GMR

The Original by FUNPEC-RP

Establishment of *Capsicum frutescens* core collections based on morphological and molecular descriptors and on virus incidence

S.I.C. Carvalho[1,2], C.F. Ragassi[2], F.G. Faleiro[1,3], G.S.C. Buso[4], L.B. Bianchetti[4], M.F. Lima[2], F.J.B. Reifschneider[5] and C.S.C. Ribeiro[2]

[1] Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, Brasilia, DF, Brasil
[2] Embrapa Hortaliças, Brasilia, DF, Brasil
[3] Embrapa Cerrados, Brasilia, DF, Brasil
[4] Embrapa Recursos Genéticos e Biotecnologia, Brasilia, DF, Brasil
[5] Embrapa Secretaria de Inteligência e Relações Estratégicas, Brasilia, DF, Brasil

Corresponding author: S.I.C. Carvalho
E-mail: sabrina.carvalho@embrapa.br

Genet. Mol. Res. 19 (1): gmr18503
Received October 16, 2019
Accepted March 11, 2020
Published March 31, 2020
DOI http://dx.doi.org/10.4238/gmr18503

ABSTRACT. Malagueta (*Capsicum frutescens*) is one of the most widely consumed and cultivated Brazilian hot peppers. It is an important crop for smallholder farmers throughout the country. Currently, the demand for new hot pepper cultivars is increasing. A germplasm collection of *C. frutescens* is maintained at Embrapa Vegetables, Brasilia, Brazil, the branch for vegetable crops of the Brazilian Agricultural Research Corporation (EMBRAPA). This is considered to be the main collection representing the variability of this species in the country. Four core collections of 13 accessions each were established through 1) random selection, 2) selection based on morphological and 3) SSR similarity groups and 4) selection based on SSR similarity groups associated with virus incidence. Characterizing the original germplasm collection (103 accessions) through 57 morphological characters, 239 alleles of 24 microsatellite (SSR) *loci* and incidence of six virus species provided the information used for selecting the accessions. Discriminating *C. chinense* and *C. frutescens* species proved to be inaccurate when relying only on morphological characterization for 5% of the accessions, whereas molecular characterization was decisive

for the species identification of all accessions. The SSR allelic variability within each core collection was compared with the full *C. frutescens* collection. Selection based on SSR grouping associated with data on viruses incidence provided the highest allelic representativeness among the four strategies (77% of the allelic variability present in the full collection), in addition to satisfactorily representing the Brazilian geographic diversity. The core collection based on morphological characters was also highly representative of the allelic variability (73%) in the original full collection.

Key words: Germplasm collection; Genetic variability; Peppers; Microsatellites

INTRODUCTION

Brazil is an important diversity center of the genus *Capsicum*, including species with all levels of domestication (domesticated, semi domesticated and wild). There are five domesticated species of *Capsicum*: *C. annuum* L. var. *annuum*; *C. baccatum* L. var. *pendulum* (Willd.) Eshbaugh; *C. chinense* Jacq.; *C. frutescens* L. and *C. pubescens* Ruiz & Pavon. Only *C. pubescens* does not occur in Brazil (Reifschneider et al., 2015); the primary diversity center of this species is Bolivia (Eshbaugh, 1979). *C. frutescens* is found from the lowlands of Southeastern Brazil to Central America, including the Antilles, in the Caribbean. In addition to the Americas, *C. frutescens* is also cultivated in Africa, India, China, Japan, and Thailand (Mongkolporn and Taylor, 2011).

The most common morphological types of *C. frutescens* are Malagueta in Brazil and Tabasco in the United States and Central America. In addition to the well-known morphological types, spontaneous forms of Malagueta occur with relative frequency in South America, especially in the Northern Brazil (Bianchetti and Carvalho, 2005). Recently, the occurrence of a new morphological type was reported in Brazil and named 'Malaguetinha' (small Malagueta pepper) (Carvalho et al., 2017).

