

<http://dx.doi.org/10.15361/1984-5529.2019v47n1p83-90>

## Genetic divergence between pepper accessions based on quantitative fruit traits

### Divergência genética entre acessos de pimentas com base em caracteres quantitativos dos frutos

Ana Carolina Ribeiro de OLIVEIRA<sup>1</sup>; Paulo Roberto CECON<sup>2</sup>; Moysés NASCIMENTO<sup>3</sup>; Fernando Luiz FINGER<sup>4</sup>; Giselda Maria PEREIRA<sup>5</sup>; Guilherme Alves PUIATTI<sup>6</sup>

<sup>1</sup> Autor para correspondência, Mestra em Estatística Aplicada e Biometria, Universidade Federal de Viçosa, Departamento de Estatística, Avenida Peter Henry Rolfs, s/n - Campus Universitário, Viçosa - MG, 36570-900, [ana.oliveira8@ufv.br](mailto:ana.oliveira8@ufv.br)

<sup>2</sup> Doutor em Estatística, Universidade Federal de Viçosa, Departamento de Estatística, [cecon@ufv.br](mailto:cecon@ufv.br)

<sup>3</sup> Doutor em Estatística e Experimentação Agropecuária, Universidade Federal de Viçosa, Departamento de Estatística, [moysesnascim@ufv.br](mailto:moysesnascim@ufv.br)

<sup>4</sup> Doutor em Horticultura, Universidade Federal de Viçosa, Departamento de Fitotecnia, [ffinger@ufv.br](mailto:ffinger@ufv.br)

<sup>5</sup> Doutora em Genética e Melhoramento, Universidade Federal de Pelotas, Departamento de Matemática e Estatística, [gmpereira08@gmail.com](mailto:gmpereira08@gmail.com)

<sup>6</sup> Doutor em Estatística Aplicada e Biometria, Universidade Federal de Viçosa, Departamento de Estatística, [guilherme@dpi.ufv.br](mailto:guilherme@dpi.ufv.br)

Recebido em: 03-07-2018; Aceito em: 13-12-2018

#### Abstract

The development of new *Capsicum* cultivars aiming to meet market requirements will depend, above all, on the genetic diversity of the study population. To quantify this genetic divergence, several multivariate techniques assessing quantitative traits have been employed. This study aimed to: i. estimate the genetic diversity among *Capsicum chinense* accessions from the Active Germplasm Bank of Plants of the Federal University of Viçosa (BGH-UFV); ii. indicate promising accessions for prospective studies of specific market niches; iii. evaluate the disposal of redundant traits. The experiment was conducted in a completely randomized design with four replicates, in which 11 *C. chinense* accessions were evaluated, based on 11 quantitative fruit traits. The data were subjected to cluster analysis by the UPGMA and Tocher methods, based on the quadratic Euclidean distance, to assess diversity. Afterwards, we used principal component analysis, Jolliffe's method and procrustes analysis for the disposal of traits. The highest genetic dissimilarity was obtained between accessions 2 and 10. The phenotypic correlation coefficients obtained were 0.75 (UPGMA) and 0.91 (Tocher), the latter being significant by the Mantel test ( $p < 0.05$ ). Six clusters were formed by using the Tocher method, four of which were composed by a single accession. Regarding the disposal of variables, traits TFDW, TFFW, PUN, %DM, FW, PT, and FL were shown to be disposable, and do not affect diversity prediction in terms of graphic dispersion. Accessions 9, 10, and 11 are indicated for *in natura* consumption, while accessions 2 and 3 are indicated for industrial purposes. These accessions showed the best results among the evaluated traits for the mentioned niches.

**Additional keywords:** *Capsicum chinense* J.; multivariate analysis; procrustes; genetic variability.

#### Resumo

O desenvolvimento de novas cultivares de *Capsicum*, visando a atender às exigências do mercado, dependerá, sobretudo, da diversidade genética presente na população em estudo. Para a quantificação dessa divergência genética, a partir de caracteres quantitativos, várias técnicas multivariadas têm sido empregadas. Este estudo teve por objetivos: i. estimar a diversidade genética entre acessos de *Capsicum chinense* do Banco Ativo de Germoplasma de Hortaliças da Universidade Federal de Viçosa (BGH-UFV); ii. indicar acessos promissores para estudos prospectivos de nichos específicos de mercado; iii. avaliar o descarte de caracteres redundantes. O experimento foi conduzido em delineamento inteiramente casualizado, com quatro repetições, sendo avaliados 11 acessos de *C. chinense*, com base em 11 caracteres quantitativos do fruto. Os dados foram submetidos à análise de agrupamento pelos métodos UPGMA e Tocher, com base na distância euclidiana quadrática padronizada, para avaliar a diversidade. Após, adotaram-se os métodos de componentes principais, Jolliffe e procrustes para descarte de caracteres. A maior dissimilaridade genética foi obtida entre o par de acessos 2 e 10. Os coeficientes de correlação cofenética obtidos foram 0,75 (UPGMA) e 0,91 (Tocher), sendo o último significativo pelo teste de Mantel ( $p < 0,05$ ). Utilizando o método Tocher, houve a formação de seis grupos, sendo quatro destes compostos por um único acesso. Em relação ao descarte de variáveis, detectou-se que os caracteres MTS, MTF, PUN, %MS, LAR, ESP e CMP são passíveis de descarte e não afetam a predição da diversidade em termos de dispersão gráfica. Os acessos 9, 10 e 11 foram indicados para o consumo *in natura*, e os acessos 2 e 3 para a indústria, pois apresentaram os melhores resultados, dentre os caracteres avaliados, para os nichos mencionados.

