Selection of Morphoagronomic Descriptors in *Physalis angulata* L. Using Multivariate Techniques

Hortência Kardec da Silva[1], Adriana Rodrigues Passos[2], Alessandra Selbach Schnadelbach[3],
Ricardo Franco Cunha Moreira[4], Antonio Leandro da Silva Conceição[4] & André Pinto Lima[2]

[1] Vegetable Genetic Resources Post-Graduation Program, State University of Feira de Santana, Feira de Santana, BA, Brazil

[2] Biological Sciences Department, State University of Feira de Santana, Feira de Santana, BA, Brazil

[3] General Biology Department, Federal University of Bahia, Salvador, BA, Brazil

[4] Center of Environmental and Biological Sciences, Federal University of Recôncavo da Bahia, Cruz das Almas, BA, Brazil

Correspondence: Hortência Kardec da Silva, Vegetable Genetic Resources Post-Graduation Program, State University of Feira de Santana, Av. Transnordestina, s/n, Labio, Novo Horizonte, Feira de Santana, BA, CEP: 44036900, Brazil. Tel: 55-75-3161-8000. E-mail: hortenciakardec@hotmail.com

Abstract

This study aimed at selecting determinant morphoagronomic descriptors to characterize and evaluate *Physalis angulata* L. germplasm. Twelve quantitative and twenty-two qualitative descriptors were analyzed in six accessions of *P. angulata* coming from the physalis germplasm collection belonging to the State University of Feira de Santana-BA. The selection and discharge of quantitative descriptors was based on the direct selection and on the Singh method, while qualitative descriptors were analyzed through entropy. The statistic analyses were carried out using the GENES and R programs. Ten quantitative descriptors were excluded through direct selection and five through the Singh method. However, only four descriptors were considered redundant by both methods: east-west fruit, weight of five ripe fruits, width of leaf blade and total soluble solids. Although the total soluble solids descriptor was appointed for discharge, it was included in the group of descriptors selected due to its importance in the characterization of physalis fruit. The list of minimum descriptors to describe physalis accessions comprised 15 descriptors: plant height, stem diameter, north-south fruits, number of fruits per plant, leaf blade length, internode length, fruit longitudinal length, fruit transversal length, total soluble solids, growth habit, stem color, leaf margin shape, unripe calyx color, unripe fruit shape and color. These were nine quantitive and six qualitative descriptors, respectively. The discharge of 55.88% of the descriptors did not cause significant loss of information and might allow the reduction of time and resources spent to characterize and evaluate physalis germplasm.

Keywords: morphoagronomic characteristics, multivariate analysis, physalis, principal components, variability

1. Introduction

The organization and maintenance of species genetic resources, considering the activities of collection, characterization and conservation, are essential to sustain the genetic basis of genetic improvement programs (Silva, Moura, Neto-Farias, Ledo, & Sampaio, 2017). Characterization is a vital activity to generate knowledge about germplasm, preserved in bases or collections, for providing better management of accessions and subsidizing their conservation and preservation, as well as for the use in genetic improvement programs (Dantas, Nunes, Araújo, & Albuquerque, 2012). It can be carried out aided by botanic, morphological and agronomic descriptors, which can be either quantitative or qualitative (Veiga & Queiróz, 2015).

When the number of characters is high, it is possible that some of them are redundant and do not contribute much for the discrimination of the individuals under analysis, since these properties are usually correlated (Daher, Moraes, Cruz, Pereira, & Xavier, 1997), or might be neglected for representing a non-representative fraction of the total variation (Alves, Garcia, Cruz, & Figueira, 2003). Therefore, the use of a large number of descriptors might result in an increase in work without improving the accuracy of the characterization, in addition to making the data analysis and interpretation more complex. Moreover, the elimination of redundant descriptors which are difficult to measure becomes desirable in order to facilitate studies, reducing experimentation time and costs, without, however, affecting the reliability and quality of the results (Pereira, Vencovsky, & Cruz, 1992).

Multivariate analysis are useful tools to identify the more informative and the less relevant descriptors in the characterization and evaluation of a germplasm to be used in genetic improvement programs, since it supplies information to exclude descriptors that represent little contribution to the total variation (Cruz, Carneiro, & Regazzi, 2004). The use of more than one procedure to discard redundant characters has been adopted to provide higher security to the selection of descriptors (Castro, Neves, Jesus, & Oliveira, 2012; Afonso et al., 2014; E. Oliveira, F. Oliveira, & Santos, 2014; Silva, Carvalho, & Duarte, 2013; Silva et al., 2017). Thus, other methodologies have been used in the evaluation of the discharge efficiency such as the Singh Method (1981), based on the Mahalanobis distance (D^2), which considers less important the characteristics that present lower variability (Chagas, Alexandre, Schimildt, Bruckner, & Faleiro, 2016).

Physalis angulata L., family Solanaceae, is an annual herbaceous species that reaches up to one meter in height, and a calyx formed by five sepals surrounds its fruit (N. Sultana, Hassan, Begum, & M. Sultana, 2008). This plant produces fruits with high nutraceutical value and quality characteristics that favor its fresh consumption , as well as its potential for commercial cultivation and use in the food industry (Oliveira, Martins, Vasconcelos, Pena, & Carvalho, 2011; P. Silva, S. Silva, J. Silva, Mendonça, & Pereira, 2018). In addition, the species outstands for the production of secondary metabolites with great medicinal potential (Tomassini, Barbi, Ribeiro, & Xavier, 2000).

Techniques to select descriptors were used with species of the Solanaceae family such as tomato (Gonçalves, Rodrigues, Amaral Júnior, Karasawa, & Sudré, 2009) and *Capsicum* spp. (Ortiz, Flor, Alvarado, & Crossa, 2010; Sudré et al., 2010; Silva et al., 2013). However, even with the evident importance of *P. angulata*, no studies were found in the literature about the discriminating capability of the descriptors or the suitable number of descriptors able to discriminate the accessions. Therefore, more information is needed that might guide further studies, both in the area of genetic resources and the improvement of plants of this species.

For this reason, the objective of this study was to select relevant morphoagronomic descriptors to characterize and evaluate *P. angulata* germplasm.

2. Method

2.1 Vegetable Material

Six *P. angulata* accessions were used, originated from the Germplasm collection belonging to the University State of Feira de Santana (UEFS), Feira de Santana, Brazil. Such accessions came from collections carried out in Bahia and Piauí, and from which, the three accessions coming from Anguera have already been submitted by Silva (2007) and Araujo (2012) to three autofecundation selection cycles (Table 1).