Little information is available on the morphology and variability of *C. frutescens*, as well as other traits that may be valuable in the development of new Malagueta cultivars. Wild *C. frutescens* trace have been found at archaeological sites in Central and South America, but ethnobotanists believe that domestication of the Tabasco pepper occurred in Panama, subsequently being spread to Mexico and the Caribbean (DeWitt and Bosland, 1997). Most Tabasco varieties grown in the United States resulted from human selection within the existing varieties; though the scientific literature poorly covers the variability of agronomic characteristics within the *C. frutescens* genetic pool (Jarret et al., 2007).

Germplasm Banks (GB) of *Capsicum* spp. maintained by various institutions have the functions of preserving genetic diversity and promoting their use in breeding programs. The largest collections are in the United States, South America, Asia and Europe. The United States Department of Agriculture (USDA) collection maintains about 5,000 accessions of *Capsicum* spp., and about 590 accessions are *C. frutescens*, 62 of which were originated from Brazil (GRIN-USDA, 2019). Approximately 8,000 accessions of *Capsicum* spp. are preserved in the Asian Vegetable Research and

Development Center (AVRDC) collection in Taiwan, of which 726 accessions are *C. frutescens* and 18 are from Brazil (AVGRIS, 2019).

The *Capsicum* GB maintained by Embrapa Vegetables, Brasilia, Federal District, Brazil, which is the branch for vegetable crops of the Brazilian Agricultural Research Corporation (Embrapa), under the aegis of the Brazilian Ministry of Agriculture, Livestock, and Food Supply, was initiated almost four decades ago and nowadays has about 2,000 accessions, representing five domesticated *Capsicum* species and dozens of semi-domesticated and wild species, from many countries and from various regions of Brazil. This *Capsicum* GB has served as a genetic basis for a broad breeding program of Embrapa and also partners in Brazil and abroad. Over 30 thousand lines and populations of domesticated and semi domesticated species and dozens of cultivars of several types of spicy and low pungency peppers have been made available to different market segments (Reifschneider et al., 2015, 2016).

An important focus for the development of new pepper cultivars has been resistance to diseases, especially to viruses. Viruses are among the most important and complex diseases that affect *Capsicum* species in the world, and especially for Malagueta pepper in Brazil, causing significant losses in production. Several viruses infect species in the genus *Capsicum*, the most important being the tospoviruses (*Tomato spotted wilt virus* - TSWV, *Groundnut ringspot virus* - GRSV), the potyviruses (*Potato virus* Y - PVY, *Pepper yellow mosaic* virus - PepYMV), a tobamovirus (*Pepper mild mottle virus* - PMMoV) and a cucumovirus (*Cucumber mosaic virus* – CMV). *Capsicum frutescens* hybrids and lines from the breeding program and genotypes from the *Capsicum* GB of Embrapa are potential sources of virus resistance (Lima et al. 2017).

The *Capsicum* GB of Embrapa Vegetables has about 112 accessions registered as *C. frutescens*, collected under different ecological conditions in the North, Northeast, Southeast, Center-West and South of Brazil. To contribute for characterizing the genetic variability within *C. frutescens* and, consequently, enable its use by breeding programs, the establishment of core collections is mandatory.

The establishment of core collections, as suggested by Frankel (1984), consists of organizing collections that represent the genetic diversity of a crop species and its relatives with a minimum of repetitiveness. Efficient employment of the genetic diversity of the GB in breeding programs, as well as obtaining information on the representativeness of a collection in relation to the genetic diversity of a crop species are advantages of establishing core collections (Ferreira et al., 2007). A sampling of 10% of the accessions present in the original collection is a reference size for a core collection (Brown, 1989).

Core collections have been developed in GB in Brazil and other countries around the world for many plant species and different purposes, for representing genetic, cultural, ecological or geographical diversity, as well as the diversity of characters of agronomic importance, such as resistance to pests and diseases. Core collections of *Capsicum* have already been established using phenotypic data (Zewdie et al., 2004), genotypic data (Mongkolporn et al., 2015) and phenotypic and genotypic data compilations (Nicolaï et al., 2013; Lee et al., 2016).

This study aimed at establishing and comparing core collections of *C. frutescens* obtained from the *Capsicum* GB of Embrapa Vegetables, which is comprised of 112 accessions registered as *C. frutescens*, by using different selection strategies based on morphological (57 descriptors) and molecular (24 single sequence repeat - SSR *loci*) information, as well as incidence of six virus species, aiming to potentiate the use of *C. frutescens* germplasm in the development of new Malagueta pepper cultivars by breeding programs.