**Palavras-chave adicionais:** análise multivariada; *Capsicum chinense* J.; procrustes; variabilidade genética.

## Introduction

Species *Capsicum chinense* Jacq. is native to the Amazon and is consumed in South America, the United States, and the Caribbean (Bosland & Votava, 2012; Moses et al., 2014). This species is characterized by high climatic adaptation in tropical and equatorial regions, and by great diversity, mainly in fruits, which have different shapes, colors, sizes and pungency levels (Lannes et al., 2007; Alvares et al., 2012). In addition, the species has great economic relevance, since its fruits are used as raw material in the food, spice, and pharmaceutical industries (Faria et al., 2013). In general, fruits of the genus *Capsicum* are important sources of three natural antioxidants: vitamin C, carotenoids, and vitamin E (Rêgo et al., 2012).

The pepper market is widely diversified, ranging from commercialization for in natura consumption and homemade preserves, to the export of industrialized products (Ferrão et al., 2011). In both niches, pepper fruit traits associated with the taste, odor, texture, and nutritional value of fruits are important attributes (Braga et al., 2013) because they are directly related to the quality of the product. Among the traits, soluble solids content plays a key role, since high contents of this constituent imply lower sugar addition, shorter water evaporation time, lower energy expenditure, and higher product yield, resulting in greater savings in the process (Faria et al., 2013).

To obtain genotypes of economic interest through a breeding program, a broad characterization of the genetic diversity of available resources is required (Hill et al., 2013). This will lead to hybrid combinations of superior strains, with greater heterotrophic effect (Pessoa et al., 2018). Thus, using this diversity through genetic combinations will allow the emergence of genotypes that are more adapted, productive, and resistant to diseases, among other traits of interest to the market (Souza et al., 2011). Furthermore, it will be possible to differentiate accessions and make them more useful (Padilha et al., 2016a), identifying effective traits in several processes, such as growth regulation, biomass, and fruit production (Souza et al., 2015).

For the study of genetic diversity, several

methods are used. Multivariate analysis, which encompasses different methods, is represented by cluster analysis, principal component analysis, canonical variable analysis, and others. These analyses allow to predict genetic diversity for several traits simultaneously, consequently leading to a more extensive and efficient evaluation for the identification of divergent individuals, aiming at the orientation of crosses (Cruz et al., 2012).

Although there are a lot of divergence studies with *Capsicum chinense* Jacq. described in the literature, such as Finger et al. (2010), Faria et al. (2012, 2013), Vasconcelos et al. (2014), Souza et al. (2015), Padilha et al. (2016b) and Pessoa et al. (2018), which only evaluate or characterize genetic diversity for the choice of superior parents, none of them evaluates genetic diversity aiming at recommendations for specific market niches. Moreover, studies that aim to observe traits with redundant information for the prediction of genetic diversity are of interest, since it can lead to cost reduction in new experiments.

In view of the above, this study aimed to: i.- estimate the genetic diversity between *Capsicum chinense* accessions from the Active Germplasm Bank of Plants of the Federal University of Viçosa (BGH-UFV); ii.- indicate promising accessions for prospective studies of specific market niches; iii. evaluate the disposal of redundant traits.

## Material and methods

The experiment was conducted in a completely randomized experimental design with four replicates and one plant per plot, in the experimental area of the Olericulture Sector of the Department of Plant Sciences of the Federal University of Viçosa (UFV), Viçosa city, located in the Zona da Mata of Minas Gerais State, Brazil, whose geographic coordinates are: 20° 45'S and 42° 51' W, and 650 m altitude.

Eleven *Capsicum chinense* Jacq. accessions from the Active Germplasm Bank of Plants of the Federal University of Viçosa (BGH-UFV) (Table 1) were evaluated for 11 quantitative fruit traits.

