

## Genetic diversity among pepper (*Capsicum* spp.) accessions estimated by microsatellite markers

A.P. Dantas<sup>1</sup>, F. da S. Quemel<sup>1</sup>, J.S. Romero<sup>2</sup>, J.O. Santos<sup>3</sup> and A.D. Lopes<sup>1</sup>

<sup>1</sup> Programa de Pós-graduação em Biotecnologia aplicada à agricultura - Universidade Paranaense, UNIPAR, Umuarama, PR, Brasil

<sup>2</sup> Programa de Pós-graduação em Genética e Melhoramento - Universidade Estadual de Maringá, UEM, Maringá, PR, Brasil

<sup>3</sup> Universidade Federal do Maranhão, UFMA, MA, Brasil

Corresponding author: A.D. Lopes  
E-mail: anadanielalopes@prof.unipar.br

Genet. Mol. Res. 21 (1): gmr18953  
Received August 02, 2021  
Accepted November 25, 2021  
Published January 03, 2022  
DOI <http://dx.doi.org/10.4238/gmr18953>

**ABSTRACT.** It is essential that the genetic variability of preserved accessions germplasm banks be evaluated and quantified, which allows, for example, the detection of duplicates, generating more reliable data, especially for groups of plants that have a wide genetic variability, as in the case of the *Capsicum* genus. . We examined the DNA from the leaves of nine accessions of *Capsicum* spp. from the Federal University of Maranhão Germplasm Bank. Amplification was performed using 78 microsatellite primers previously described as potential molecular markers for *Capsicum* accessions. Thirty-three microsatellite primers produced amplification products, seven of which were polymorphic and used to investigate genetic variability among nine accessions. The proportion of polymorphic loci in pepper accessions was estimated at 21%. A mean of 2.28 alleles was detected per polymorphic locus and the number of alleles ranged from 1 to 3. The highest effective number of alleles was 1.62, and the highest values of mean observed heterozygosity ( $H_o$ ) and Shannon's diversity index ( $I$ ) were 0.533 and 0.421, respectively. Private alleles were detected in two accessions: UFMA17 and UFMA35. The

highest  $H_o$  value was found in the UFMA10 accession, whereas the lowest value was recorded in the UFMA35 accession. Fixation indices ( $F_{ST}$ ) indicated that 65% of the total allele frequency variation occurred within the analyzed loci. The polymorphism information content ranged from 0.011 to 0.354. Clustering of samples based on Nei's genetic distance matrix revealed the formation of three genetically distinct groups that allowed identifying genetically distinct groups based on SSR markers and discriminating the most closely related accessions of *Capsicum* spp. Neighbor-joining cluster analysis and PCoA allowed identifying genetically distinct groups based on SSR markers and discriminating the most closely related accessions of *Capsicum* spp.

**Key words:** Variability; Molecular characterization; Simple sequence repeat; Principal coordinate analysis; UPGMA

## INTRODUCTION

Brazil has the greatest plant genetic diversity in the world (Campos et al., 2016) and is known as a diversity center of peppers of the genus *Capsicum*. Throughout Brazil, *Capsicum* plants are either domesticated (widely cultivated), semi-domesticated (little cultivated), or wild (not commercially cultivated) (Maciel et al., 2016; Carvalho et al., 2017). In the state of Maranhão, peppers of the *Capsicum* genus are cultivated due to their mystical value, but also because they make up a large part of the typical foods of the region. They are used as pickled peppers, with babassu milk and tucupi, and are commonly found in local stores. In addition, its importance is also attributed to the value it adds to family farming, since its cultivation requires a high demand for labor, from cultivation to harvest.

Given the genetic variability of the genus *Capsicum*, species differ in color, fruit size, shape, chemical composition, and, especially, degree of pungency or piquancy (Chuah et al., 2008). The genus encompasses a wide diversity of bell and chili peppers with worldwide economic importance because of their use as food, medicine, ornamental plants, and raw materials for pharmaceutical applications (Rêgo et al., 2012; Gomes et al., 2019).

The expansion of agricultural boundaries associated with the growth of urban centers has increased the extinction risk of *Capsicum* spp. in some regions, thereby narrowing the genetic base (Maciel et al., 2016; Carvalho et al., 2017). Germplasm banks, together with institutions such as the National Registry of Cultivars, contribute to conservation by characterizing morphological, geographical, and molecular data of accessions (Hill et al., 2013; Maciel et al., 2016; Carvalho et al., 2017).

Morphological characterization comprises the quantitative and qualitative analysis of morphological characters, which are easily differentiable (e.g., color, fruit shape, and level of pungency) (Costa, 2015). Such characterization should be performed using descriptors as soon as individuals are inserted into the collection. This procedure generates useful information for collection management. However, there are some limitations with regard to characters that show non-additive inheritance, as they are highly influenced by the environment (Burle et al., 2010).

Evaluation of genetic variability, on the other hand, is based on molecular data, allowing the generation of a series of information about intrinsic characteristics of the species' genome. Such data are of great importance for conservation, whether *in situ* or *ex situ*, and genetic improvement programs (Azevedo, 2010). Molecular markers are important tools for characterizing plant species, including *Capsicum* spp. In particular, microsatellite markers, such as simple sequence repeats (SSRs), have contributed to the characterization of a variety of accessions in genetic improvement programs, including chili peppers, bell peppers, and their landrace varieties (Pereira et al., 2012; Oliveira et al., 2012; Naz et al., 2013; Azevedo et al., 2019).