MATERIAL AND METHODS

Morphological and molecular characterization

One hundred and twelve accessions registered as *C. frutescens* (original collection) originated from the North, Northeast, Center-West, Southeast and South of Brazil preserved at the Embrapa Vegetables *Capsicum* GB were cultivated as described by Carvalho et al. (2017) in a greenhouse. Verification of the taxonomic classification was made through a key for identification of domesticated and semi domesticated *Capsicum* species and varieties occurring in Brazil (Bianchetti and Carvalho, 2005). Morphological characterization was carried out using 53 descriptors recommended for *Capsicum* (International Plant Genetic Resources Institute - IPGRI, 1995) and still four additional descriptors were added for this study: fruit position, pungency, aroma and segregation, totalizing 57 morphological descriptors. The set of descriptors included: 17 passport/vegetative part descriptors (origin, species, plant height, plant width, leaf color, leaf shape, leaf density, stem shape, stem color, stem length, stem diameter, branching habit, nodal anthocyanin, growth habit, tillering, leaf pubescence, stem pubescence), 16 inflorescence/seed descriptors (male sterility, calyx margin, number of flowers/axil, calyx pigmentation, flower position, stigma exsertion, calyx annular constriction, corolla spot color, anther color, filament color, corolla color, days to flowering, corolla shape, seed color, number of seeds/fruit, seed surface) and 24 fruit descriptors (fruit persistence, number of locules, fruit wall thickness, fruit pedicel length, fruit weight, fruit width, pungency, fruit shape, days to fruiting, fruit color at immature stage, fruit color at mature stage, placenta length, aroma, fruit length, fruit blossom end appendage, varietal mixture condition, segregation, fruit shape at pedicel attachment, fruit position, anthocyanin spot, neck at base of fruit, fruit shape at blossom end, cross-sectional corrugation and fruit surface). These data are presented in Carvalho et al. (2017).

Molecular characterization was carried out using the 24 SSR primer sets from Carvalho et al. (2017). Besides accessions identified as *C. frutescens*, 11 accessions belonging to other *Capsicum* species were included in the molecular characterization, eight *C. chinense* (CNPH 4315, CNPH 4316, CNPH 4325, CNPH 4327, CNPH 4328, CNPH 4332 A, CNPH 4360 and CNPH 4361), one *C. praetermissum* (CHPH 3825) and two *C. annuum var. annuum* (CNPH 30062 and CNPH 40013), totalizing 123 accessions. Based on morphological characterization data, genetic distance between accessions was estimated by simple correspondence analysis. The genetic distances

obtained through microsatellite (SSR) *loci* were calculated with the Genes software, following the methodology described by Carvalho et al. (2017).

Genetic dissimilarity matrices obtained through morphological descriptors or SSR *loci* were used to perform cluster analysis using the unweighted pair group mean averaging (UPGMA) method. Graphical dispersion analyses according to multidimensional scales were also performed via main coordinates, with SAS and Statistica softwares. The correlation between the genetic distances and its significance (t-test) were estimated according to SSR *loci* and distances calculated based on morphological descriptors by Pearson's correlation coefficient, with the statistical software Genes.

Assessment of virus incidence

The incidence of viruses was evaluated on plants growing under open field conditions with natural infection, without artificial inoculation of viruses. Field trials were performed at Embrapa Vegetables, Brasilia, DF, Brazil. Pepper seedlings were produced in the greenhouse in Styrofoam trays and transplanted to the field 40 days after sowing.

Infection of seedlings in the field trial was favored by presence of older pepper plants naturally infected with viruses and showing characteristic disease symptoms. The presence of viruses in those pepper plants was checked by serology using polyclonal antibodies against the coat protein of each virus species (produced at Embrapa Vegetables), by DAS-ELISA (double-antibody sandwich-Enzyme-linked immunosorbent assay; Clark and Adams, 1977). These infected plants served as virus inoculum to infect pepper seedlings of different accessions in the field.