**Table 1** - Identification of the 11 pepper accessions (*Capsicum chinense* J.).

Pepper accessions	Name	Origin
1	BGH 1716-17	Pindaré - Mirim – MA
2	BGH 4289-44	Rondonópolis – MT
3	BGH 4733-54	Manaus – AM
4	BGH 5012-72	Água Branco – AL
5	BGH 6228-79	Brasília – DF
6	BGH 7295-101	Viçosa – MG
7	BGH 1716-16	Pindaré - Mirim – MA
8	BGH 4201-32	Belém – PA
9	BGH 4223-39	IAC – SP
10	BGH 6233-85	Brasília – DF
11	BGH 6378-98	Boca do Janacanam – AM

The following traits were evaluated: TFW: total fruit weight per plant (g); TFFW: total fruit fresh weight (g); TFDW: total fruit dry weight (g); %DM: percent dry matter ( $g\ 100g^{-1}$ ); FL: fruit length (mm); FW: fruit width (mm); PT: pericarp thickness (mm); TSS: total soluble solids ( $^{\circ}$ Brix); PUN: pungency level ( $mg\ g\ MS^{-1}$ ); VIT C: vitamin C content ( $mg\ 100g^{-1}$  of fresh fruit), and NSF: number of seeds per fruit.

TFW was obtained through three monthly harvests, in the period of higher plant yield. The other traits were obtained by the average of five mature fruits per plant.

To estimate the genetic divergence among accessions, the quadratic Euclidean distance ( $d^2$ ) was used as dissimilarity measure, according to Equation 1:

$$d_{ii'}^2 = \|y_i - y_{i'}\|_{\Psi}^2 \quad (1)$$

in which:  $d_{ii'}^2$  is the quadratic distance between accessions  $i$  and  $i'$ , where  $i, i' = 1, 2, \dots, 11$ ;  $y_i$  and  $y_{i'}$  are the  $p$ -dimensional vectors of the means of the observations of accessions  $i$  and  $i'$  (in cases of repetition);  $\Psi$  is denoted as metric, represented by the inverse diagonal matrix of residual variances and covariances of the sample of 11 accessions. This measurement promotes an increase in the weight of the most distant accessions, highlighting the difference between clusters. Moreover, the data are standardized by using the scale in terms of standard deviation, being indicated for variables that have different scales (Ferreira, 2011; Cruz et al., 2012). Standardization will allow the traits to contribute equally in the evaluation of the similarity between accessions, that is, preventing the variables with great dispersion to dominate the classification of the distances (Cruz et al., 2011).

Subsequently, to characterize the accessions, we adopted the UPGMA (average linkage between groups) and Tocher methods.

For the UPGMA method, the determination of the optimal number of clusters in the dendrogram adopted the Mojena (1977) procedure, which consists in selecting the number of clusters in step  $j$  that satisfies:  $\alpha_j > \theta_k$ , in which  $\alpha_j$  is the distance value of the fusion level corresponding to steps  $j$  ( $j = 1, 2, \dots, g - 1$ ), and  $\theta_k$  is the cutoff value, according to Equation 2:

$$\theta_k = \bar{\alpha} + k\hat{\sigma}_{\alpha} \quad (2)$$

in which:  $\bar{\alpha}$  is the mean value of  $\alpha$ ,  $\hat{\sigma}_{\alpha}$  is the standard deviation of  $\alpha$ ; and  $k = 1, 2, 5$ , according to Milligan & Cooper (1985).

To validate the consistency of the methods, the phenotypic correlation coefficient was applied, which measures the degree of fit between the dissimilarity matrix and the phenotypic matrix, that is, the preservation of the distances resulting from the clustering in relation to the original distances (Sneath & Sokal, 1973). For the Tocher method, the phenotypic

matrix was obtained by the method of Silva & Dias (2013).

The efficiency of the classification of methods was evaluated by the Mantel (1967) test, which tests the hypothesis of a null correlation between the phenotypic matrix and the original distance matrix, with 1,000 permutations, considering the 5% level of significance; and by Fisher's discriminant analysis (1936), based on the results of each clustering, considering the apparent error rate (AER) (Equation 3):

$$AER = \frac{1}{n} \sum_{i=1}^g e_i \quad (3)$$

in which:  $n$  is the total number of classification;  $e_i$  is the number of wrong classifications in each cluster ( $i = 1, \dots, g$ ).

For the disposal of traits, principal component analysis was used in association with the direct selection method proposed by Jolliffe (1972, 1973). This selection method allows the elimination of traits with higher coefficient (eigenvectors) for the last components (eigen values lower than 0.7). Finally, the disposal effect was evaluated directly in the graphic configuration of the genetic diversity, through procrustes analysis (Krzanowski, 1987). This analysis can be used to determine how much the new subset of variables (which can be represented by a subset of the original variables) represents the structure of the original data (Guedes & Ivanqui, 1998). The similarity between configurations can be evaluated by means of the R coefficient (Krzanowski, 1987), in which values close to 1 (one) indicate a great similarity between the evaluated configurations.

