *Capsicum* species is at the same level as the anther or is inserted, thus increasing the likelihood of self-pollination. However, when they are at the same level, cross-pollination is still possible. Therefore, precautions should be taken to protect bud flowers in order to ensure selfing, especially during seed regeneration in gene banks to avoid undesirable cross-pollination, resulting in the loss of original accessions. When growing *C. baccatum* to produce seeds, germplasm curators should be more attentive in places where high temperatures are predominant, since this environmental condition may stimulate stigma growth, which would eventually be exserted and ready for pollination by pollen from other plants.

Variability within this collection of *C. baccatum* var. *pendulum* accessions was observed for several key domestication-related characteristics, such as pubescence, anthocyanin presence, branching density, flower and stigma positions, and corolla spot color. Pubescence, anthocyanin content, sparse branching (Albrecht et al., 2012), erect flower and fruit (Pickersgill,

1971), in addition to green spots on the corolla (Albrecht et al., 2012) are related to wild-like species and were found in some of the evaluated accessions in the present study.

An intermediate persistence of fruit and pedicels was also noted. This feature is important for crops intended for dried ground chili pepper processing, which is one way of using "*dedo-de-moça*" pepper (lady's finger). In highly persistent fruit, stalk removal is critical, and if it is ground with fruits, the final product becomes impaired due to changes in the color and flavor. However, in fruit with low pedicel persistence, natural dehiscence may be present, causing crop losses (Neitzke et al., 2008).

Long fruit, as well as fruit with high values of FDMC and SSC, are crucial for *Capsicum*-breeding programs, especially for industrial purposes, focusing dehydrated products for food industry, such as paprika, or for cosmetic and pharmaceutical uses that often need raw material dried fruit. Lannes et al. (2007) studied 49 *C. chinense* accessions and noted that smaller fruits are most suitable for dehydration because they have about 3.5-fold more dry matter than larger fruits. Rêgo et al. (2009) studied general and specific combining ability from reciprocal diallelic crosses with eight parental *C. baccatum* var. *pendulum* accessions and concluded that non-additive effects are involved in the genetic control of fruit length.

Significant differences were detected for all quantitative morphoagronomic traits, indicating the presence of wide variation among accessions (Table 2). Therefore, all of these accessions have potential use in future breeding studies of *C. baccatum* var. *pendulum* depending on the traits to be improved.

The negative correlations to FFM x NFP and DFM x NFP were expected, since plants that produce more fruits generate smaller fruit. The high correlation between the width of the petal and fruit length (0.71), can be justified because larger flowers have larger ovaries and consequently larger fruits. The negative correlation between the width of the petal and total number of fruits per plant (-0.59) led to expectation that larger petals should produce less fruits. Therefore, it is possible to select genotypes for fruit size at the beginning of flowering, enabling indirect selections for this feature. Early descriptors that are significantly correlated with fruit development stage are desirable in characterization studies, and for selection in breeding purposes, this reducing labor, input, and time.

Rêgo et al. (2015), found positive correlations between FFM and FWT, FL, and DFM, relating fruit quality to yield for 40 *C. baccatum* accessions. However, negative correlations were observed for NFP and FDMC. NFP was negatively correlated with SSC and yield-related variables, while SSC was negatively correlated with FWT but positively correlated with FDMC. These results are consistent with those observed in the present study, generating essential information for breeding programs. According to Moreira et al. (2013), the identification of correlations between easily measured and yield-related variables with fruit quality is a main goal of breeding programs, which helps breeders by facilitating and speeding up the selection of superior plants.

According to Lannes et al. (2007), FWT is a characteristic of paramount importance to *in natura* markets, since thick-skinned peppers are the most resistant to damage during handling and post-harvest transport. Furthermore, fruits with thicker pericarps look fresher to consumers when compared to those with a thinner peel. Fruit with a thinner pericarp are ideal for processing, leading to lower costs during the dehydration process for paprika production.

The correlation matrices formed to compare groups based on morphoagronomic characteristics revealed inconsistencies between quantitative and qualitative data, which had poor correlation between distance matrices. This low correlation suggests that both characterizations are important for broad and improved discrimination among *C. baccatum* accessions.