Assessment of virus incidence on plants of the 112 accessions was performed on leaf samples collected from each individual plant just before the flowering stage and analyzed by DAS-ELISA test. Young leaves were collected from at least three different branches of the same plant to increase the chances of virus detection. Plants were evaluated for the presence of tospoviruses (*Tomato spotted wilt virus* – TSWV and *Groundnut ringspot virus* – GRSV), potyviruses (*Potato virus Y* – PVY and *Pepper yellow mosaic virus* – PepYMV), a tobamovirus (*Pepper mild mottle virus*– PMMoV) and a cucumovirus (*Cucumber mosaic virus* – CMV), following the methodology described by Lima et al. (2017). Leaf extracts were prepared in extraction buffer (1.4 M NaCl; 0.02 M KH_2PO_4; 0.08 M $Na_2HPO_4.12H_2O$; 0.02 M KCl; pH 7.4), at proportion of 1g/10 mL. Antibodies and conjugates were used at concentration 1mg/mL. Infected indicator plants (*Nicotiana tabacum* cv. TNN infected with PVY; *Datura stramonium* - TSWV, GRSV; *C. annuum* cv. Ikeda – CMV, PepYMV, PMMoV) and healthy plants were the positive and the negative controls for each virus species, respectively. Absorbance readings at 405 nm were measured in ELISA reader (Titertek Multiskan). Samples were considered as positive when their absorbance values were at least three times higher than the absorbance values of the negative control.

Composition of core collections

The method dependent on genetic variability was used for selecting accessions to compose the nuclear collections. Genetic variability data was used to cluster accessions into groups of genetic similarity; then, at least one accession was selected to represent each group, so that the selected accessions represented more than 70% of the allelic variability of the original collection (Faleiro et al., 2007).

Different selection strategies based on morphological characters, SSR *loci* and virus incidence were used for composition of four core collections: 1) random selection (CoreColl-1); 2) selection based on the representativeness of morphological similarity groups (CoreColl-2); 3) selection based on representativeness of similarity groups defined by SSR *loci* (CoreColl-3) and 4) selection based on the representativeness of similarity groups defined by SSR *loci* associated with virus incidence data (CoreColl-4).

The choice of accessions from the similarity groups established through morphological data, SSR *loci* and SSR *loci* associated with virus incidence data was based on the representation of all groups established by 1 up to 4 accessions for each group. Each core collection comprised a total of 13 accessions.

The percentage of the alleles of the full *C. frutescens* collection present in each core collection was used to calculate the core collection allelic representativeness. For comparison purposes, an additional core collection was established contemplating 100% of the allelic variability present in the full *C. frutescens* collection.

RESULTS

The 112 accessions registered in the Embrapa Vegetable's GB as belonging to the species *C. frutescens* were reclassified using a classification key on morphological characteristics according to Bianchetti and Carvalho (2005). After reclassification, only 96 accessions were confirmed as being *C. frutescens*. Among the remaining accessions, 14 were reclassified as *C. chinense*, one as *C. baccatum* var. *pendulum* and one as *C. annuum* var. *glabriusculum*. Based on molecular characterization data, accessions also clustered into four groups; however, different from morphological analysis, 103 were *C. frutescens*, seven *C. chinense*, one *C. baccatum* var. *pendulum* and one *C. annuum* var. *glabriusculum* (Table 1). The 103 accessions identified molecularly as *C. frutescens* were then considered as the full *C. frutescens* collection for all the subsequent analyses.

Cluster analysis based on the matrix of distances obtained from morphological descriptors divided the 96 *C. frutescens* accessions into six groups of genetic similarity (Figure 1). The group of 103 accessions established by the molecular analysis according to SSR *loci* comprised the 96 accessions morphologicaly classified as *C. frutescens* and seven accessions classified as *C. chinense* according to morphological traits (Figure 2). The seven accessions classified as *C. chinense* according to morphological traits were, then, assumed to be *C. frutescens* considering molecular data. The group of 103 *C. frutescens* accessions was also divided into six similarity clusters. Further details on these groups are in Carvalho et al. (2017).

Table 1. Identification of the accessions of the *Capsicum* Germplasm Bank (GB) of Embrapa Vegetables, evaluated in this research based on morphological characteristics and molecular traits.