The disposal of redundant variables, that is, of low informative content, reduces the effort of data collection without restricting the accuracy of the morphological characterization and, consequently, of the genetic diversity analysis (Galate et al., 2012).

All analyses were performed using the R software (R Development Core Team, 2018).

## Results and discussion

According to Table 2, there was greater genetic dissimilarity between accessions 2 and 10 ( $d^2 = 72.16$ ). In contrast, accessions 1 and 5 are the most genetically similar ( $d^2 = 1.78$ ).

The clustering analysis by the UPGMA method resulted in two clusters, cluster I (CI), composed of eight accessions (72.73%); and cluster II (CII), with three accessions (27.27%). The Tocher method, in turn, divided the accessions into six clusters: CI, composed of five accessions (45.4%); CII, with two accessions (18.2%); and CIII to CVI, with only one accession (9.09%) each (Table 3).

**Table 2** - Estimates of dissimilarity measures between 11 pepper accessions based on standardized quadratic euclidean distance.

	2	3	4	5	6	7	8	9	10	11
1*	8.79	10.58	2.23	1.78	8.46	8.44	7.36	24.57	41.89	27.82
2		34.11	6.40	15.13	13.13	21.04	23.35	48.08	72.16	47.13
3			16.84	6.67	24.11	12.89	9.55	13.42	23.90	24.72
4				4.16	6.92	11.84	9.53	34.65	52.59	35.65
5					11.43	6.09	4.21	19.63	36.45	26.31
6						25.07	22.02	50.08	68.54	52.11
7							10.92	11.07	26.70	20.66
8								21.79	33.93	20.22
9									10.40	11.04
10										11.18

\*(1): BGH 1716-17; (2): BGH 4289-44; (3) BGH 4733-54; (4) BGH 5012-72; (5) BGH 6228-79; (6) BGH 7295-101; (7) BGH 1716-16; (8) BGH 4201-32; (9) BGH 4223-39; (10) BGH 6233-85; (11) BGH 6378-98.

**Table 3** - Clustering of the 11 accessions of pepper through the UPGMA and Tocher methods from the standardized quadratic euclidean distance.

Method	Group	Number of accessions	Accession
UPGMA	I	8	1 2 3 4 5 6 7 8
	II	3	9 10 11
Tocher	I	5	1 4 5 7 8
	II	2	9 10
	III	1	2
	IV	1	3
	V	1	6
	VI	1	11

The phenotypic correlation coefficients obtained for the UPGMA and Tocher methods were 0.72 and 0.91, respectively, the latter being significant by the Mantel test ( $p < 0.05$ ). Thus, it was found that the best representation of the genetic distances was obtained by using the Tocher method and, therefore, only the results referring to such methodology were discussed. In addition, Fisher's discriminant analysis applied to Tocher's clustering did not detect erroneous classifications, with the apparent error rate (AER) being zero.

In the Tocher method, CIII showed the highest values of TSS (10.4 °Brix), responsible for giving aroma and flavor to the fruit, and %DM (20.7%). CIV,

in turn, showed the highest values of VIT C (123.3 mg 100g<sup>-1</sup>) and FL (59.8 mm) (Table 4). Moura et al. (2010) obtained mean values of VIT C similar to those found in this study, and verified the viability of using germplasm per se as a source of vitamin C. These two clusters together presented the highest values of important traits for industrial purposes such as production of sauces, paprika, and preserves, as mentioned by Faria et al. (2013) and Leite et al. (2016), in addition to cosmetic and pharmaceutical uses (Leite et al., 2016). Thus, accessions belonging to clusters CIII and CIV, accessions 2 and 3, are indicated for prospective studies aimed at the industrial market.

**Table 4** - Mean performance of the 11 traits for each group obtained by the Tocher method.

Method	Group	Traits <sup>1</sup>										
		TFW	TFFW	TFDW	%DM	TSS	VIT C	FL	FW	PT	NSF	PUN
Tocher	I	958.8	3.4	0.5	15.7	10.0	107.2	34.8	15.2	1.9	42.2	2.3
	II	1520.5	11.3	1.2	10.8	7.4	100.7	53.1	29.3	2.6	56.6	0.0
	III	309.1	1.6	0.3	20.7	10.4	81.1	25.0	12.2	1.6	33.5	6.1
	IV	1344.4	5.4	0.7	13.7	8.6	123.3	59.8	16.8	1.9	48.8	4.8
	V	550.1	1.3	0.2	18.3	9.3	120.4	11.9	10.9	1.8	32.5	9.5
	VI	1386.4	11.1	1.3	12.6	8.8	91.2	44.0	27.8	3.2	33.9	0.0

<sup>1</sup> TFW: total fruit weight per plant (g); TFFW: total fruit fresh weight (g); TFDW: total fruit dry weight (g); % DM: percent dry matter (g 100g<sup>-1</sup>); TSS: soluble solids (°Brix); VIT C: vitamin C content (mg 100g<sup>-1</sup>); FL: fruit length (mm); FW: fruit width (mm); PT: pericarp thickness (mm); NSF: number of seeds per fruit; and PUN: pungency level (mg g<sup>-1</sup> MS).