SSR markers represent short regions of tandemly repeated units of one to six nucleotides whose level of polymorphism depends on differences in the number of repetitions (Litt and Luty, 1989). Microsatellite sequences are abundantly and uniformly dispersed in genomes of all organisms and are efficiently detected by polymerase chain reaction (PCR), independent of environmental influences and stage of plant development (Morgante and Olivieri, 1993).

SSRs can be developed directly from genomic DNA libraries or microsatellite libraries enriched for specific microsatellite sequences. They can also be found in online databases or obtained on the basis of microsatellite transferability (Kalia et al., 2011). One of the strategies for microsatellite analysis of species that do not have specific primers is to transfer primers of evolutionarily close species. The transferability of SSR markers is higher between species from the same genus or closely related genera than between species from distant genera of the same family (Cheng, 2016). This study aimed to assess genetic diversity among accessions of the genus *Capsicum* by using microsatellite markers.

## MATERIAL AND METHODS

### Germplasm

The biological material used for molecular characterization was obtained from the *Capsicum* spp. collection of the Center for Agricultural Sciences of the Federal University of Maranhão (CCA/UFMA), Brazil. Most accessions maintained in the germplasm bank were collected in Maranhão State. Of the nine accessions analyzed in this study, four were morphologically characterized by Silva et al. (2021), namely UFMA17, UFMA18, UFMA35, and UFMA 70. Accessions were morphologically characterized based on 24 qualitative descriptors developed by the International Plant Genetic Resources Institute (IPGRI, 1995). The choice of these materials was based on the availability of seeds and accessions with greater germination potential, besides having already been characterized morpho agronomically.

The collection maintained by the Federal University of Maranhão has approximately 100 accessions of peppers of the *Capsicum* genus, whose seeds are kept under refrigeration, under controlled humidity and temperature conditions. The accessions are constituted of the species *C. annum*, *C. chinense*, and *C. bacatum*, cultivated in the state of Maranhão, by small or medium producers or in urban gardens and domestic backyards (Silva et al., 2016). It is also noteworthy that the collection does not include wild material, only cultivated varieties. The objective is to maintain the accesses already cultivated by local producers, and not to provide material for cultivation.

## DNA extraction and quantification

Genomic DNA was extracted from young leaves collected from 15 pepper plants per accession of nine selected accessions, totaling 135 samples. Extraction was performed by the 2% cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1990), with modifications. Seeds were sown in polystyrene trays containing commercial substrate (Carolina Soil<sup>®</sup>) and maintained in a greenhouse at Paranaense University, campus 3, Umuarama, Paraná, Brazil, for 30 days.

Approximately 200 mg of fresh leaf tissue was ground in liquid nitrogen using a mortar and pestle. Subsequently, the resulting powder was collected into 2 mL microtubes and mixed with 700  $\mu\text{L}$  of 2% CTAB extraction buffer and 2-mercaptoethanol. After homogenization, microtubes were kept in a water bath at 65 °C for 30 min. Then, 600  $\mu\text{L}$  of a 24:1 mixture of chloroform/isoamyl alcohol was added, and samples were homogenized by manual inversion for 5 min and centrifuged for 10 min at 12,000 rpm. The upper phase was transferred to a new microtube and mixed with about 50  $\mu\text{L}$  (1/10 of the volume) of a solution containing 10% CTAB and 1.4 M NaCl. The aqueous phase was collected into clean microtubes and mixed with 400  $\mu\text{L}$  of cold isopropanol (−20 °C) by gently inverting the tubes for 1 min. Samples were placed in a freezer (−20 °C) overnight and, after this period, centrifuged for 15 min at 12,000 rpm. The supernatant was discarded, and the DNA pellet was washed twice with 70% ethanol and twice with absolute ethanol. Microtubes were allowed to air dry at room temperature, and the pellet was resuspended in TE solution (1 M Tris-HCl and 0.5M EDTA, pH 8.0) containing 10  $\mu\text{g mL}^{-1}$  RNase. DNA samples were stored in a refrigerator until use.

The concentration and purity of extracted DNA was estimated by absorbance at 260 nm and the 260 to 280 nm absorbance ratio, respectively. Absorbance readings were taken on a Spectra Max Plus 384 Microplate Reader. Subsequently, samples were diluted in TE to a concentration of 50  $\text{ng } \mu\text{L}^{-1}$ .

## Amplification reactions

A total of 78 microsatellite primers were tested in nine accessions of *Capsicum* spp. (Lee et al., 2004; Mimura et al., 2012) (Table 1). Amplification reactions were performed in 0.2 mL microtubes containing 10 $\times$  PCR buffer, 50 mM  $\text{MgCl}_2$ , 12.5 mM dNTPs, 10  $\text{mg } \mu\text{L}^{-1}$  bovine serum albumin, 1.0  $\mu\text{L}$  of each primer (10  $\mu\text{M}$ ), 5 U of Taq DNA polymerase, and 1.0  $\mu\text{L}$  of DNA (50  $\text{ng } \mu\text{L}^{-1}$ ). PCR was performed using an Eppendorf Mastercycler Personal system set at 94 °C for 4 min and 40 s for initial denaturation, followed by 30 cycles of 15 s at 94 °C, 10 s at 50–57 °C (depending on the primer used, see Table 1), 20 s at 72 °C, and final extension at 72 °C for 5 min.