	Identification N° CNPH[/1]	Origin (Brazilian Regions)	Morphologically identified species	Molecularly identified species
1	63	Southeast	*C. frutescens*	*C. frutescens*
2	287	Center-West	*C. frutescens*	*C. frutescens*
3	595	Southeast	*C. frutescens*	*C. frutescens*
4	597	Center-West	*C. frutescens*	*C. frutescens*
5	1386	Northeast	*C. frutescens*	*C. frutescens*
6	2631	Southeast	*C. frutescens*	*C. frutescens*
7	2744	North	*C. frutescens*	*C. frutescens*
8	2841	North	*C. frutescens*	*C. frutescens*
9	2866 A	North	*C. frutescens*	*C. frutescens*
10	2866 B	North	*C. chinense*	*C. chinense*
11	2869	North	*C. frutescens*	*C. frutescens*
12	2870	North	*C. frutescens*	*C. frutescens*
13	2871	North	*C. chinense*	*C. chinense*
14	3241	Center-West	*C. frutescens*	*C. frutescens*
15	3257	Northeast	*C. frutescens*	*C. frutescens*
16	3286	North	*C. frutescens*	*C. frutescens*
17	3349	Southeast	*C. frutescens*	*C. frutescens*
18	3374	Center-West	*C. frutescens*	*C. frutescens*
19	3399	North	*C. frutescens*	*C. frutescens*
20	3410	Center-West	*C. frutescens*	*C. frutescens*
21	3414	Southeast	*C. frutescens*	*C. frutescens*
22	3440	Southeast	*C. frutescens*	*C. frutescens*
23	3446	Southeast	*C. frutescens*	*C. frutescens*
24	3448	North	*C. frutescens*	*C. frutescens*
25	3453	North	*C. chinense*	*C. chinense*
26	3462	North	*C. frutescens*	*C. frutescens*
27	3470	North	*C. frutescens*	*C. frutescens*
28	3484	North	*C. frutescens*	*C. frutescens*
29	3499	North	*C. frutescens*	*C. frutescens*
30	3535 A	North	*C. frutescens*	*C. frutescens*
31	3535 B	North	*C. chinense*	*C. chinense*
32	3539	North	*C. frutescens*	*C. frutescens*
33	3546	North	*C. frutescens*	*C. frutescens*
34	3550	North	*C. frutescens*	*C. frutescens*
35	3606 A	North	*C. frutescens*	*C. frutescens*
36	3606 B	North	*C. frutescens*	*C. chinense*
37	3612	North	*C. frutescens*	*C. frutescens*
38	3621	Southeast	*C. frutescens*	*C. frutescens*
39	3630	-	*C. frutescens*	*C. frutescens*
40	3645	Northeast	*C. frutescens*	*C. frutescens*
41	3646	Northeast	*C. frutescens*	*C. frutescens*
42	3647	Northeast	*C. frutescens*	*C. frutescens*
43	3648	Northeast	*C. frutescens*	*C. frutescens*
44	3649	Northeast	*C. frutescens*	*C. frutescens*
45	3667	Southeast	*C. frutescens*	*C. frutescens*
46	3696	Center-West	*C. frutescens*	*C. frutescens*
47	3697	Southeast	*C. frutescens*	*C. frutescens*
48	3698	Center-West	*C. frutescens*	*C. frutescens*
49	3715	North	*C. frutescens*	*C. frutescens*
50	3716	North	*C. chinense*	*C. frutescens*
51	3746	South	*C. frutescens*	*C. frutescens*
52	3804	Center-West	*C. chinense*	*C. frutescens*
53	3805	Center-West	*C. chinense*	*C. frutescens*
54	3806	Center-West	*C. chinense*	*C. frutescens*
55	3813	Center-West	*C. chinense*	*C. frutescens*
56	3815	Center-West	*C. chinense*	*C. frutescens*
57	3816	Center-West	*C. frutescens*	*C. frutescens*
58	3818	Center-West	*C. chinense*	*C. frutescens*
59	3819	Southeast	*C. frutescens*	*C. frutescens*
60	3820	Northeast	*C. frutescens*	*C. frutescens*
61	3821	Northeast	*C. frutescens*	*C. frutescens*
62	3835	Southeast	*C. frutescens*	*C. frutescens*
63	3847	Southeast	*C. frutescens*	*C. frutescens*
64	3861	Northeast	*C. frutescens*	*C. frutescens*
65	3880	North	*C. frutescens*	*C. frutescens*
66	3885	North	*C. frutescens*	*C. frutescens*
67	3891	North	*C. frutescens*	*C. frutescens*
68	3894	North	*C. frutescens*	*C. frutescens*
69	3906	Center-West	*C. frutescens*	*C. frutescens*
70	3932	North	*C. frutescens*	*C. frutescens*
71	3944	Northeast	*C. frutescens*	*C. frutescens*
72	3984	South	*C. frutescens*	*C. frutescens*