Accessions belonging to clusters CV and CVI (9.5 mg gMS<sup>-1</sup>) and PT (3.2 mm), respectively (Table 4). Domenico et al. (2012) reported the existence of

environmental influence on fruit pungency and plant development, so comparisons are recommended only under similar conditions. For Paulus et al. (2015), pungency is an important commercial attribute and a quality requirement for fresh peppers and processed products. In the pharmaceutical field, *Capsicum* is used as an active principle for the production of several drugs (Domenico et al., 2012).

Based on the pungency level, accessions belonging to clusters CII and CVI are indicated for in natura consumption in the form of salad, due to the absence of capsaicin. Furthermore, they show high values of TFDW, FL, PT, and NSF, both related to firmness - a requirement that is essential for the in natura market, as mentioned by Casali et al. (1984). Therefore, accessions 9, 10 (CII), and 11 (CVI) are indicated for prospective studies aiming at in natura consumption (Table 4).

In the principal component analysis, we observed that the first two components (PC1 and PC2)

account for 79% of the total variation of the data. This result is close to that recommended by Cruz et al. (2012), that is, values around 80% are considered satisfactory. PC1 alone accounts for 66.93% of the total variation, while PC2 accounts for 12.13% (Table 5). Among the studies of genetic divergence between *Capsicum* spp. by means of principal components analysis, Bento et al. (2007) observed that about 72% of the total variation was explained by the first two components. On the other hand, Leite et al. (2011), when studying the divergence between 60 accessions of *Capsicum* spp., found that only 47% of the total variation was explained by the first two principal components.

In Table 5, of the eleven principal components, seven have a variance lower than 0.7 (eigenvalue less than 0.7). According to Jolliffe (1972, 1973), these components correspond to the number of traits that can be discarded; therefore, seven traits (63.3% of the 11 traits evaluated) can be discarded.

**Table 5** - Principal components (PC), eigenvalues ( $\lambda_i$ ), percentage of variance explained by components (%VEC) and cumulative percentage of variance explained by components (%VECa) of the evaluated traits.

PC	$\lambda_i$ *	%VEC	%VECa
1	7.36	66.93	66.93
2	1.33	12.13	79.06
3	0.85	7.69	86.75
4	0.73	6.63	93.39
5	0.32	2.88	96.27
6	0.26	2.38	98.65
7	0.06	0.57	99.22
8	0.06	0.52	99.74
9	0.03	0.25	99.99
10	0.00	0.01	100.00
11	0.00	0.00	100.00

\* Bold values indicate the eigenvalues less than 7.

Analyzing the last seven principal components, we observed that the traits indicated for disposal have the largest eigenvectors per component, namely: TFDW (-0.67), TFFW (-0.86), PUN (0.53), %DM (0.77),

FW (-0.54), PT (0.62), and FL (0.68), as shown in Table 6. Hence, optimal selection was given by the following subset of variables: TFW, TSS, VIT C, and NSF.