Amplified fragments were separated by horizontal electrophoresis on 3.5% agarose gel in 1 $\times$  TBE buffer (90 mM Tris-borate and 2 mM EDTA). Gels were stained with 0.5  $\mu\text{g mL}^{-1}$  ethidium bromide and photographed under ultraviolet light using an L-PIX EX-Locus system. The size of PCR fragments was determined by comparison with a 1 kb DNA ladder.

**Table 1.** Microsatellite primers used for analysis of genetic diversity in the accessions UFMA08, UFMA10, UFMA17, UFMA14, UFMA15, UFMA18, UFMA35, UFMA70 and UFMA 105 of *Capsicum* spp.

Locus	Sequence	Repeat motif	Linkage group	Reference	$T_a$ (°C)
CAMS-117	F: TTGTGGAGGAAACAAGCAAA R: CCTCAGCCCAGGAGACATAA	(TG) <sub>21</sub> (TA) <sub>3</sub>	11	Mimura et al. (2012)	50
CAMS-142	F: TCCATATAGCCCGTGTGTGA R: CGTGGGAATACAATGCTAGA	(TA) <sub>3</sub> (AC) <sub>7</sub> ...(AC) <sub>12</sub> A(TA) <sub>8</sub>	7	Mimura et al. (2012)	50
CAMS-162	F: GGACCGTTCAGGAGGTTACA R: CCATCATTCAAAACCGAAT	(TA) <sub>7</sub> (TG) <sub>7</sub> TA(TG) <sub>6</sub>	1	Mimura et al. (2012)	50
CAMS-405	F: TTCTTGGGTCCCACACTTTC R: AGGTTGAAAGGAGGGCAATA	(TC) <sub>18</sub>	11	Mimura et al. (2012)	50
CAMS-806	F: GTCACAAGTGTCAAGGTAGGAG R: CCCCAAAAATTTCCCTCAT	(AGA) <sub>19</sub>	10	Mimura et al. (2012)	50
CAMS-864	F: CTGTTGTGGAAGAAGAGGACA R: GCTTCTTTTCAACCTCCTCCT	(AGA) <sub>32</sub>	7	Mimura et al. (2012)	50
Bd12	F: CAACAGTAGGACCCGAAAATCC R: ATGAAGGCTACTGCTGCGATCC	(CT) <sub>2</sub> (AG) <sub>15</sub>	1	Lee et al. (2004)	52

F, forward; R, reverse;  $T_a$ , annealing temperature.

## Data analysis

Only polymorphic microsatellite markers were considered for analysis. Genetic variability among *Capsicum* accessions was analyzed using GenAlEx software (Peakall and Smouse, 2012). The following parameters were determined: number of polymorphic loci, actual number of alleles per locus ( $N_a$ ), effective number of alleles ( $N_e$ ), number of private alleles ( $P_a$ , defined as alleles detected exclusively in a given population; Jiang et al., 2011), mean observed heterozygosity ( $H_o$ ), mean expected heterozygosity ( $H_e$ ), Shannon's diversity index ( $I$ ), fixation index ( $F_{ST}$ ), genetic distance between genotypes (Nei, 1972), and analysis of molecular variance (AMOVA). Nei's distance matrix for SSR data was constructed using PowerMarker software version 3.25 (Liu and Muse, 2005). Bootstrapping was performed with 1,000 replications, and a neighbor-joining tree was constructed using Mega X software version 10.2 (Kumar et al., 2018). The polymorphism information content (PIC) was calculated using PowerMarker software version 3.25. Principal coordinate analysis (PCoA) was performed to confirm dissimilarities using GenAlEx software (Peakall and Smouse, 2012).

## RESULTS

Proportion of polymorphic loci, number of alleles, effective number of alleles, observed and expected heterozygosities, and private alleles in pepper accessions

Thirty-three microsatellite primers (42%) designed for the genus *Capsicum* produced amplification products, seven of which were polymorphic and used to assess genetic variability among the nine pepper accessions. The estimated proportion of polymorphic loci (%P) in the nine pepper accessions was 21%. The percentage of polymorphic loci ranged from 14% (UFMA14 and UFMA70) to 71% (UFMA10), with a mean of 38% in the nine accessions analyzed.

A total of 16 alleles and a mean of 2.28 alleles per polymorphic locus were detected in the nine pepper accessions using seven microsatellite primers. Actual ( $N_a$ ) and effective ( $N_e$ ) numbers of alleles per polymorphic locus and mean observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosities for the seven loci are presented in Table 2. The number of alleles detected by microsatellite markers ranged from 1 to 3, and the highest effective number of alleles was 1.62. The highest  $H_o$  and  $I$  values were 0.533 and 0.421, respectively, both detected in the CAMS-806 locus, which contains repeat motifs of (AGA)<sub>19</sub> nucleotides. The lowest effective number of alleles (1.01) was detected in the CAMS-405 locus, which contains (TC)<sub>18</sub> repeats (Table 2). The mean genetic diversity for the seven SSR loci was 0.155.  $H_e$  values above 0.50 (indicating high genetic diversity) were not detected in any of the analyzed microsatellite loci (Table 2).

**Table 2.** Actual number of alleles ( $N_a$ ), effective number of alleles ( $N_e$ ), Shannon's index ( $I$ ), mean observed heterozygosity ( $H_o$ ), mean expected heterozygosity ( $H_e$ ), polymorphism information content (PIC), and fixation index ( $F_{ST}$ ) of single sequence repeat loci in the accessions UFMA08, UFMA10, UFMA17, UFMA14, UFMA15, UFMA18, UFMA35, UFMA70 and UFMA 105 of *Capsicum* spp.