73	4005	North	*C. frutescens*	*C. frutescens*
74	4011	Center-West	*C. frutescens*	*C. frutescens*
75	4020	Southeast	*C. frutescens*	*C. frutescens*
76	4037	North	*C. frutescens*	*C. frutescens*
77	4052	Northeast	*C. frutescens*	*C. frutescens*
78	4069	Center-West	*C. frutescens*	*C. frutescens*
79	4082	Center-West	*C. baccatum* var. *pendulum*	*C. baccatum* var. *pendulum*
80	4083	Center-West	*C. frutescens*	*C. frutescens*
81	4084 A	Center-West	*C. frutescens*	*C. frutescens*
82	4084 B	Center-West	*C. chinense*	*C. chinense*
83	4085	Center-West	*C. frutescens*	*C. frutescens*
84	4095	South	*C. frutescens*	*C. frutescens*
85	4105	Center-West	*C. frutescens*	*C. frutescens*
86	4138	North	*C. frutescens*	*C. frutescens*
87	4154 A	Southeast	*C. frutescens*	*C. frutescens*
88	4154 B	Southeast	*C. chinense*	*C. chinense*
89	4161	Southeast	*C. frutescens*	*C. frutescens*
90	4184	North	*C. frutescens*	*C. frutescens*
91	4191	North	*C. frutescens*	*C. frutescens*
92	4195	North	*C. frutescens*	*C. frutescens*
93	4212	Center-West	*C. annuum* var. *glabriusculum*	*C. annuum* var. *glabriusculum*
94	4224	Center-West	*C. frutescens*	*C. frutescens*
95	4231	Center-West	*C. frutescens*	*C. frutescens*
96	4237	Center-West	*C. frutescens*	*C. frutescens*
97	4263	Northeast	*C. frutescens*	*C. frutescens*
98	4264	Northeast	*C. frutescens*	*C. frutescens*
99	4265	Northeast	*C. frutescens*	*C. frutescens*
100	4266	Northeast	*C. frutescens*	*C. frutescens*
101	4267	Northeast	*C. frutescens*	*C. frutescens*
102	4268	Northeast	*C. frutescens*	*C. frutescens*
103	4269	Northeast	*C. frutescens*	*C. frutescens*
104	4270	Northeast	*C. frutescens*	*C. frutescens*
105	4271	Northeast	*C. frutescens*	*C. frutescens*
106	4272	Northeast	*C. frutescens*	*C. frutescens*
107	4273	Northeast	*C. frutescens*	*C. frutescens*
108	4274	Center-West	*C. chinense*	*C. chinense*
109	4283	North	*C. frutescens*	*C. frutescens*
110	4304	North	*C. frutescens*	*C. frutescens*
111	4353	North	*C. frutescens*	*C. frutescens*
112	4364	North	*C. frutescens*	*C. frutescens*

[71] CNPH= Embrapa Vegetables.