**Table 6** - Weighting coefficients of the last seven principal components (eigenvectors)<sup>1</sup>.

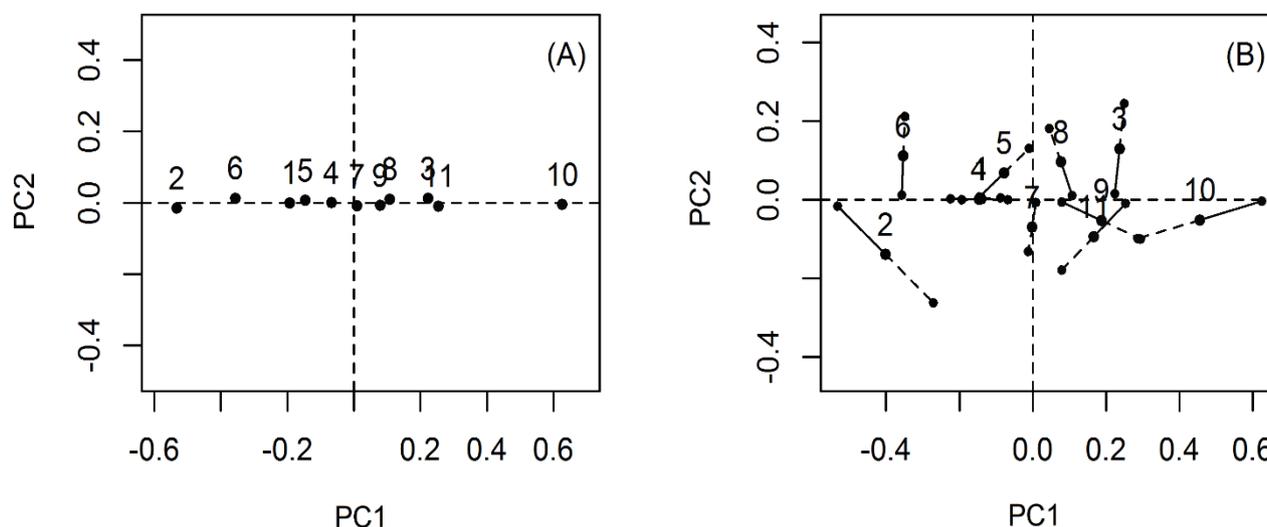
Traits <sup>2</sup>	Principal components						
	CP5	CP6	CP7	CP8	CP9	CP10	CP11
TFW	-0.36	-0.34	0.46	0.36	0.43	-0.05	0.09
TFFW	0.07	-0.00	-0.25	0.07	-0.10	<b>-0.86</b>	0.06
TFDW	0.18	0.18	-0.32	0.19	0.36	0.21	<b>-0.67</b>
%DM	0.06	-0.09	-0.37	<b>0.77</b>	0.02	0.07	0.29
TSS	-0.04	0.08	-0.13	-0.29	0.49	-0.21	0.15
VIT C	-0.13	0.38	-0.27	0.07	-0.09	0.00	0.05
FL	<b>0.68</b>	-0.42	-0.02	-0.01	-0.11	0.12	0.18
FW	-0.32	-0.25	<b>-0.54</b>	-0.31	0.07	0.33	0.36
PT	0.24	<b>0.62</b>	0.26	0.06	0.04	0.17	0.46
NSF	-0.06	0.20	-0.09	0.05	0.30	-0.04	0.18
PUN	0.40	-0.08	-0.01	-0.15	<b>0.53</b>	-0.08	0.09

<sup>1</sup>In bold, the seven eigenvectors less important to explain the total variation.

<sup>2</sup> TFW: total fruit weight per plant (g); TFFW: total fruit fresh weight (g); TFDW: total fruit dry weight (g); % DM: percent dry matter (g 100g<sup>-1</sup>); TSS: total soluble solids (°Brix); VIT C: vitamin C content (mg 100g<sup>-1</sup>); FL: fruit length (mm); FW: fruit width (mm); PT: pericarp thickness (mm); NSF: number of seeds per fruit; and PUN: pungency level (mg g<sup>-1</sup> MS).

Subsequently, to evaluate the effect of trait disposal in the prediction of genetic diversity by means of graphic dispersion of the accessions, procrustes analysis was used. First, it is observed that when considering the configuration with eleven traits versus four traits - excluding the seven traits that can be discarded by Jolliffe's method (TFDW, TFFW, PUN, %DM, FL, PT, and FL), the graphic dispersion of the accessions

was not altered, since visually only one point represents the position of the accessions (Figure 1A). Therefore, the disposal of the traits mentioned above did not influence the graphic dispersion, confirming their disposability, as observed by Jolliffe's method. In addition, the R coefficient was equal to 0.99, evidencing the high similarity between the two configurations evaluated.



**Figure 1** - Graphical dispersion of 11 pepper accessions (*Capsicum chinense* J.) considering the first and two principal components (PC1 and PC2) after the procrustes transformation, with eleven traits vs. four traits (A) and with eleven traits vs. ten traits (B).

In contrast, when analyzing the configuration with eleven traits versus ten traits - excluding TFW, which according to the literature is extremely important both for the determination of genetic diversity and for yield (Moreira et al., 2009), not being indicated for disposal by Jolliffe's method; there was a large change in the graphic dispersion of the accessions (Figure 1B). Therefore, this trait is of paramount importance in diversity analysis. The R coefficient was equal to 0.57, indicating a smaller association between the evaluated configurations. Specifically, this figure shows the existence of three distinct points for each accession, one of which represents the analysis with all eleven traits, while another considers ten traits (except TFW); moreover, between the two previously described points, a third point is given by the average between the two previous configurations.

The results of the methods for trait disposal agreed with each other, demonstrating that it would be redundant to analyze these traits in future research of genetic diversity in peppers under experimental conditions similar to those used in this study; saving time, labor, and financial resources (Yamaki et al., 2009). This magnitude of trait reduction (63.3%) was similar to that observed by Silva et al. (2013), who, when identifying minimum descriptors for the characterization of *Capsicum* spp. germplasms, obtained reductions of almost 50%.

### Conclusions

In the Tocher method, the 11 accessions were partitioned into six clusters, due to genetic similarity.