Rêgo et al. (2011), who studied 40 genotypes of *C. baccatum*, also observed the largest contribution of fruit length (32.30%), followed by increased and decreased fruit diameters, fruit fresh mass, and soluble solids content, with contributions of 20.58, 12.20, 7.64, and 5.75%, respectively. Albrecht et al. (2012) assessed phenotypic differences among 170 accessions of *C. baccatum* using 40 descriptors and found that fruit mass was the most informative, followed by fruit diameter, as well as fruit and peduncle lengths. Fruit characteristics contributed the most to divergence. This can occur because the fruit is the portion consumed by humans and also by birds, and is responsible for the dissemination of *Capsicum* seeds. This leads to the hypothesis that selection and domestication have been based on this plant organ, which has a diverse range of colors, shapes, sizes, degree of pungency, and aromas (Paran and van der Knaap, 2007).

Molecular data

Regarding microsatellite markers available for *C. annuum*, of the 30 primers tested, 12 amplified in *C. baccatum* var. *pendulum* and were considered polymorphic, corresponding to a transferability rate of 40%. According to Varshney et al. (2005), SSR flanking regions are highly preserved and generally exhibit high levels of transferability between species. However, the failure of amplification could be attributed to genotypic differences between *C. annuum* and *C. baccatum* considering some specific primer binding sites.

The average number of alleles per locus was consistent with previous reports. Villela et al. (2014) genotyped 20 *C. baccatum* landraces with 8 SSR loci and identified 43 alleles with an average of 5 alleles per locus. Rai et al. (2013) studied 48 accessions belonging to four *Capsicum* species, 79% of which were *C. annuum* species, and obtained 76 alleles, with the number of alleles per locus ranging from two to five, with an average of 3.04. Nicolaï et al. (2013) genotyped 1352 *Capsicum* spp accessions with 28 microsatellite primers and observed 2.68 to 12.57 alleles per locus once the redundant primers were eliminated. In this particular study, the number of alleles per locus was considered to be biased because of the unbalanced number of accessions among species.

The results obtained for H_E are consistent with the autogamous reproduction system of the species. Pacheco-Olvera et al. (2012) studied *C. annuum* and observed values of 0.47 and 0.22 for H_E and H_O, respectively, in wild species, 0.42 and 0.20 in local varieties, and 0.44 and 0.24 in hybrids. González-Pérez et al. (2014) studied the genetic divergence among 102 accessions of *Capsicum* spp belonging to the Vegetable Germplasm Bank of Zaragoza in Spain, and reported values of 0.12 for H_O and 0.43 for H_E, with H_O averaging 0.074 for all accessions.

In the present study, heterozygosity related to species was consistent with the values reported in most studies on *Capsicum*. For example, Carvalho et al. (2015), studying *C. chinense* and *C. frutescens* accessions, observed a smaller number of alleles and increased homozygosity for *C. frutescens*, which was attributed to low genetic variability, since this species has more conserved fruit morphology than reported for domesticated species. Another major factor is how these accessions are kept in collections and germplasm banks, since the seeds should be multiplied periodically. An active germplasm bank of *Capsicum* should regenerate the seeds every 2 years from preventing germination losses. During this procedure, crosses between accessions could take place if the gene bank curator does not take care to ensure self-fertilization, thus maintaining the accession genetic purity. *C. baccatum* is considered to be autogamous species with high rates of cross-pollination, especially under specific conditions such as high temperature.

PIC was considered highly informative for seven of the analyzed loci in the present study. Rai et al. (2013), studying different species of *Capsicum* and primers for *C. annuum*, found PIC values between 0.29 and 0.92, with an average of 0.69. Villela et al. (2014) characterized 20 local varieties of *C. baccatum* using nine microsatellite loci and obtained PIC values ranging from 0.54 to 1.00, with an average value of 0.66.

Unexpectedly, the SSR markers were not efficient at distinguishing accessions of *C. annuum* from *C. baccatum* complexes. UENF 1491 (*C. frutescens*), UENF 1719 (*C. chinense*), and UENF 1795 (*C. annuum*) were clustered into Group II with 19 other *C. baccatum* accessions. However, there was a large distance between *C. chinense* and *C. annuum* and accessions of Group II.