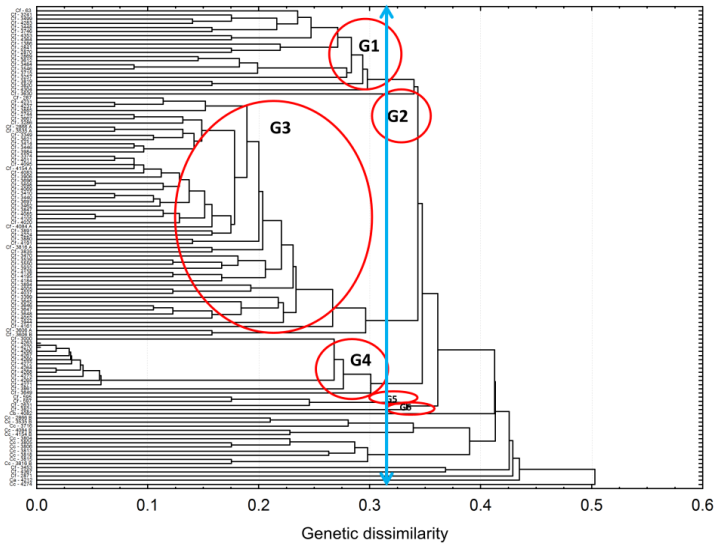

Figure 1. Cluster analysis of 112 *Capsicum* accessions (96 *C. frutescens* accessions clustered into six groups of genetic similarity) based on a matrix of genetic distances calculated using 57 morphological descriptors. The UPGMA method was used as clustering criterion. Brasília, Embrapa Vegetables, 2011.

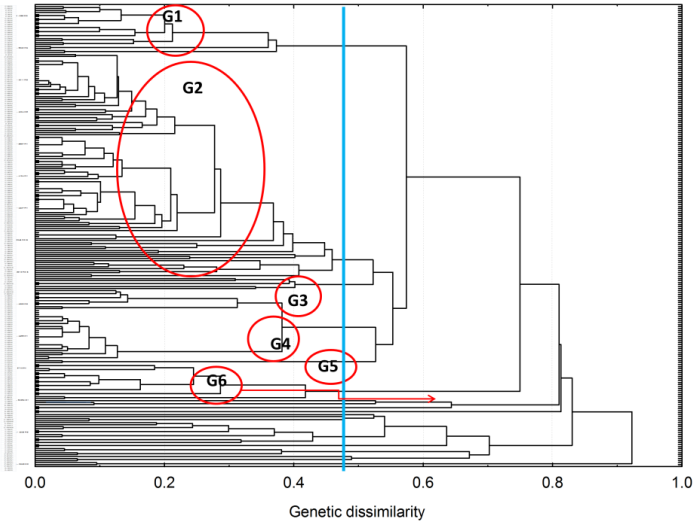

Figure 2. Cluster analysis of 123 *Capsicum* accessions (103 *C. frutescens* accessions divided into six groups of genetic similarity) based on a matrix of genetic distances obtained from 239 alleles of 24 SSR *loci*. The UPGMA method was used as clustering criterion. Brasília, Embrapa Vegetables, 2013.

A variety of virus symptoms was observed in plants in the field trial according to the resistance level of each accession, age of plants at the time of infection and occurrence of multiple infections in a same plant with more than a virus species. Pepper plants from different accessions exhibited a wide range of symptoms including mottling, green mosaic, yellow mosaic, vein banding, chlorotic ringspots, and necrotic ringspots on leaves, leaf distortion and stunting of plants. Virus symptoms were more severe on plants when infected at seedling stage or when mix infected.

Selection of *C. frutescens* accessions with low virus incidence was based on serological test results obtained from evaluation of the original collection (112 accessions) to natural infection with six virus species (TSWV, GRSV, PVY, PepYMV, PMMoV and CMV) in a field trial. Total 543 plant samples were tested, from which 42% were infected with one or more virus species. Basically, potyviruses (PepYMV=27.2%; PVY=4.8%) and tospoviruses (TSWV=12.5%; GRSV=28.2%), which are transmitted by aphids and trips, respectively, occurred in similar percentages. Both groups of viruses are amongst the most important pathogens to hot peppers and considered a limiting factor to the pepper crop cultivation in Brazil, affecting production and quality of fruits (Lima et al., 2011a). However, PMMoV (25.9%) which is found very often infecting the pepper crop in Brazil (Lima et al., 2011b), was one of the most detected viruses. The potential importance of PMMoV is attributed mainly because it is seed-borne and can be easily transmitted from plant to plant during cultivation practices. CMV was not found in the collected samples.