Locus	$N_a$	$N_e$	$I$	$H_o$	$H_e$	PIC	$F_{ST}$
CAMS-117	2.00	1.20	0.199	0.159	0.128	0.162	0.278
CAMS-142	2.00	1.06	0.078	0.056	0.046	0.160	0.730
CAMS-162	2.00	1.33	0.230	0.322	0.166	0.243	0.385
CAMS-405	2.00	1.01	0.022	0.011	0.011	0.011	0.045
CAMS-806	3.00	1.62	0.421	0.533	0.289	0.354	0.327
CAMS-864	3.00	1.38	0.290	0.153	0.189	0.318	0.509
Bd12	2.00	1.47	0.360	0.452	0.254	0.279	0.275
Mean	2.28	1.29	0.229	0.241	0.155	0.218	0.364

Allele frequency was estimated for each of the seven SSR loci. Private alleles ( $P_a$ ) were detected in two accessions, namely UFMA17 (CAMS-806 and CAMS-864) and UFMA35 (CAMS-405). In addition to being an indicator of the occurrence of gene flow, private alleles reflect the level of genetic relationship between accessions or populations (Szpiech and Rosenber, 2011). The effective number of alleles ranged from 1.04 (UFMA35) to 1.67 (UFMA17). The highest  $H_o$  value (0.419) was found in the UFMA10 accession, whereas the lowest value (0.043) was recorded in the UFMA35 accession (Table 3).

$F_{ST}$  values were classified into the following classes, as proposed by Nei (1978): low ( $F_{ST} < 0.05$ ), intermediate ( $0.05 < F_{ST} < 0.15$ ), and high ( $>0.15$ ), indicating low, medium, and high population structures, respectively. The analyzed loci were found to have high genetic differentiation ( $>0.15$ ), except CAMS-405 (Table 2).  $F_{ST}$  values indicated that 65% of the total allele frequency change at loci is due to genetic differences between *Capsicum* accessions.

PIC ranged from 0.011 (CAMS-405) to 0.354 (CAMS-806), with a mean value of 0.218 (Table 2). Three of the seven loci analyzed were moderately polymorphic ( $0.25 < PIC < 0.50$ ) (CAMS-806, CAMS-864, and Bd12). CAMS-117, CAMS-142, CAMS-162, and CAMS-405 were uninformative ( $PIC \leq 0.25$ ) (Table 2). PIC values were higher than  $H_e$  and  $H_o$  in three of the seven loci, lower in three of the seven loci, and equal in CAMS-405.

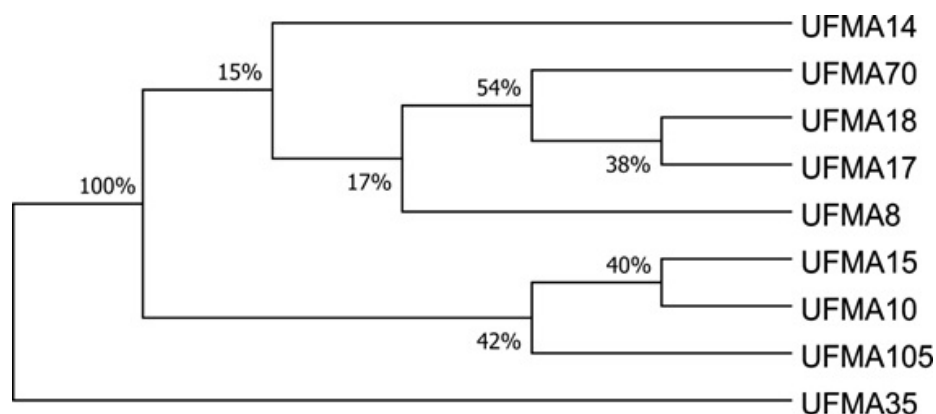
**Table 3.** Actual number of alleles ( $N_a$ ), effective number of alleles ( $N_e$ ), mean observed heterozygosity ( $H_o$ ), and mean expected heterozygosity ( $H_e$ ) in the accessions UFMA08, UFMA10, UFMA17, UFMA14, UFMA15, UFMA18, UFMA35, UFMA70 and UFMA 105 of *Capsicum* spp., as estimated by microsatellite markers.

Accessions	$N_a$	$N_e$	$H_o$	$H_e$
UFMA08	1.42	1.33	0.257	0.180
UFMA10	1.71	1.50	0.419	0.286
UFMA17	1.85	1.67	0.357	0.299
UFMA14	1.14	1.14	0.143	0.071
UFMA15	1.42	1.31	0.286	0.169
UFMA18	1.28	1.25	0.196	0.133
UFMA35	1.42	1.04	0.043	0.041
UFMA70	1.14	1.14	0.143	0.071
UFMA105	1.28	1.28	0.286	0.143

Fixation index, polymorphism information content, and AMOVA

### Neighbor-joining cluster analysis

Sample clustering using Nei's genetic distance matrix by the neighbor-joining method revealed three genetically distinct groups (Figure 1). The first comprised UFMA14, UFMA70, UFMA18, UFMA17, and UFMA08; the second, UFMA15, UFMA10, and UFMA105; and the third, UFMA35 only. UFMA35 was the most distant accession in the sample, whereas UFMA18 and UFMA17, in Group I, and UFMA15 and UFMA10, in Group II, had the greatest genetic similarity.