Accessions 9, 10, and 11 were indicated for in natura consumption, since they show the highest values of traits requested by the market. Moreover, accessions 2 and 3 are indicated for the industry.

Traits TFDW, TFFW, PUN, %DM, FW, PT, and FL can be discarded by Jolliffe's method, and do not affect diversity prediction in terms of graphic dispersion.

### References

Alvares RC, Reis EF, Pinto JFN (2012) Genetic divergence in pepper genotypes from southwest Goiás. *Ciência e Agrotecnologia* 36(5): 498-506. doi: 10.1590/S1413-70542012000500002

Bento CS, Sudré CP, Rodrigues R, Riva EM, Pereira MG (2007) Descritores qualitativos e multicategóricos na estimativa da variabilidade fenotípica entre acessos de pimentas. *Scientia Agraria* 8(2):149-156. doi: 10.5380/rsa.v8i2.8379

Bosland PW, Votava EJ (2012) Peppers: vegetable and spice capsicums. CABI Publishing, 230 p.

- Braga TR, Pereira RCA, Silveira MRS, Silva LR, Oliveira MMT (2013) Caracterização físico-química de progênies de pimentas (*Capsicum frutescens* L.). Revista de la Facultad de Agronomía 112(1):6-10.
- Casali WD, Pádua JG, Braz LT (1984) Melhoramento de pimentão e pimenta. Informe Agropecuário 10(113):19-22.
- Cruz CD, Ferreira FM, Pessoni LA (2011) Biometria aplicada ao estudo da diversidade genética, 1ed, Editora Suprema. 620p.
- Cruz CD, Regazzi AJ, Carneiro PCS (2012) Modelos biométricos aplicados ao melhoramento genético, 4ed, Editora UFV. 514p.
- Domenico CI, Coutinho JP, Godoy HT, Melo AMT (2012) Caracterização agrônômica e pungência em pimenta de cheiro. Horticultura Brasileira 30(3):466-472.
- Faria PN, Cecon PR, Silva AR, Finger FL, Silva FF, Cruz CD, Sávio FL (2012) Métodos de agrupamento em estudo de divergência genética de pimentas. Horticultura Brasileira 30(3):428-432.
- Faria PN, Laia GA, Cardoso KA, Finger FL, Cecon PR (2013) Estudo da variabilidade genética de amostras de pimenta (*Capsicum chinense* Jacq.) existentes num banco de germoplasma: um caso de estudo. Revista de Ciências Agrárias 36(1):17-22.
- Ferrão LFV, Cecon PR, Finger FL, Silva FF, Puiatti M (2011) Divergência genética entre genótipos de pimenta com base em caracteres morfo-agrônomicos. Horticultura Brasileira 29(3): 354-358.
- Ferreira DF (2011) Estatística multivariada, 2ed, Editora UFLA. 676p.
- Finger FL, Lannes SD, Schuelter AR, Doege J, Comerlato AP, Gonçalves LSA, Ferreira FRA, Clovis LR, Scapim CA (2010) Genetic diversity of *Capsicum chinense* accession based on molecular markers and morphological and agronomic traits. Journal Genetics and Molecular Research 9(3):1852-1864. doi: 10.4238/vol9-3gmr891
- Fisher RA (1936) The use of multiple measurements in taxonomic problems. Annals of Eugenics 7(2):179-188.
- Galate RS, Mota MGC, Gaia JMD, Costa MSS (2012) Caracterização morfoagronômica de germoplasma de açaizeiro no nordeste paraense. Revista Brasileira de Fruticultura. 34(2):540-550.
- Guedes TA, Ivanqui IL (1998) Análise procrustes aplicada à seleção de variáveis. Acta Scientiarum 20(4):505-509.
- Hill TA, Ashrafi H, Wo SRC, Yao J, Stoffel K, Truco MJ, Kozik A, Michelmore RW, Deynze AV (2013) Characterization of *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism discovery with a 30k unigene pepper genechip. Plos One 8(2):1-16. doi: 10.1371/journal.pone.0056200
- Jolliffe IT (1972) Discarding variables in a principal component analysis. I. Artificial data. Applied Statistics 21(2):160-173.
- Jolliffe IT (1973) Discarding variables in a principal component analysis. II. Real data. Applied Statistics 22(1):21-31.
- Krzanowski WJ (1987) Selection of variables to preserve multivariate data structure, using principal components. Applied Statistics 36(1):22-33.
- Lannes SD, Finger FL, Schuelter AR, Casali VWD (2007) Growth and quality of Brazilian accessions of *Capsicum chinense* fruits. Scientia Horticulturae, 112(3):266-270. doi:10.1016/j.scienta.2006.12.029
- Leite PSS, Rêgo ER, Nascimento NFF, Nascimento MF, Santos RMC, Fortunato FLG (2011) Variabilidade fenotípica em geração F2 de pimenteira ornamental. Horticultura Brasileira 29(2):S3030-S3035.
- Leite PSS, Rodrigues R, Silva RNO, Pimenta S, Medeiros AM, Bento CS, Gonçalves LSA (2016) Molecular and agronomic analysis of intraspecific variability in *Capsicum baccatum* var. pendulum accessions. Genetics and Molecular Research 15(4):1-16. doi: 10.4238/gmr.15048482
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Research 27(2):209-202.
- Milligan GW, Cooper MC (1985) An examination of procedures for determining the number of cluster in a data set. Psychometrika 50(2):159-179.
- Mojena R (1977) Hierarchical grouping methods and stopping rules: an evaluation. The Computer Journal 20(4):359-363.
- Moreira SO, Rodrigues R, Araújo ML, Sudré CP, Souza EMR (2009) Desempenho agrônômico de linhas endogâmicas recombinadas de pimenta em dois sistemas de cultivo. Ciência Rural 39(5):1387-1393.
- Moses M, Umaharan P, Dayanandan S (2014) Microsatellite based analysis of the genetic structure and diversity of *Capsicum chinense* in the Neotropics. Genetic Resources and Crop Evolution 61(4):741-755. doi: 10.1007/s10722-013-0069-y