Table 1. List of the accesses characterized with the respective denominations, origins, geographical coordinates and year of collection. UEFS. Feira de Santana, BA, 2018

Access	Origin		
	City/State	Geographical coordinates	Year
BGF 16-001	Feira de Santana/BA	12º 26′ 67″S; 38º 96′ 67″W	2016
BGF 17-002	Teresina/PI	05º 00′ 25″S; 42º 49′ 55″W	2017
BGF 16-003	Candeias/BA	12º 77′ 05″S; 38º46′ 52″W	2016
BGF 05-004	Anguera/BA	12º 09′ 04″S; 39º 14′ 47″W	2005
BGF 05-005	Anguera/BA	12º 09′ 04″S; 39º 14′ 47″W	2005
BGF 05-006	Anguera/BA	12º 09′ 04″S; 39º 14′ 47″W	2005

2.2 Study Area

The experiment was carried out in the period from April to September 2017 at the Unidade Experimental Horto Florestal (Forest Garden Experimental Unit) (12°16′087″S; 38°56′346″W; 243 m altitude), belonging to the State

University of Feira de Santana (UEFS) in Feira de Santana, BA. According to Thornthwaite and Mather classification (1955), the region has a sub-humid climate, megahermic (C2rA'a'), with an annual rainfall average of 848 mm and an annual mean temperature of 24 ºC. During the summer, the region can reach an average monthly temperature of 27 ºC, and during the winter, 21 ºC (INMET, 2017). The soil of this region is a Plantasol Haplic Salic (Dias, Souza, Oliveira, & Santos, 2010).

2.3 Morphoagronomic Characterization

The experiment was carried out on an experimental field, in complete randomized blocks design, with three replications and a useful plot of six plants, totalling 18 plants per treatment. The space between rows was 1.0 m and 0.8 m between plants. To form the seedlings, three seeds were sowed in 300 mL disposable plastic cups, containing the commercial substrate TechnsVivato. Thinning was carried out 15 days after seeding, keeping only the most vigorous plants.

Later on, the plants were kept in screen house with daily manual watering early in the morning and late afternoon. When the seedlings were around 20 cm high, they were transplanted to open field conditions, with drip irrigation. The adjustment of the fertilizing process was carried out based on the chemical analysis of the soil (Appendix A) and following the recommendation for the tomato crop. For the pests control, it was used neem oil, which was applied with manual sprayer at the beginning of the plant development.

The evaluations started two months after the transplanting, from the flowering. Morphoagronomic descriptors, proposed by the National University of Colombia, (González, Torres, Cano, Arias, & Arboledo, 2008), were used with adjustments for the species. The physiologically mature fruits were randomly collected at each accession. The mature stage was evidenced by the intensification of the yellow color of the calyx. The fruits were packed in plastic bags, labeled and taken to the Laboratory of Molecular Genetics at the Horto Florestal Experimental Unit in Feira de Santana. At the laboratory, the fruits were separated from the calyxes, washed in running water, and dried. Then, the evaluations were carried out.

Twelve quantitative and twenty- two qualitative descriptors were used (Table 2). It is important to state that, for the fruit number analysis, besides the descriptor "number of fruits per plant", two other descriptors were included, "fruits of the north-south" and "east-west" axis, due the fact that some plants lacked branches in a specific orientation. Qualitative variables were carried out based on the phenotypical class of the descriptors for *P. angulata* (González et al., 2008) and according to the color catalogue by the Royal Horticultural Society (The Royal Horticultural Society, 2001). Quantitative and qualitative descriptors were analyzed in five leaves and fruits of six plants from each accession, totalling 18 plants evaluated per accession.

Table 2. Quantitative and qualitative descriptors for the characterization of six accessions of *P. angulata* from UEFS. UEFS, Feira de Santana, BA, 2018

Quantitive descriptors	
Descriptor	Methodology
Plant height (PH)	It was measured from the bottom to the top of the main branch with a measuring tape (cm).
Stem diameter (SD)	It was measured from the bottom to 5 cm from the soil with a digital caliper (mm).
North-south fruits (NSF)	It was determined the number of fruits from the north-south axis in four branches.
East-west fruits (EWF)	It was counted the number of fruits from the east-west axis in four branches
Weight of five ripe fruits (RFW)	It was calculated the mean of 5 randomly picked ripe fruit, using a precision scale (g)
Number of fruits per plant (NFP)	It was determined the number of fruits per plant
Leaf blade length (LBL)	It was calculated the mean of the leaf length of 5 randomly picked leaves, from the base to the top of the blade, with a millimeter ruler (cm).
Leaf blade width (LBW)	It was calculated the mean of the width of the leaf blade, on the base, of 5 randomly picked leaves, using a millimeter ruler (cm).
Entrenode length (EL)	Mean of the length between 5 randomly picked nodes, using a measuring tape (cm).
Fruit longitudinal lenght (FLL)	The mean of the longitudinal axis of 5 randomly picked ripe fruits was measured with a digital caliper (mm).
Fruit transversal length (FTL)	The mean of the transversal axis of 5 randomly picked ripe fruits was measured with a digital caliper (mm).
Total soluble solids (TSS)	The mean of the total soluble solids content of five random fruits selected at random, expressed in ºBrix, was determinated by using a drop of fruit pulp extract with a digital refractometer with the temperature correction at 20 ºC.

Qualitative descriptors	
Descriptor	Classes
Growth habit (GH)	(1) Erect; (2) Semi-erect; (3) Prostrated.
Stem color (SC)	(1) Green; (2) Crimson.
Stem pubescence (SP)	(1) Absent; (2) Weak; (3) Average; (4) Abounding.
Stem anthocyanin (SA)	(1) Absent; (2) Present.
Leaf blade shape (LBS)	(1) Lanceolate; (2) Cordate; (3) Assymetrical.
Leaf margin shape (LMS)	(1) Serrate; (2) Undulate; (3) Repando; (4) Undulate; (5) Sinuate.
Leaf base shape (SBS)	(1) Inequilateral; (2) Cuneate; (3) Obliquous; (4) Cordate; (5) Subcordate.
Leaf apex shape (LAS)	(1) Acute; (2) Apiculate; (3) Acuminate.
Leaf bundle pubescence (LBP)	(1) Absent; (2) Weak; (3) Average; (4) Abounding.
Pubescence around the leaves (PAL)	(1) Absent; (2) Weak; (3) Average; (4) Abounding.
Leaf vein anthocyanin (LVA)	(1) Absent; (2) Present.
Flower position (FlP)	(1) Erect; (2) Pendulous.
Corolla color (Ccl)	(1) RHS 8C; (2) Yellow 5C; (3) Yellow 5B; (4) Yellow 6B.
Color of the corolla stains (CCS)	(1) Coffe; (2) Brown.
Color of the unripe calyx (CUC)	(1) Green without anthocyanin; (2) Green with light anthocyanin; (3) Green with dark anthocyanin.
Color of pedicel (CP)	(1) Green; (2) Crimson.
Fruit shape (FS)	(1) Round; (2) Slightly flat; (3) Flat; (4) Heart-shaped.
Unripe fruit color (UFC)	(1) RHS 144A; (2) RHS145B; (3) Yellow green 152C; (4) Yellow orange 23A.
Ripe fruit color (RFC)	(1) RHS160B; (2) RHS 145C; (3) Purple; (4) Green purple.
Type of calyx (Tcal)	(1) Accrescent; (2) Maccrescent; (3) Globular.
Calyx shape (Scal)	(1) Enlongated; (2) Slightly flat; (3) Flat.
Calyx division (Dcal)	(1) Absent; (2) Present.