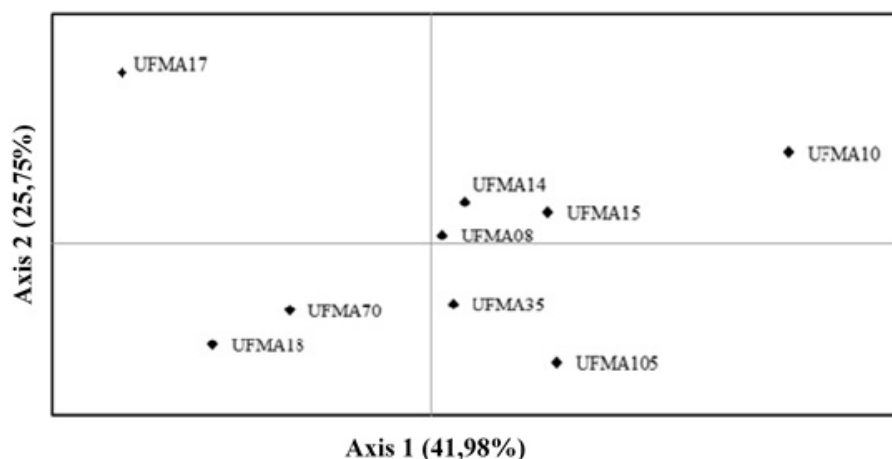


**Figure 1.** Neighbor-joining tree of the *Capsicum* accessions UFMA08, UFMA10, UFMA17, UFMA14, UFMA15, UFMA18, UFMA35, UFMA70 and UFMA 105 assessed using simple-sequence repeat markers. The tree was constructed using Nei's distance matrix. Bootstrap values at the nodes were obtained in 1,000 replicates.

### Principal coordinate analysis

Accessions were classified into two groups according to genetic relationships and, probably, the species to which they belonged (Figure 2). The first (horizontal) axis (42%) indicated that UFMA10, UFMA15, UFMA14, UFMA08, UFMA35, and UFMA105, which were closest to each other, differed from the group containing UFMA17, UFMA70, and

UFMA18. This result demonstrated that approximately 68% of the total variability contained in the evaluated accessions can be explained by the two-dimensional plane.



**Figure 2.** Principal coordinate analysis of the *Capsicum* accessions UFMA08, UFMA10, UFMA17, UFMA14, UFMA15, UFMA18, UFMA35, UFMA70 and UFMA 105 assessed using simple-sequence repeat markers. Axis 1 accounts for the highest amount of total variance, separating the sample into two principal groups.

## DISCUSSION

Of the 33 amplified fragments, only 21% were polymorphic. Although the number of polymorphic loci was not high, it was possible to discriminate accessions, suggesting that microsatellite markers are promising for the study of diversity in *Capsicum*.

The mean number of alleles estimated for the nine *Capsicum* accessions (2.28 alleles per locus) is consistent with the results of Dhaliwal (2014), who reported a mean of 2.78 alleles per locus in 64 accessions of *Capsicum*. Of the 64 accessions, 49 were from indigenous sources and 15 were of exotic origin, belonging to *C. annuum*, naturally occurring landrace varieties of *C. chinense* and *C. frutescens*, and *C. frutescens* (tabasco pepper). Sharmin et al. (2018) observed a mean of 2.00 alleles per locus in 20 accessions of peppers grown in Bangladesh. According to Lopes et al. (2015), comparison with other studies should be made with caution, by taking into account differences in sample sizes; furthermore, the same number of alleles may not always indicate the same amount of variability.

Studies based on the efficiency of transferability of SSR markers between species of the genus *Capsicum* have been developed and revealed promising results, as well as those described by Buso et al. (2016), who evaluated the transferability of 45 SSR markers to 48 accessions of *Capsicum* elite in the germplasm bank. The authors detected an average of 6 alleles, ranging from 2 to 11 alleles. According to the authors, the new markers developed were highly informative and their transferability between species was confirmed, suggesting that these markers are powerful tools for syntax analysis in peppers. In addition, new candidate SSR markers for *C. flexuosum*, *C. recurvatum*, and *C. campylopodium* were prepared.



Furthermore, Moulin et al. (2015) proposing the construction of an integrated genetic map for *Capsicum baccatum* found that 62 of the 152 pairs of microsatellite markers available for *C. annuum* were successfully transferred to *C. baccatum*. Furthermore, 42 of these markers were mapped, allowing the introduction of *C. baccatum* in synteny studies with other species of the genus *Capsicum*.

The lowest effective number of alleles ( $N_e = 1.04$ ) and the lowest diversity ( $H_e = 0.041$ ) were detected in UFMA35. Azevedo et al. (2019), in evaluating *Capsicum* spp. with 33 SSR markers, found a range of 1.15 to 3.00 effective alleles, with a mean of 1.691. These values are in line with those expected for accessions that are genetically close and from autogamous species (Azevedo et al., 2019).

Ulhoa et al. (2014) found that, of the 15 SSR markers used to evaluate 24 strains of yellow jalapeño (*C. annuum*) from the Embrapa Hortaliças germplasm bank, 73% were homozygous and the remaining 27% had heterozygosity values of 5–73% (73% heterozygosity was observed in CAMS-024 and CAMS-117). In our study, however, CAMS-117 did not exhibit high heterozygosity and was not a highly informative locus for accession discrimination.