- Moura MCCL, Gonçalves LSA, Sudré CP, Rodrigues R, Amaral Júnior AT, Pereira TNS (2010) Algoritmo de Gower na estimativa da divergência genética em germoplasma de pimenta. *Horticultura Brasileira* 28(2):155-161.
- Padilha HKM, Sigales CV, Villela JCB, Valgas RA, Barbieri RL (2016a) Agronomic evaluation and morphological characterization of chili peppers (*Capsicum annuum*, Solanaceae) from Brazil. *Australian Journal of Basic and Applied Sciences* 10(13):63-70.
- Padilha HKM, Sosinski EE, Barbieri RL (2016b) Morphological diversity and entropy of peppers (*Capsicum baccatum* and *Capsicum chinense*, Solanaceae). *International Journal of Current Research* 8(12):42758-42768.
- Paulus D, Valmorbida R, Santin A, Toffoli E, Paulus E (2015) Crescimento, produção e qualidade de frutos de pimenta (*Capsicum annuum*) em diferentes espaçamentos. *Horticultura Brasileira* 33(1):091-100. doi: 10.1590/S0102-053620150000100015
- Pessoa AMS, Rêgo ER, Carvalho MG, Santos CAP, Rêgo MM (2018) Genetic diversity among accessions of *Capsicum annuum* L. through morphoagronomic characters. *Genetics and Molecular Research* 17(1):gmr16039883. doi: 10.4238/gmr16039883
- R Development Core Team (2018) R: a language and environment for statistical computing. Vienna, Austria: the R Foundation for Statistical Computing.
- Rêgo ER, Finger FL, Rêgo MM (2012) Consumption of pepper in Brazil and its implications on nutrition and health of humans and animals. In: Salazar MA, Ortega JM (ed) *Peppers: Nutrition, Consumption and Health*, Nova Science Publishers. p.159-170.
- Silva AR, Dias CTS (2013) A cophenetic correlation coefficient for Tocher's method. *Pesquisa Agropecuária Brasileira* 48(6):589-596. doi: 10.1590/S0100-204X2013000600003
- Silva WCJ, Carvalho SIC, Duarte JB (2013) Identification of minimum descriptors for characterization of *Capsicum* spp. germplasm. *Horticultura Brasileira* 31(2): 190-202.
- Sneath PH, Sokal RR (1973) *Numerical taxonomy: The principles and practice of numerical classification*. W.H. Freeman. 573 p.
- Souza LR, Scossa F, Chaves IS, Kleessen S, Salvador LFD, Milagre JC, Finger FL, Bhering LL, Sulpice R, Araújo WL, Nikoloski Z, Fernie AR, Nesi AN (2015) Exploring natural variation of photosynthetic, primary metabolism and growth parameters in a large panel of *Capsicum chinense* accessions. *Planta* 242(3):677-691. doi:10.1007/s00425-015-2332-2
- Souza SAM, Martins KC, Pereira TNS (2011) Polimorfismo cromossômico em *Capsicum chinense*. *Ciência Rural* 41(10):1777-1783. doi:10.1590/S0103-84782011001000017
- Vasconcelos CS, Barbieri RL, Neitzke RS, Priori D, Fischer SZ, Mistura CC (2014) Distância genética entre variedades crioulas de *Capsicum chinense*. *Magistra* 26(2):178-185.
- Yamaki M, Menezes GRO, Paiva ALC, Barbosa L, Silva RF, Teixeira RB, Torres RA, Lopes PS (2009) Estudo de características de produção de matrizes de corte por meio da análise de componentes principais. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia* 61(1):227-231.