2.4 Statistical Analyses

2.4.1 Selection of Quantitative Descriptors

The data obtained through the quantitative characters were analyzed using descriptive statistics, position measurements (mean) and dispersion measurements (maximum and minimum values, standard deviation and variation coeficiente). In addition, the Shapiro-Wilk Normality test was employed.

The recognition of redundant descriptors was carried out through two methods: (1) Direct selection (Jolliffe, 1972, 1973), which is indicated to discard any descriptor that presents the highest weighting coefficient in absolute value (eigenvector), in the principal componente of lowest eigenvalue, from the last component to that whose eigenvalue did not exceed 0.70; (2) Selection according to the Singh coefficient (1981), taking into consideration the relative contribution of each of the characteristics to the genetic divergence.

Considering the coincident information in both techniques, the final discharge of descriptors was carried out, excluding those suggested as redundant in both methods. The Pearson correlation coefficients were calculated, among all descriptors, to verify the association between the discarded descriptors and the remaining ones, helping the discharge decision. The significance of the correlation coefficient was verified through the t test.

2.4.2 Selection of Qualitative Descriptors

For the qualitative characters, the percentage frequency of each of the classes and the entropy level of the descriptors were calculated using the Renyi entropy coefficient (Renyi, 1961). The higher the number of phenotypical classes and the more homogeneous the balance between the accessions frequency in the different phenotypical classes are, the higher the entropy value of any descriptor is (Vieira, Fialho, & Faleiro, 2007). Descriptors presenting level of entropy 0.00 were discarded.

The statistical analyses were aided by the programs GENES (Cruz, 2013) and R (R Core Team, 2013).

3. Results

Table 3 shows the amplitude of the values presented from the descriptive statistics of the quantitative characters investigated. The coefficient of variation (CV) oscillated between 8.78% and 68.21%, corresponding to the fruit transversal lenght (FTL) and number of fruit per plant (NFP), respectively. Restrepo and Vallejo (2003) estimated the genetic variability in tomato accessions, and obtained a coefficient of variation equal to 82.34%,

30.82% and 30.04% for the number of fruits per plant, average fruit weight and plant height, respectively. The high coefficient of variation recorded for the NFP character represents a high genetic variability characteristic of the evaluated material. Thus, the CV values found suggested a wide variation between the results, presenting variability and heterogeneity between the accessions, and a possibility to be explored in breeding programs.

The variable number of fruits per plant (NFP) showed the greatest variation, whose difference between the maximum and minimum values was 436.57, presenting a 188.96 mean, while the leaf blade width (LBW) presented the least variation between maximum and minimum values (2.96 cm and 5.12 cm, respectively), with a 3.94 mean (Table 3). Moreno, Fischer and Sánchez (2012) obtained results that confirm the ones found in this study, with the greatest variation for the variable number of fruits per plant (NFP), in 54 accessions of a collection of *Physalis peruviana* L. germplasm, originated in the central and northeast regions of Colombia.

Regarding the normality test, most of the variables were seen to present normal distribution, since the variables were not significant in the Shapiro-Wilks test with 5% and 1% significance levels (Table 3).

Table 3. Descriptive statistics and Shapiro-Wilk normality test for the quantitative characters. UEFS, Feira de Santana, BA, 2018

Variables	Minimum	Maximum	Mean	Standard deviation	CV (%)	W
PH	33.94	75.02	46.08	11.47	24.90	0.86**
SD	13.53	27.46	18.40	3.55	19.28	0.89*
NSF	5.50	10.33	7.75	1.53	19.71	0.95[ns]
EWF	4.83	11.00	7.55	1.76	23.25	0.97[ns]
RFW	8.09	16.40	12.00	2.93	24.41	0.91[ns]
NFP	79.60	516.17	188.96	128.88	68.21	0.79**
LBL	6.81	10.16	8.20	1.03	12.51	0.94[ns]
LBW	2.96	5.12	3.94	0.61	15.46	0.95[ns]
EL	3.46	6.35	4.74	0.86	18.14	0.95[ns]
FLL	11.31	19.01	15.18	1.76	11.62	0.90*
FTL	12.16	16.47	13.86	1.22	8.78	0.86**
TSS	8.37	12.19	10.40	1.15	11.02	0.96[ns]

Note. Plant height (PH) in cm; Stem diameter (SD) in mm; North-south fruits (NSF) in units; East-west fruits (EWF) in units; Weight of five ripe fruits (RFW) in kg/plant^{-1}; Number of fruits per plant (NFP) in units; Leaf blade length (LBL) in cm; Leaf blade width (LBW) in cm; Entrenode length (EL) in cm; Fruit longitudinal length (FLL) in mm; Fruit transversal length (FTL) in mm; Total soluble solids (TSS) in ºBrix. ** significant at 1%, * significant at 5% probability in the Shapiro-Wilks test. [ns] not significant. Normality test (W), Coefficient of variation (CV).

Based on the Singh coefficient (1981), the characters that provided greater relative contributions regarding the genetic diversity of the accessions were north-south fruits (NSF), with 20.14% contribution, followed by the descriptors of stem diameter (SD), with 17.23%, entrenode length (EL), with 11.60%, leaf blade length (LBL), with 10.92%, plant height (PH), with 9.15%, number of fruits per plant (NFP), with 6.98% and fruit transversal axis (FTL), with 6.49%. These seven characteristics contributed with 82.51% of the total distribution (Table 4).