The relatively high heterozygosity observed in SSRs, especially for UFMA10 ( $H_o = 0.419$ ), is potentially significant, because high heterozygosity suggests a substantial amount of adaptive genetic variation in the plant population, allowing plants to escape the effects of agents that limit their development and survival (Allendorf and Luikart, 2007). Other plant populations showed low levels of heterozygosity, such as UFMA35 ( $H_o = 0.043$ ). Azevedo et al. (2019), in investigating six varieties of *Capsicum* spp., found minimum and maximum  $H_o$  values of 0.0 and 0.857, respectively, surpassing those found in the present study.

$H_o$  values higher than  $H_e$  values, observed in the current study for all accessions, suggest excess homozygotes. An increase in homozygosity tends to increase inbreeding coefficient, self-fertilization rate, and crosses between related individuals, which makes them inbred and, therefore, more susceptible to genetic erosion. Expansion of the genetic base occurs mainly from crosses between less similar accessions, that is, contrasting accessions (Sánchez, 2008).

The number of private alleles detected here was lower than that found by Azevedo et al. (2019), who reported a high number of private alleles in *Capsicum* varieties when testing SSR markers (18 in 'De Cayenne', 17 in 'Cayenne Long Slin', 16 in 'UENF Campista' and 'Amarela Comprida', 15 in 'Cascadura Ikeda', 8 in 'UENF 2154', and 6 in 'Malagueta'). According to the authors, the higher the number of private alleles, the lower the gene flow; consequently, accessions with the highest number of private alleles tend to be more genetically distant from other accessions of the same species, genus, or population (Azevedo et al., 2019). This was observed in the current study, in which the accession with the highest occurrence of private alleles (UFMA35) had great dissimilarity with other accessions by the neighbor-joining method (Figure 3).

Shannon's index ( $I$ ) calculates species diversity or richness and phenotypic or genotypic diversity within a population. It is possible to assess the proportion of total diversity that is concentrated within and between groups (Sambatti et al., 2000). Low  $I$  values were observed in the present study, ranging from 0.022 in CAMS-405 to 0.421 in CAMS-806. A possible explanation for these findings is the origin and morphological characteristics of accessions. UFMA35, for instance, as shown by the neighbor-joining method, has some peculiarities that do not allow clustering with the other accessions. Added

together, these findings confirm the existence of variability in the CCAA/UFMA collection of *Capsicum* spp.  $I$  values were lower than those reported by Azevedo et al. (2019) for *Capsicum* spp., which ranged from 0.410 to 1.099.

$F_{ST}$  values close to one indicate a high degree of difference between populations. Values equal to zero indicate identical allele frequencies, whereas values equal to one represent different alleles. The mean value of the parameter was 0.364, which means that the simulated ideal population had similar allelic frequencies, probably caused by genetic drift.

The estimates generated by SSR markers demonstrated that there is a higher proportion of genetic variability within (65%) *Capsicum* spp. accessions than among them (35%). One possible explanation for this result is that alleles with low frequencies contribute little to differentiation between populations. High divergence within populations and low divergence between populations have been reported in previous studies on tropical tree species (Hamrick and Godt, 1990).

PIC quantifies the value of polymorphism information of a locus marker; the parameter considers the number and relative frequency of alleles (Romero-Severson et al., 2001). Low PIC values indicate a high frequency of one or two alleles. Loci with high PIC values can be used to discriminate germplasm. Loci can be classified as very informative ( $PIC > 0.5$ ), moderately informative ( $0.25 < PIC < 0.5$ ), or uninformative ( $PIC < 0.25$ ) (Botstein et al., 1980). The mean PIC values for four of the seven loci were lower than 0.25; therefore, these loci were uninformative (CAMS-806, CAMS-864, and Bd12) had PIC values of 0.25–0.5, being classified as moderately informative. Sharmin et al. (2018) reported mean PIC values of 0.371 (range of 0.255 to 0.500) for 20 accessions of peppers. According to the authors, the highest PIC value was obtained for CAMS-806, as also observed here; the locus had the highest polymorphic content and was therefore considered informative for diversity studies in *Capsicum* spp. The polymorphism information content was on average 0.60, ranging from 0.20 to 0.83 in studies of transferability of SSR markers to accessions of the genus *Capsicum* (Buso et al., 2016).

The genetic diversity between the nine *Capsicum* spp. accessions, estimated from the dissimilarity matrix (Nei, 1972), ranged from 0.040 (UFMA14 × UFMA35 and UFMA35 × UFMA70) to 0.389 (UFMA10 × UFMA17). The narrow genetic base did not lead to the formation of heterotic groups; however, the dendrogram revealed that crosses between plants of the UFMA35 accessions with accessions of Group I, for instance, are promising for enlargement of the genetic base.

The results of the neighbor-joining tree allowed us to infer that cultivars with the highest values of genetic distance can be applied in reciprocal recurrent selection programs, given that a pair of accessions with high genetic distance are more different from each other than a pair with a low value (Sharmin et al., 2018). However, recommendations for future crosses based on the neighbor-joining tree should be considered with caution, because isolated analysis of the mean allele number and AMOVA results (which suggests that most of the genetic variation occurs within cultivars) hinders the selection of pairs of cultivars with sufficient genetic distance to provide variations for heterotic hybrids.

Both clustering methods (neighbor-joining tree and PCoA) provided similar results. UFMA70, UFMA18, and UFMA 17 were grouped in the same cluster in the dendrogram and remained closely related in the PCoA plot, in agreement with the results of Silva et al. (2021), who characterized the morphology of *Capsicum* spp. using 24 qualitative

descriptors proposed for the genus by IPGRI (1995). The same occurred with the accessions UFMA15, UFMA10, and UFMA105, which had the same grouping pattern in both analyses. UFMA35 was the most distant from the others, forming an isolated group. The accession was the only accession from Espírito Santo State, suggesting that sampling locality was responsible for differences between this accession and the others. These findings also confirm the existence of variability in the CCAA/UFMA collection of *Capsicum* spp.