Comparison of the results obtained using molecular markers and morphological descriptors (qualitative and quantitative variables) demonstrated distinct patterns of dissimilarity among the evaluated accessions between the generated matrices (r = -0.002 and r = -0027, respectively). While microsatellite markers cover a larger proportion of the genome, including non-coding regions, morphological descriptors represent the expression of coding regions. Consequently, the molecular profile may not represent regions correlated with phenotypic traits. As already noted for other *Capsicum* species, both molecular and morphoagronomic characterization complement each other in discriminating accessions in germplasm banks (Costa et al., 2006). When using accessions from gene banks as a source of variability in breeding programs, the choice of parent lines should take into account genotyping and phenotyping data, including agronomic traits, in order to meet the expectation of breeders in terms of high genetic value, heterosis, and genetic variability.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Albrecht E, Zhang D, Saftner RA and Stommel JR (2012). Genetic diversity and population structure of *Capsicum baccatum* genetic resources. *Genet. Resour. Crop Evol.* 59: 517-538. http://dx.doi.org/10.1007/s10722-011-9700-y

Baba VY, Rocha KR, Gomes GP, Ruas CF, et al. (2015). Genetic diversity of *Capsicum chinese* accessions based on fruit morphological characterization and AFLP markers. *Genet. Resour. Crop Evol.* 40: 1-11.

Botstein D, White RL, Skolnick M and Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.

Carvalho SIC, Ragassi CF, Oliveira IB, Amaral ZPS, et al. (2015). Transferability of microsatellite markers of *Capsicum annuum* L. to *C. frutescens* L. and *C. chinense* Jacq. *Genet. Mol. Res.* 14: 7937-7946. http://dx.doi.org/10.4238/2015. July.17.1

Costa FR, Pereira TNS, Vitória AP, Campos KP, et al. (2006). Genetic diversity among *Capsicum* accessions using RAPD markers. *Crop Breed. Appl. Biotechnol.* 6: 18-23. http://dx.doi.org/10.12702/1984-7033.v06n01a03

Cruz CD (2013). GENES - a software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum* 35: 271-276.

D'Arcy WG and Eshbaugh WH (1974). New world peppers (*Capsicum* - Solanaceae) north of Colombia: A resume. *Baileya* 19: 93-105.

_____

Daher RF, Pereira MG, Pereira AV and Antônio Jr TA (2002). Genetic divergence among elephantgrass cultivars assessed by RAPD markers in composite samples. *Sci. Agric.* 59: 623-627. http://dx.doi.org/10.1590/S0103-90162002000400001

Doyle JJ and Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19: 11-15.

Earl DA and vonHoldt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4: 359-361. http://dx.doi.org/10.1007/s12686-011-9548-7

Eshbaugh WH (1970). A biosystematic and evolutionary study of *Capsicum baccatum* (Solanaceae). *Brittonia* 22: 31-43. http://dx.doi.org/10.2307/2805720

Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620. http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x

FAO (Food and Agriculture Organization of the United Nations) (2012). Statistcs: [http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567]. Accessed January 10, 2016.

Filgueira FAR (2012). Novo manual de olericultura: agrotecnologia moderna na produção e comercialização de hortaliças. 3ª. edn. Viçosa: UFV. 420.

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

González-Pérez S, Garcés-Claver A, Mallor C, Sáenz de Miera LE, et al. (2014). New insights into *Capsicum* spp. relatedness and the diversification process of *Capsicum annuum* in Spain. *PLoS One* 9: e116276. http://dx.doi.org/10.1371/journal.pone.0116276

Ibiza VP, Blanca J, Cañizares J and Nuez F (2012). Taxonomy and genetic diversity of domesticated *Capsicum* species in the Andean region. *Genet. Resour. Crop Evol.* 6: 1077-1088. http://dx.doi.org/10.1007/s10722-011-9744-z

IPGRI (1995). Descriptors for *Capsicum* (*Capsicum* spp.). Rome: International Plant Genetic Resources Institute, 49.

Lannes SD, Finger FL, Schuelter AR and Casali VWD (2007). Growth and quality of Brazilian accessions of *Capsicum chinense* fruits. *Sci. Hortic.* 112: 266-270. http://dx.doi.org/10.1016/j.scienta.2006.12.029

Minamiyama Y, Tsuro M and Hirai M (2006). SSR-based linkage map of *Capsicum annuum. Mol. Breed.* 18: 157-169. http://dx.doi.org/10.1007/s11032-006-9024-3

Moses M, Umaharan P and Dayanandan S (2014). Microsatellite based analysis of the genetic structure and diversity of *Capsicum chinense* in the Neotropics. *Genet. Resour. Crop Evol.* 61: 741-755. http://dx.doi.org/10.1007/s10722-013-0069-y