Table 4. Relative contribution of twelve descriptors to the genetic divergence between six *P. angulata* accessions, through the method proposed by Singh (1981). UEFS, Feira de Santana, BA, 2018

Descriptors	S.j.	Value (%)
Plant height (PH)	1314048	9.15
Stem diameter (SD)	2474142	17.23
North-south fruits (NSF)	2891718	20.14
East-west fruits (EWF)	280 442.4	1.95
Weight of five ripe fruits (RFW)	69 172.55	0.48
Number of fruits per plant (NFP)	1003530	6.98
Leaf blade length (LBL)	1569175	10.92
Leaf blade width (LBW)	758 567.7	5.28
Entrenode length (EL)	1666158	11.60
Fruit longitudinal length (FLL)	804197	5.60
Fruit transversal length (FTL)	931 860.4	6.49
Total soluble solids (TSS)	594 632.8	4.14

Sudré, Rodrigues, Riva, Karasawa, and Amaral Júnior (2005), when evaluating the genetic divergence between 56 accessions of the *Capsicum* spp. germplasm collection of the State University of Norte Fluminense (UENF), using eleven quantitative descriptors, verified a different result from the one found in this work. They reported fruit length, stem diameter, number of seeds per fruit and fruit mean weight as the descriptors that most contributed to the genetic difference (32%, 32%, 13% and 12%, respectively). Therefore, it seems very relevant to carry out studies in specific collections, as well as to consider different environmental conditions at the time of the germplasm characterization and evaluation through morphoagronomic descriptors (Miguel, 2017).

The descriptors with the least contribution were: weight of five ripe fruits (RFW), with 0.48%; east-west fruits (EWF), with 1.95%; total soluble solids (TSS), with 4.14%; leaf blade width (LBW), with 5.28% and fruit longitudinal length (FLL), with 5.60% (Table 4). These characteristics resulted in only 17.45% relative importance, and, therefore were considered of little relevance in the characterization of the physalis accessions. Thus, these variables might be initially identified to be discarded, since according to E. Rego, M. Rego, Cruz, and Cecon (2003) those characteristics that contributed with a very low percentage or that did not contribute at all to the variability found should be discarded. As a criterion for this analysis, variables with relative contribution below 6% were discarded. These were five descriptors, that is, 41.66% of the descriptors were discarded using this method (Table 4).

The eigenvalues estimates associated to the principal components and their respective relative and accumulated variances obtained for the 12 quantitative characters are represented in Table 5. The first two principal components explained 89.49% of the accumulated total variation, while the relative variations and their respective percentages showed that most of the variation was concentrated up to the 4th principal component, corresponding to 98.35% of the whole variation available in the germplasm collection.

Table 5. Estimates of eigenvalues associated to the Principal Components and their total and accumulated variance, obtained from the twelve descriptors evaluated in the *P. angulata* six accessions under study. UEFS, Feira de Santana, BA, 2018

Principal Components	Eigenvalues	Variance (%)	Accumulated variance (%)
3	0.68	5.66	95.15
4	0.38	3.20	98.35
5	0.20	1.65	100.00
6	0.00	0.00	100.00
7	0.00	0.00	100.00
8	0.00	0.00	100.00
9	0.00	0.00	100.00
10	0.00	0.00	100.00
11	0.00	0.00	100.00
12	0.00	0.00	100.00

A percentage close to the one observed in this work, was reported by Moreira (2012), who analyzed the first two principal components and obtained 91.07% when characterizing lines of *Capsicum annum* L. using quantitative and qualitative descriptors. According to Cruz et al. (2004), for the study of diversity using principal components, the first couple of variables must retain most of the total variation, in general, over 80%. To Pereira et al. (1992), the variance distribution is associated to the nature and number of characteristics employed in the analysis, and it is only concentrated in the first couple of components when few characteristics of agricultural interest are evaluated or when they belong to specific parts of the plant.

According to the preliminary discharge through the direct method proposed by Jolliffe (1972, 1973), 10 out of 12 characteristis (83.33%) that presented the highest weighting coefficients, in absolute value from the last principal component might be discarded, due to the number of components that presented eigenvalues below 0.7 (Table 6). According to Hair Junior, Black, Babin, Anderson, and Tatham (2009), the Principal Components technique is advantageous, because it enables the evaluation of the importance of each of the characteristics investigated to the total variation between the accessions evaluated, allowing the elimination of less informative characters, since they are already correlated to the remaining variables or due to their invariance.

Table 6. Estimates of the weight coefficients associated to the Principal Components of eigenvectors below 0.70 and identification of the characteristics to be discarded in each component, through direct selection of *P. angulata* six accessions under evaluation. UEFS, Feira de Santana, BA, 2018

Variables	Principal Components									
	CP3	CP4	CP5	CP6	CP7	CP8	CP9	CP10	CP11	CP12
PH	0.31	0.29	0.31	0.32	0.32	0.30	0.24	0.22	0.28	0.19
SD	0.03	0.04	0.00	-0.04	0.05	0.02	-0.41	**0.56**	0.44	-0.53
NSF	-0.31	-0.44	-0.10	0.02	0.07	0.44	**0.56**	0.05	0.02	-0.42
EWF	-0.19	-0.34	-0.41	-0.03	0.15	-0.23	0.08	0.51	0.16	**0.56**
RFW	-0.02	-0.29	0.55	-0.37	**0.40**	0.08	-0.13	0.01	-0.16	0.13
NFP	-0.15	0.22	-0.24	-0.16	-0.12	**0.78**	-0.29	0.00	0.00	0.32
LBL	-0.01	0.00	-0.04	**0.73**	-0.03	0.00	0.05	-0.04	-0.03	0.05
LBW	-0.25	-0.14	**0.60**	0.10	-0.51	0.02	0.08	0.23	0.13	0.23
EL	-0.07	-0.12	-0.01	-0.11	0.03	-0.06	0.02	-0.54	**0.81**	0.11
FLL	-0.01	-0.01	-0.01	-0.09	-0.61	0.00	-0.02	0.05	0.03	-0.04
FTL	**0.80**	-0.23	-0.10	-0.22	-0.24	0.10	0.26	0.10	0.04	0.05
TSS	-0.21	**0.62**	0.01	-0.35	0.00	-0.16	0.52	0.16	0.06	-0.01

Note. Plant height (PH); Stem diameter (SD); North-south fruits (NSF); East-west fruits (EWF); Weight of five ripe fruits (RFW); Number of fruit per plant (NFP); Leaf blade length (LBL); Leaf blade width (LBW); Entrenode length (EL); Fruit Longitudinal length (FLL); Fruit transversal length (FTL); Total soluble solids (TSS).

The first descriptor indicated for discharge through the direct selection was the variable "east-west fruits", which presented the highest weighting coefficient in module with the last principal component (0.56) followed by the characteristics "entrenode length", "stem diameter" and "north-south fruits", whose highest eigenvectors in module occurred in the Principal Components 11, 10 and 9, respectively (Table 6).

Using this procedure, ten characters were found redundant, according to the discharge sequence: EWF, EL, SD, NSF, NFP, RFW, LBL, LBW, TSS and FTL. The procedure might be considered too strict, since it eliminated 10 of the 12 quantitative characters used as descriptors for physalis. Similar results were reported by Alvares (2011), who evaluated the genetic difference between 137 *Capsicum chinense* Jacq. accessions and verified that 14 out of the 25 morphoagronomic descriptors should be discarded based on the method proposed by por Jolliffe (1972, 1973).