Molecular-level characterization suggested that UFMA35 plants can be crossed with Group I accessions (Figure 3) to expand the genetic base of *Capsicum* spp., increase heterozygosity, and introduce useful genes into commercial cultivars, which have important agronomic characteristics. Molecular characterization and study of the genetic diversity of these accessions may complement morphological analyses, being useful for breeding programs and future studies on genetic resources within the genus *Capsicum* spp.

## CONCLUSIONS

Neighbor-joining cluster analysis and PCoA allowed identifying genetically distinct groups based on SSR markers and distinguishing closely related accessions of *Capsicum* spp. These tools are useful to estimate genetic diversity and guide the crossing of contrasting parents in genetic improvement programs for *Capsicum* spp. This preliminary study can be used to support the development of a germplasm conservation program for the genus *Capsicum* and may serve as a basis for future investigations on marker-assisted selection.

## ACKNOWLEDGMENTS

The authors thank the Universidade Federal do Maranhão for the concession of the biological material used, the Universidade Paranaense and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for funding this research.

## CONFLICTS OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

- Allendorf FW, Luikart G and Aitken SN (2007). Conservation and the genetics of populations. *Mammalia*. 6: 189-197.
- Azevedo CDO, Rodrigues R and Sudré CP (2019). Microsatellites for detecting inconsistencies in *Capsicum* cultivars registration in Brazilian database: more than meets the eye. *Hortic. Bras.* 37: 285-293. <https://doi.org/10.1590/S0102-053620190306>.
- Azevedo VCR (2010). Manual de curadores de germoplasma vegetal: Caracterização Molecular. Embrapa Recursos Genéticos e Biotecnologia, Brasília, 17p.
- 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.
- Burle ML and Oliveira M (2010). Manual de Curadores de Germoplasma Vegetal: Caracterização Morfológica. Embrapa Recursos Genéticos e Biotecnologia, Brasília, 16p.
- Buso GS, Reis AM, Amaral ZP and Ferreira ME (2016). Novel and highly informative *Capsicum* SSR markers and their cross-species transferability. *Genet. Mol. Res.* 15: gmr.15038689.
- Campos AL, Marostega TN, Cabral NSS, Araújo KL, et al. (2016). Morphoagronomic and molecular profiling of *Capsicum* spp. from southwest Mato Grosso, Brazil. *Genet. Mol. Res.* 15: gmr.15038167.