Moreira SO, Rodrigues R, Oliveira HS, Medeiros AM, et al. (2013). Phenotypic and genotypic variation among *Capsicum annuum* recombinant inbred lines resistant to bacterial spot. *Genet. Mol. Res.* 12: 1232-1242. http://dx.doi.org/10.4238/2013.April.17.2

Moulin MM, Rodrigues R, Bento CS, Gonçalves LSA, et al. (2015). Genetic dissection of agronomic traits in *Capsicum baccatum* var. pendulum. *Genet. Mol. Res.* 14: 2122-2132. http://dx.doi.org/10.4238/2015.March.20.23

Neitzke RS, Barbieri RL, Heiden G and Castro CM (2008). Divergência genética entre variedades locais de *Capsicum baccatum*, utilizando caracteres multicategóricos. *Magistra* 3: 249-255.

Nicolaï M, Cantet M, Lefebvre V, Sage-Palloix A, et al. (2013). Genotyping a large collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for the wild origin of cultivated *C. annuum* and the structuring of genetic diversity by human selection of cultivar types. *Genet. Resour. Crop Evol.* 60: 2375-2390. http://dx.doi.org/10.1007/s10722-013-0006-0

Pacheco-Olvera A, Hernandez-Verdugo S, Rocha-Ramirez V, Gonzalez-Rodriguez A, et al. (2012). Genetic diversity and structure of pepper (*Capsicum annuum* L.) from Northwestern Mexico analyzed by microsatellite markers. *Crop Sci.* 52: 231-241. http://dx.doi.org/10.2135/cropsci2011.06.0319

Paran I and van der Knaap E (2007). Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *J. Exp. Bot.* 58: 3841-3852. http://dx.doi.org/10.1093/jxb/erm257

Pickersgill B (1971). Relationships between weedy and cultivated forms in some species of chili peppers (Genus *Capsicum*). *Evolution* 4: 683-691. http://dx.doi.org/10.2307/2406949

Pritchard JK, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Rai VP, Kumar R, Kumar S, Rai A, et al. (2013). Genetic diversity in *Capsicum* germplasm based on microsatellite and random amplified microsatellite polymorphism markers. *Physiol. Mol. Biol. Plants* 19: 575-586. http://dx.doi.org/10.1007/s12298-013-0185-3

Rêgo ER, Rêgo MM, Finger FL, Cruz CD, et al. (2009). A diallel study of yield components and fruit quality in chilli pepper (*Capsicum baccatum*). *Euphytica* 168: 275-287. http://dx.doi.org/10.1007/s10681-009-9947-y

Rêgo ER, Rêgo MM, Cruz CD, Finger FL, et al. (2011). Phenotypic diversity, correlation and importance of variables for fruit quality and yield traits in Brazilian peppers (*Capsicum baccatum*). *Genet. Resour. Crop Evol.* 58: 909-918. http://dx.doi.org/10.1007/s10722-010-9628-7

Rêgo MM, Sapucay MJLC, Rêgo ER and Araújo ER (2015). Analysis of divergence and correlation of quantitative traits in ornamental pepper (*Capsicum* spp.). *Acta Hortic.* 1087: 389-394. http://dx.doi.org/10.17660/ActaHortic.2015.1087.52

Rodrigues R, Gonçalves LSA, Bento CS, Sudré CP, et al. (2012). Combining ability and heterosis for agronomic traits in chili pepper. *Hortic. Bras.* 30: 226-233. http://dx.doi.org/10.1590/S0102-05362012000200008

Singh D (1981). The relative importance of characters affecting genetic divergence. *Indian J. Genet. Plant Breed.* 2: 237-245.

Sudré CP, Gonçalves LSA, Rodrigues R, do 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. http://dx.doi.org/10.4238/vol9-1gmr698

Varshney RK, Graner A and Sorrells ME (2005). Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* 23: 48-55. http://dx.doi.org/10.1016/j.tibtech.2004.11.005

Villela JCB, Barbieri RL, Castro CM, Neitzke RS, et al. (2014). Caracterização molecular de variedades crioulas de pimentas (*Capsicum baccatum*) com marcadores microssatélites. *Hortic. Bras.* 2: 131-137. http://dx.doi.org/10.1590/S0102-0536201400020002