The information obtained using the two methodologies was in disagreement since in the procedure proposed by Jolliffe (1972, 1973) 10 characters were appointed to be discarded against only five in the Singh analysis. Therefore, aiming at reducing inconsistencies in the elimination of descriptors, the adoption of two procedures is common to indicate the most relevant descriptors. For example, Afonso et al. (2014), and Castro et al. (2012) employed the same direct selection criteria proposed by Jolliffe (1972, 1973) and Singh coeficiente to select morphoagronomic descriptors for *Manihot esculenta* Crantz and yellow passion fruit, respectively.

Based on the simultaneous analyses by both procedures, four characters were coincidente in relation to the discharge: east-west fruit (EWF), weight of five ripe fruits (RFW), leaf blade width (LBW) and total soluble solidds (TSS) (Table 7). Although the descriptor total soluble solids (TSS) had been appointed for discharge by both the Singh (1981) and Jolliffe Methods (1972, 1973), it should be included in the group of descriptors selected for not presenting significant correlations with the other descriptors selected (Table 8). The use of this descriptor is highly relevant to the characterization of accessions in which the fruits are used in the agroindustry, since physalis fruit constitute a product that might obtain aggregated value. Therefore, the descriptors that were included in the final discharge were: east-west fruits (EWF), weight of five ripe fruits (RFW) and leaf blade width (LBW) (Table 7). The combination of information obtained from both methods indicated that 25% of the quantitative descriptors was eliminated, this percentage was higher than the 21% found for *Euterpe oleracea* Mart. (Oliveira, Ferreira, & Santos, 2006) and lower than the 40% found in *Carica papaya* L. (Oliveira, Dias, & Dantas, 2012).

Table 7. Pre-selected and selected variables based on the Singh (1981) and Jolliffe Methods (1972). UEFS, Feira de Santana, BA, 2018

Variables	Pre-selected		Selected
	Jolliffe (1972)	Singh (1981)	
PH	Selec.	Selec.	Selected
SD	Disc. (3)	Selec.	Selected
NSF	Disc. (4)	Selec.	Selected
EWF	Disc. (1)	Disc. (2)	Discarded
RFW	Disc. (6)	Disc. (1)	Discarded
NFP	Disc. (5)	Selec.	Selected
LBL	Disc. (7)	Selec.	Selected
LBW	Disc. (8)	Disc. (4)	Discarded
EL	Disc. (2)	Selec.	Selected
FLL	Selec.	Disc. (5)	Selected
FTL	Disc. (10)	Selec.	Selected
TSS	Disc. (9)	Disc. (3)	Selected

Note. Plant height (PH); Stem diameter (SD); North-south fruits (NSF); East-west fruits (EWF); Weight of five ripe fruits (RFW); Number of fruits per plant (NFP); Leaf blade length (LBL); Leaf blade width (LBW); Entrenode length (EL); Fruit longitudinal length (FLL); Fruit transversal length (FTL); Total soluble solids (TSS). Discarded (Disc.); Selected (Sel.).

The correlations established were seen to be significant and positive for most of the descriptors investigated and demonstrated that the discharge of the appointed characteristics did not result in significant loss of information when they were not used in subsequent evaluations carried out in the physalis collection, since the characteristic indicated to be discharged (east-west fruits, weight of five ripe fruits and leaf blade width) were correlated to at least one of the selected descriptors (Table 8).

Similar phenomenon was observed by Oliveira et al. (2006), where four relevant characters (fruit weight, fruit widht, fruit mass per bunch and fruit yield) were discarded for the evaluation and selection of *Euterpe oleracea* Mart., with very little loss of information, since the discarded characteristics were strongly associated to the remaining characteristics investigated.

A positive and highly significant correlation was observed, with strong magnitude between the characters east-west fruits and plant height ($r = 0.70**$), stem diameter ($r = 0.75**$), north-south fruits ($r = 0.85**$), number of fruits per plant ($r = 0.84**$), leaf blade length ($r = 0.72**$) and moderate magnitude with fruit transversal length ($r = 0.67**$). The character east-west fruits also presented positive, significant correlation of moderate magnitude with the fruit longitudinal length ($r = 0.55*$) (Table 8).

Positive and significant correlation was also found, however, with moderate magnitude between the characters weight of five ripe fruits and the fruit transversal length ($r = 0.55*$). The character leaf blade width was observed to present positive and highly significant correlation, with strong magnitude with plant height ($r = 0.80**$), stem diameter ($r = 0.83**$), number of fruits per plant ($r = 0.79**$), leaf blade length ($r = 0.89**$) and moderate

magnitude with north-south fruits (0.68**) and entrenode length (0.67**). The character leaf blade width also presented positive, significant and moderate correlation with fruit transversal length (r = 0.54*) (Table 8).

Table 8. Estimates of the Pearson correlation coefficients, with their respective significances, between the 12 variables. UEFS, Feira de Santana, BA, 2018

Variables	PH	SD	NSF	EWF	RFW	NFP	LBL	LBW	EL	FLL	FTL
SD	0.90**										
NSF	0.76**	0.71**									
EWF	0.70**	0.75**	0.85**								
RFW	0.21^{ns}	0.30^{ns}	0.35^{ns}	0.50*							
NFP	0.86**	0.94**	0.79**	0.84**	0.38^{ns}						
LBL	0.90**	0.85**	0.79**	0.72**	0.36^{ns}	0.79**					
LBW	0.80**	0.83**	0.68**	0.63**	0.47*	0.79**	0.89**				
EL	0.63**	0.64**	0.42^{ns}	0.30^{ns}	0.33^{ns}	0.62**	0.49*	0.67**			
FLL	0.40^{ns}	0.42^{ns}	0.34^{ns}	0.55*	0.42^{ns}	0.56*	0.48*	0.44^{ns}	0.06^{ns}		
FTL	0.47*	0.53*	0.42^{ns}	0.67**	0.55*	0.67**	0.51*	0.54*	0.20^{ns}	0.96**	
TSS	-0.25^{ns}	-0.12^{ns}	0.01^{ns}	0.07^{ns}	0.55*	-0.01^{ns}	-0.15^{ns}	-0.07^{ns}	0.17^{ns}	0.06^{ns}	0.12^{ns}

Note. ns = not significant; *significant at 5% probability level (p < 0,05); and **significant at 1% probability level (p < 0,01), through the t test. Plant height (PH), Stem diameter (SD), North-south fruits (NSF); East-west fruits (EWF), Weight of five ripe fruits (RFW), Number of fruits per plant (NFP), Leaf blade lenght (LBL), Leaf blade width (LBW), Entrenode length (EL), Fruit longitudinal length (FLL) and Fruit transversal length (FTL).