- Carvalho F, Maciel G, De Oliveira C, Nogueira A, et al. (2017). Comparison of clustering methods to study genetic divergence in pepper accessions at vegetative stage. *Rev. Bras. Pl. Med.* 19: 236-243.
- Cheng J, Zhao Z, Li B, Qin C, et al. (2016). A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in *Capsicum*. *Sci. Rep.* 6: 18919. 10.1038/srep18919.
- Chuah AM, Lee YC, Yamaguchi T, Takamura H, et al. (2008). Effect of cooking on the antioxidant properties of coloured peppers. *Food Chem.* 111: 20-28. 10.1016/j.foodchem.2008.03.022.
- Costa LV, Bentes JL, Lopes MT, Alves SR, et al. (2015). Caracterização de acessos de pimentas do Amazonas. *Hortic. Bras.* 33: 290-298. /10.1590/S0102-053620150000300003 .
- Dhaliwal MS, Yadav A and Jindal SK (2014). Molecular characterization and diversity analysis in chilli pepper using simple sequence repeats (SSR) markers. *Afr. J. Biotechnol.* 13: 3137-3143. 10.5897/AJB2014.13695.
- Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus.* 12: 13-15.
- Gomes GP, Baba VY, Dos Santos OP, Sudré CP, et al. (2019). Combinations of distance measures and clustering algorithms in pepper germplasm characterization. *Hortic. Bras.* 37: 172-179. 10.1590/S0102-053620190207.
- Hamrick JL and Godt MJW (1990). Allozyme diversity in plant species. In: Plant population genetics, breeding and genetic resources (Brow ADH, Clegg MT, Kahler AL, Weir BS, eds.). Sinauer, Sunderland, 10.1007/BF00985462.
- Hill TA, Ashrafi H, Reyes-Chin-Wo S, Yao J, et al. (2013). Characterization of *Capsicum annuum* genetic diversity and population structure based on parallel polymorphism discovery with a 30K unigene Pepper GeneChip. *Plos One* 8: e56200.10.1371/journal.pone.0056200.g003.
- IPGRI (International Plant Genetic Resource Institute) (1995). Descriptor for *Capsicum* (*Capsicum spp.*). IPGRI, Rome. [https://www.biodiversityinternational.org/fileadmin/\\_migrated/uploads/tx\\_news/Descriptors\\_for\\_capsicum\\_\\_Capsicum\\_spp.\\_\\_345.pdf](https://www.biodiversityinternational.org/fileadmin/_migrated/uploads/tx_news/Descriptors_for_capsicum__Capsicum_spp.__345.pdf).
- Jiang Z, Xia H, Basso B and Lu B (2011). Introgression from cultivated rice influences genetic differentiation of weedy rice populations at a local spatial scale. *Theor. Appl. Genet.* 124: 309-322. 10.1007/s00122-011-1706-5.
- Kalia RK, Rai MK, Kalia S, Singh R, et al. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica.* 177: 309-334. 10.1007/s10681-010-0286-9.
- Kumar S, Stecher G, Knayaz C and Tamura K (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35: 1547-1549.
- Lee JM, Nahm SH, Kim YM and Kim BD (2004). Characterization and molecular genetic mapping of microsatellite loci in pepper *Theor. Appl. Genet.* 108: 619-627. 10.1093/molbev/msy096.
- Litt M and Luty JA (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* 44: 397-401.
- Liu K and Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics.* 21: 2128-2129. 10.1093/bioinformatics/bti282.
- Lopes AD, Scapim CA, Machado MFPS, Mangolim CA, et al. (2015). Genetic diversity assessed by microsatellite markers in sweet corn cultivars. *Sci. Agr.* 72: 513-519. 10.1590/0103-9016-2014-0307.
- Maciel GM, Oliveira CS, Siquieroli ACS, Melo EI, et al. (2016). Genetic dissimilarity among the physiochemical characteristics of fruit from pepper accessions. *Biosci. J.* 32: 978-985. 10.14393/BJ-v32n4a2016-33893.
- Mimura Y, Inoue T, Minamiyama Y and Kubo N (2012). An SSR-based genetic map of pepper (*Capsicum annuum* L.) serves as an anchor for the alignment of major pepper maps. *Breed Sci.* 62: 93-98. doi:10.1270/jsbbs.62.93
- Morgante M and Olivieri AM (1993). PCR-amplified microsatellites as markers in plant genetics. *Plant J.* 3: 175-182.
- Moulin MM, Rodrigues R, Ramos HC, et al. (2015). Construction of an integrated genetic map for *Capsicum baccatum* L. *Genet Mol Res.* 14: 6683-6694.
- Naz S, Zafullah A, Shahzadhi K and Munir N (2013). Assessment of genetic diversity within germplasm accessions in tomato using morphological and molecular markers. *J. Anim. Plant Sci.* 23: 1099-1106. 10.1007/BF00266542.
- Nei M (1972). Genetic distance between populations. *Am. Nat.* 106: 283-292.
- Nei M (1978). Estimation of Average Heterozygosity and Genetic Distance from A Small Number of Individuals. *Genetics* 89: 583-590. 10.1093/genetics/89.3.583.
- Oliveira MVC, Baliza DP, Souza GA, Carvalho SP, et al. (2012). Caracterização de clones de mandioca utilizando marcadores microssatélites. *Cienc. Agron.* 43: 170-176. 10.1590/S1806-66902012000100021.
- Peakall R and Smouse P (2012). GenAlEx6.5: genetic analysis in Excel. Population genetic software for teaching and research, an update. *Bioinformatics.* 28: 2537-9. 10.1093/bioinformatics/bts460.
- Pereira WA, Sávio FL, Dias DCFDS, Cruz CD, et al. (2012). Fluxo gênico recíproco entre cultivares de soja convencional e geneticamente modificada. *Pesqui. Agropecu. Bras.* 47: 227-236. 10.1590/S0100-204X2012000200011.
- Rêgo ER, Finger FL and Rêgo MM (2012). Types, uses and fruit quality of Brazilian chili peppers. In: Spices: Types, uses and health benefits (Kralis JF, eds.). Nova Science, New York, 131-144.
- Sambatti JBM, Martins OS and Ando A (2000). Distribuição da diversidade isoenzimática e morfológica da mandioca na agricultura autóctone de Ubatuba. *Sci. Agr.* 57: 75-80. 10.1590/S0103-90162000000100013.
- Romero-Severson J, Smith JSC, Ziegler J, Hauser J, et al. (2001). Pedigree analysis and haplotype sharing within diverse groups of *Zea mays* L. inbreds. *Theor. Appl. Genet.* 103: 567-574. 10.1007/PL00002911.

- Sánchez CFB (2008). Diversidade entre e dentro de populações simuladas sob deriva genética. Master's thesis, UFV, Viçosa.
- Sharmin A, Hoque ME, Haque MM and Khatun F (2018). Molecular diversity analysis of some chilli (*Capsicum* spp.) genotypes using SSR markers. *Am. J. Plant Sci.* 9: 368-379. 10.4236/ajps.2018.93029.
- Silva ML, Marques SJ, Guse LG, Santos JO, et al. (2016). Prospecção da diversidade genética entre acessos de *Capsicum* sp. conservados na "coleção de pimentas do CCAA/UFMA". In: Tópicos em produção agrícola no leste Maranhense: livro comemorativo dos 10 anos do curso de agronomia do CCAA/UFMA (Farias MF, Furtado MB, eds). EDUFMA, São Luis, 265-278.
- Silva JM, Santos JO, Da Silva IM, De Oliveira RV, et al. (2021). Caracterização morfológica de acessos de pimentas (*Capsicum* spp.) conservados no estado do Maranhão. *Braz. J. Dev.* 7: 21358-21371. 10.34117/bjdv7n3-039.
- Szpiech ZA and Rosenberg NA (2011). On the size distribution of private microsatellite alleles. *Theor. Popul. Biol.* 80: 100-113. 10.1016/j.tpb.2011.03.006.
- Ulhoa AB, Pereira TN, Silva RN, Ragassi CF, et al. (2014). Caracterização molecular de linhagens de pimenta do tipo Jalapeño amarelo. *Hortic. Bras.* 32: 35-40. 10.1590/S0102-05362014000100006.