Table 9 presents the qualitative descriptors, their phenotypical classes, percentage frequency of the accessions in each of the classes and the Renyi entropy level (Renyi, 1961). Any variable that presented a 0.00 entropy level was considered as discarded.

Table 9. Qualitative descriptors evaluated, phenotypical categories (classes), percentage frequency and entropy level of the *P. angulata* accessions under study. UEFS, Feira de Santana, 2018

Descriptors	Acronym	Classes	Frequency (%)	Level of entropy (H)
Growth habit	GH	(1) Erect	50.00	
		(2) Semi-erect	50.00	0.69
		(3) Prostrated	0.00	
Stem Color	SC	(1) Green	50.00	0.69
		(2) Crimson	50.00	
Stem Pubescence	SP	(1) Absent	100.00	
		(2) Weak	0.00	0.00
		(3) Average	0.00	
		(4) Abounding	0.00	
Stem anthocyanin	SA	(1) Absent	0.00	0.00
		(2) Present	100.00	
Leaf blade shape	LBS	(1) Lanceolate	100.00	
		(2) Cordate	0.00	0.00
		(3) Asymmetric	0.00	
Leaf apex shape	LAS	(1) Acute	0.00	
		(2) Apiculate	100.00	0.00
		(3) Acuminate	0.00	
Leaf margin shape	LMS	(1) Serrate	83.33	
		(2) Ondulate	16.67	
		(3) Erose	0.00	0.45
		(4) Sinuate	0.00	
		(5) Repando	0.00	

Leaf base shape	LBS	(1) Inequilateral	100.00	0.00
		(2) Cuneate	0.00	
		(3) Obliquous	0.00	
		(4) Cordate	0.00	
		(5) Subcordate	0.00	
Leaf bundle pubescence	LBP	(1) Absent	100.00	0.00
		(2) Weak	0.00	
		(3) Average	0.00	
		(4) Abounding	0.00	
Pubescence around the leaves	PAL	(1) Absent	100.00	0.00
		(2) Weak	0.00	
		(3) Average	0.00	
		(4) Abounding	0.00	
Leaf vein anthocyanin	LVA	(1) Absent	0.00	0.00
		(2) Present	100.00	
Flower position	Pflow	(1) Erect	0.00	0.00
		(2) Pendulous	100.00	
Corolla color	Ccor	(1) RHS8C	100.00	0.00
		(2) Yellow 5C	0.00	
		(3) Yellow 5B	0.00	
		(4) Yellow 6B	0.00	
Color of the corolla stains	CCS	(1) Coffee	0.00	0.00
		(2) Brown	100.00	
Color of the unripe calyx	CUC	(1) Green without anthocyanin	0.00	0.45
		(2) Green with slight anthocyanin	83.33	
		(3) Green with dark anthocyanin	16.67	
Pedicel color	PC	(1) Verde	0.00	0.00
		(2) Crimson	100.00	
Fruit shape	FS	(1) Round	0.00	0.64
		(2) Slightly flat	66.67	
		(3) Flat	33.33	
		(4) Heart-shaped	0.00	
Unripe fruit color	UFC	(1) RHS 144A	33.33	0.64
		(2) RHS 145B	66.67	
		(3) Yellow green 152C	0.00	
		(4) Yellow orange 23A	0.00	
Ripe fruit color	RFC	(1) RHS 160B	100.00	0.00
		(2) RHS 145C	0.00	
		(3) Purple	0.00	
		(4) Green-purple	0.00	
Calyx type	Tcal	(1) Accrescent	100.00	0.00
		(2) Maccrescent	0.00	
Calyx shape	Scal	(1) Globular	100.00	0.00
		(2) Enlongated	0.00	
		(3) Slightly flat	0.00	
		(4) Flat	0.00	
Calyx division	Dcal	(1) Present	100.00	0.00
		(2) Absent	0.00	

Growth habit (0.69), stem color (0.69), leaf margin shape (0.45), unripe calyx color (0.45), fruit shape (0.64) and unripe fruit color (0.64) presented the highest entropy values, indicating genetic variability to these characteristics (Table 9).

The descriptors initially discarded for not being able to differ the accessions are presented in the entropy analysis with a 100% frequency, that is, those accessions were concentrated in the same class of the relevant descriptors, presenting entropy level equal zero, indicating that they were monomorphic for each characteristic under evaluation. The descriptors discarded were: Stem pubescence (SP), stem anthocyanin (SA), leaf blade shape (LBS), leaf apex shape (LAS), leaf base shape (LBS), leaf bundle pubescence (LBP), pubescence around the leaves (PAL), leaf vein anthocyanin (LVA), flower position (Pflow), corolla color (Ccor), color of the corolla stains (CCS), pedicel color (PC), ripe fruit color (RFC), calyx type (Tcal), calyx shape (Scal) and calyx division (Dcal) (Table 9).

Different results were found by Padilha, Sosinski, and Barbieri (2016), when evaluating genetic divergence and the entropy of descriptors in twenty-one pepper accessions from the Germplasm Active Bank at the Embrapa Clima Temperado (Temperate Climate Embrapa), verifying that the descriptors with the highest entropy values were flower position (1.48), stigma exacerbation (1.47), ripe fruit color (1.34), pungency (1.30), plant height (1.29), node anthocyanin (1.27), fruit shape in the pedicel (1.26), persistence between pedicel and fruits (1.24) and fruit shape (1.21); while, stem shape, leaf shape, length of placenta and seed color presented entropy values equal zero (0.00). These differences in entropy values of the descriptors might have been caused by the smaller number of accessions used in this study. Therefore, one may infer that although some characteristics present genetic variability, high entropy values were not found in relation to the number of accessions used in the research.

The following descriptors were selected: Growth habit (GH), stem color (SC), leaf margin shape (LMS), unripe calyx color (UCC), fruit shape (FS), unripe fruit color (UFC). Among the 22 qualitative descriptors proposed by the National University of Colombia (González, Torres, Cano, Arias, & Arboledo, 2008), with adjustments for the species (Table 1), only six descriptors were seen to be relevant to the discrimination of genetic divergence between the accessions studied (Table 9).

Further studies with a larger number of accessions and using different methodologies (Oliveira et al. 2012; Castro et al., 2012; Oliveira et al., 2014) to discharge qualitative descriptors should be carried out to confirm the discharge of the qualitative descriptors selected in this work.

4.Conclusion

The list of minimum descriptors to characterize *P. angulata* accessions might contain nine quantitative descriptors (plant height, stem diameter, north-south fruits, number of fruits per plant, leaf blade length, entrenode length, fruit longitudinal length, fruit transversal length and total soluble solids) and six qualitative descriptors (growth habit, stem color, leaf margin shape, unripe calyx color, fruit shape and unripe fruit color).

The results of this work led to the conclusion that discharging 55.88% of the descriptors does not result in significant loss of information, minimizes the costs and makes the characterization of *P. angulata* germplasm more dynamic.

References

Afonso, S. D. J., Ledo, C. A. da S., Moreira, R. F. C., Silva, S. de O. E., Leal, V. D. de J., & Conceição, A. L. da S. (2014). Selection of descriptors in a morphological characteristics considered in cassava acessions by means of multivariate techiniques. *Journal of Agriculture and Veterinary Science, 8*(1), 13-20. https://doi.org/10.9790/2380-07151320

Alvares, R. N. (2011). *Divergência genética entre acessos Capscicum chinense Jacq. coletados no sudoeste Goiano* (Dissertação de mestrado, Universidade Federal de Goiás, Jataí, Goiás). Retrieved from https://repositorio.bc.ufg.br/tede/handle/tde/369

Alves, R. M., Garcia, A. A. F., Cruz, E. D., & Figueira, A. (2003). Seleção de descritores botânico-agronômicos para caracterização de germoplasma de cupuaçuzeiro. *Pesquisa Agropecuária Brasileira, 38*(7), 807-818. https://doi.org/10.1590/S0100-204X2003000700004

Araujo, F. L. (2012). *Estudo genético e citogenético do gênero Physalis* (Unpublished Doctoral Dissertation, Universidade Estadual de Feira de Santana, Feira de Santana, Bahia).

Castro, J. A., Neves, C. G., Jesus, O. N. de, & Oliveira, E. J. de. (2012). Definition of morpho-agronomic descriptors for the characterization of yellow passion fruit. *Scientia Horticulturae, 145*, 17-22. https://doi.org/10.1016/j.scienta.2012.07.022

Chagas, K., Alexandre, R. S., Schimildt, E. R., Bruckner, C. H., & Faleiro, F. G. (2016). Divergência genética em genótipos de maracujazeiro azedo, com base em características físicas e químicas dos frutos. *Revista Ciência Agronômica, 47*(3), 524-531. https://doi.org/10.5935/1806-6690.20160063

Cruz, C. D., Carneiro, P. C. S., & Regazzi, A. J. (2004). *Modelos biométricos aplicados ao melhoramento genético* (3rd ed.). Viçosa, MG: UFV.

Cruz, C. D. (2013). GENES—A software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum, 35*(3), 271-276.

Daher, R. F., Moraes, C. F. de, Cruz, C. D., Pereira, A. V., & Xavier, D. F. (1997). Seleção de caracteres morfológicos discriminantes em capim-elefante (*Pennisetum purpureum* Schum.). *Revista Brasileira de Zootecnia, 26*(2), 247-253.

Dantas, A. C. de A., Nunes, G. H. de S., Araújo, I. S., & Albuquerque, L. B. (2012). Caracterização molecular de acessos de melão coletados no Nordeste Brasileiro. *Revista Brasileira de Fruticultura, 34*(1), 183-189. https://doi.org/10.1590/S0100-29452012000100025

Dias, R. dos S., Souza, J. L. L. de, Oliveira, A. M., & Santos, R. L. (2010). *A relação entre a morfogenese e pedogenese do munícipio de Feira de Santana (BA)*. Apresentado no VIII Simpósio Nacional de Geomorfologia, Recife, PE. Retrieved from http://lsie.unb.br/ugb/sinageo/8/7/1.pdf

Gonçalves, L. S. A., Rodrigues, R., Amaral Júnior, A. T. do, Karasawa, M., & Sudré, C. P. (2009). Heirloom tomato gene bank: Assessing genetic divergence based on morphological, agronomic and molecular data using a Ward-modified location model. *Genetics and Molecular Research, 8*(1), 364-374. https://doi.org/10.4238/vol8-1gmr549

González, O. T., Torres, J. M. C., Cano, C. I. M., Arias, M. L., & Arboleda, A. A. N. (2008). Caracterización morfológica de cuarenta y seis accesiones de uchuva (*Physalis peruviana* L.), en Antioquia (Colombia). *Revista Brasileira de Fruticultura, 30*(3), 708-715. https://doi.org/10.1590/S0100-29452008000300025

Hair Junior, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Análise multivariada de dados* (6th ed.). Porto Alegre, RS: Bookman.

Jolliffe, I. T. (1972). Discarding variables in a principal component analysis. I: Artificial data.*Journal of the Royal Statistical Society Series C: Applied Statistics*, 21(2), 160-173. https://doi.org/10.2307/2346488

Jolliffe, I. T. (1973). Discarding variables in a principal component analysis. II: Real data. *Journal of the Royal Statistical Society Series C: Applied Statistics, 22*(1), 21-31. https://doi.org/10.2307/2346300

INMET (Instituto Nacional de Meteorologia). (2017). *In Meteorological database for teaching and research*. Retrieved November 22, 2017, from http://www.inmet.gov.br/portal/index.php?r=bdmep/bdmep

Lorenzi, H. E., & Matos, M. (2008). *Plantas medicinais no Brasil: Nativas e exóticas cultivadas* (2nd ed.). Nova Odessa, SP: Plantarum.

Miguel, L. C. V. (2017). *Divergência genética em coleção didática de batata-doce por descritores morfológicos* (Dissertação de mestrado, Universidade Federal Rural do Semi-árido, Mossoró, Rio Grande do Norte). Retrieved from http://bdtd.ufersa.edu.br/handle/tede/757

Moreno, A. M. H., Fischer, G., & Sánchez, M. I. C. (2012). Agronomical evaluation of cape gooseberries (*Physalis peruviana* L.) from central and north-eastern Colombia. *Agronomía Colombiana, 30*(1), 15-24.

Moreira, S. O. (2012). *Caracterização morfológica e molecular de pré-cultivares de Capsicum annuum L. com resistência à mancha-bacteriana* (Dissertação de mestrado, Universidade Estadual Norte Fluminense "Darcy Ribeiro", Campos Goytacazes, Rio de Janeiro). Retrieved from http://uenf.br/postgraduacao/gmp/wp-content/uploads/sites/6/2012/06/Tese-Sarah-Ola-Moreira.pdf

Oliveira, M. do S. P. de, Ferreira, D. F., & Santos, J. B. dos. (2006). Seleção de descritores para caracterização de germoplasma de açaizeiro para produção de frutos. *Pesquisa Agropecuária Brasileira, 41*(7), 1133-1140. https://doi.org/10.1590/S0100-204X2006000700009

Oliveira, J. A. R. de, Martins, L. H. da S., Vasconcelos, M. A. M. de, Pena, R. da S., & Carvalho, A. V. (2011). Caracterização física, físico-química e potencial tecnológico de frutos de camapu (*Physalis angulata* L.). *Revista Brasileira de Tecnologia Agroindustrial, 5*(2), 573-583. https://doi.org/10.3895/S1981-36862011000200009

Oliveira, E. J. de, Dias, N. L. P., & Dantas, J. L. L. (2012). Selection of morpho-agronomic descriptors for characterization of papaya cultivars. *Euphytica, 185*, 253-265. https://doi.org/10.1007/s10681-011-0565-0

Oliveira, E. J. de, Oliveira Filho, O. S. de, & Santos, V. da S. (2014). Selection of the most informative morphoagrnomic descriptors for cassava germplasm. *Pesquisa Agropecuária Brasileira, 49*(11), 891-900. https://doi.org/10.1590/S0100-204X2014001100008

Ortiz, R., Flor, F. D. de L., Alvarado, G., & Crossa, J. (2010). Classifying vegetable genetic resources—A case study with domesticated *Capsicum* spp. *Scientia Horticulturae, 126*, 186-191. https://doi.org/10.1016/j.scienta.2010.07.007

Padilha, H. K. M., Sosinki, E. E., & Barbieri, R. L. (2016). Morphological diversity and entropy of peppers (*Capsicum baccatum* and *Capscicum chinense*, Solanaceae). *Internacional Journal of Current Research, 8*, 42758-42766.

Pereira, A. V., Vencovsky, R., & Cruz, C. D. (1992). Selection of botanical and agronomical descriptors for the characterization of cassava (*Manihot esculenta* Crantz.) germplasm. *Revista Brasileira de Genética, 15*(1), 115-124.

Rêgo, E. R. do, Rego, M. M. do, Cruz, C. D., & Cecon, P. R. (2003). Genetic diversity analysis of peppers: A comparison of discarding variable methods. *Crop Breeding and Applied Biotechnology, 3*(1), 19-26. https://doi.org/10.12702/1984-7033.v03n01a03

Renyi, A. (1961). On Measures of Entropy and Information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1*, 547-561

Restrepo, E. F. S., & Vallejo, F. A. C. (2003). Diversidadgenética del tomate cultivado tipo "chonto" *Lycopersicon esculentum* Mill, en las zonas productoras de Colombia. *Acta Agronómica, 52*(1-4), 11-17.

R-Team. (2013). *R: A language and environment for statistical computing R Foundation for Statistical Computing*. Retrieved from http://www.R-project.org

Silva, A. H. B. (2007). *Seleção e variabilidade genética para caracteres qualitativos e quantitativos em progênie de Physalis angulata L. (Solanaceae)* (Dissertação de Mestrado, Universidade Estadual de Feira de Santana, Feira de Santana, Bahia). Retrieved from https://www.ppgbot.uefs.br/.../selecao-e-variabilidade-genetica-para-caracteres-qualitativos-e-quantitativos-em-progenies-de-physalis-angulata-l-solanaceae-%20(3).pdf

Silva, W. C. J. E., Carvalho, S. I. C. de, & Duarte, J. B. (2013). Identification of minimum descriptors for characterization of *Capsicum* spp. germplasm. *Horticultura Brasileira, 31*(2), 190-202. https://doi.org/10.1590/S0102-05362013000200004

Silva, R. S., Moura, E. F., Neto-Farias, J. T., Ledo, C. A. S., & Sampaio, J. E. (2017). Selection of morphoagronomic descriptors for the characterization of accessions of cassava of the Eastern Brazilian Amazon.*Genetics and Molecular Research, 16*(2), 1-11. https://doi.org/10.4238/gmr16029595

Silva, P. B. da, Silva, S. de M., Silva, J. A. da, Mendonça, R. M. N., & Pereira, W. E. (2018). Changes in quality during maturation of physalis fruit. *Semina: Ciências Agrárias, 39*(4), 1503-1516. https://doi.org/10.5433/1679-0359.2018v39n4p1503

Singh, D. (1981). The relative importance of characters affecting genetic divergence. *The Indian Journal of Genetic and Plant Breeding, 41*(1), 237-245.

Sudré, C. P., Rodrigues, R., Riva, E. M., Karasawa, M., & Amaral Junior, A. T. do. (2005). Divergência genética entre acessos de pimenta e pimentão utilizando técnicas multivariadas. *Horticultura Brasileira, 23*(1), 22-27. https://doi.org/10.1590/S0102-05362005000100005

Sudré, C. P., Gonçalves, L. S. A., Rodrigues, R., Amaral Júnior, A. T. do, Riva-Souza, E. M., & Bento, C. dos S. (2010). Genetic variability in domesticated *Capsicum* spp. as assessed by morphological and agronomic data in mixed statistical analysis. *Genetics and Molecular Research, 9*(1), 283-294. https://doi.org/10.4238/vol9-1gmr698

Sultana, N., Hassan, M. A., Begum, M., & Sultana, M. (2008). *Physalis angulata* L. (Solanaceae)- A new angiospermic record for Bangladesh. *Bangladesh Journal of Botany, 37*(2), 195-198. https://doi.org/10.3329/bjb.v37i2.1731

The Royal Horticultural Society. (2001). *RHS colour chart*. London, UK: The Royal Horticultural Society.

Thornthwaite, C. W., & Mather, J. R. (1955). *The water balance*. Centerton, NJ: Drexel Institute of Technology, Laboratory of Climatology.

Tomassini, T. C. B., Barbi, N. S., Rbeiro, I. M., & Xavier, D. C. D. (2000). Gênero *Physalis*-uma revisão sobre vitaesteróides. *Química Nova, 23*(1), 47-57. https://doi.org/10.1590/S0100-40422000000100011

Veiga, R. F. A., & Queiróz, M. A. (2015). *Recursos fitogenéticos: A base da agricultura sustentável no Brasil* (1st ed.). Brasília, DF: Sociedade Brasileira de Recursos Genéticos.

Vieira, E. A., Fialho, J. F., & Faleiro, F. G. (2007). Variabilidade genética do banco ativo de germoplasma de mandioca do cerrado acessada por meio de descritores morfológicos. *Boletim de Pesquisa e Desenvolvimento, 179*, 5-28.

Appendix A

Soil chemical analysis from the Experimental Horticultural Unit of the UEFS. UEFS, Feira de Santana, BA, 2018

pH	P	K	Ca	Mg	Ca+Mg	Al	Na	H+ Al	SB	CTC	V
In water	mg/dm^3	--------------------------------------- Cmol$_c$/dm^3 ---------------------------------------									%
6.1	95	0.15	2.80	1.05	3.85	0.0	0.05	1.98	4.04	6.02	67

Note. Analysis carried out at the Laboratory of Soils and Nutrition of Plants at Embrapa Mandioca and Fruticultura. Cruz das Almas-BA.

Copyrights