Mixed infection of plants was very common and occurred in 30% of hot pepper plants, in particular, those transmitted by insects (e. g. tospoviruses and potyviruses) with tobamovirus. According to serological test results, plants of selected accessions

that compose the CoreColl-4 tested negative for at least one virus group (tospovirus; potyvirus) and presented low virus incidence for the other groups (Table 2). Then, CNPH 597, CNPH 3698, CNPH 3885, CNPH 3894 and CNPH 3944 were positive for tospoviruses (e. g. GRSV) but not for potyviruses (PepYMV; PVY). On the other hand in CNPH-3715, just the presence of potyviruses (PepYMV; PVY) was detected, and no tospoviruses. At least five accessions tested positive for species of both group of viruses (CNPH 3470: GRSV and PVY; CNPH 3820: GRSV and PepYMV; CNPH 3821: GRSV and PepYMV; CNPH 3861: GRSV and PVY; CNPH 3891: GRSV and PepYMV). Similar behavior of CNPH 597 and CNPH 3820 to virus resistance in the field was observed by Lima et al. (2017).

Table 2. Evaluation of natural infection caused by tospoviruses (*Tomato spotted wilt virus* - TSWV; *Groundnut ringspot virus* - GRSV), potyviruses (*Pepper yellow mosaic virus* - PepYMV; *Potato virus Y* - PVY) and a tobamovirus (*Pepper mild mottle virus*- PMMoV) in plants of 112 accessions of Malagueta hot pepper (*Capsicum frutescens*) grown under open field conditions, using polyclonal antibodies, in DAS-ELISA. Brasilia, DF - Embrapa Vegetables.

Acessions Identification (CNPH #)	Sample (#)	Infected plants (%)[1]				
		TSWV	GRSV	PepYMV	PVY	PMMoV
0063	12	58.3	8.3	33.3	8.3	58.3
0595	12	0	25	16.7	0	41.7
287	3	0	0	33.3	0	3.33
597	9	33.3	0	0	0	0
1386	13	30.8	15.4	15.4	30.8	53.8
2631	12	25	33.3	58.3	8.3	8.3
2744	13	46.2	0	7.7	23.1	61.5
2841	9	11.1	77.8	33.3	44.4	77.8
2866 A	3	0	0	0	0	100
2866 B	nt[2]	nt	nt	nt	nt	nt
2869	11	0	18.2	18.2	36.4	27.3
2870	12	50	16.7	58.3	50	58.3
2871	8	0	50	0	37.5	100
3241	13	0	0	0	0	15.4
3257	13	0	23.1	7.7	0	53.8
3349	3	0	0	0	0	100
3374	12	33.3	25	0	8.3	50
3286	18	0	5.6	16.7	0	44.4
3399	2	0	0	0	0	100
3410	3	0	0	0	0	100
3414	3	0	0	0	0	100
3440	4	0	0	25	0	100
3446	2	0	0	0	0	50.0
3453	11	0	9.1	0	27.3	18.2
3448	13	53.8	23.1	15.4	0	38.5
3462	12	8.3	25	41.7	0	8.3
3499	12	8.3	0	0	0	0
3470	11	0	18.8	0	18.2	45.5.0
3484	nt	nt	nt	nt	nt	nt
3546	12	0	40	0	41.6	50
3535 A	3	0	0	0	0	0
3535 B	nt	nt	nt	nt	nt	nt
3539	3	0	0	33.3	0	33.3
3550	3	0	0	66.7	0	66.7
3606 A	5	0	20	0	0	20
3606 B	nt	nt	nt	nt	nt	nt
3612	8	37.5	0	12.5	0	50
3621	3	0	33.3	0	0	66.7
3630	12	0	16.7	0	8.3	25
3649	5	0	0	20.0	0	0
3645	3	0	0	33.3	0	100
3646	3	0	0	0	0	33.3
3647	3	0	0	0	0	100
3648	3	66.7	0	0	0	66.7
3696	3	0	0	33.3	0	0
3667	15	0	13.3	0	6.7	6.7

3697	2	0	50.0	0	0	0
3698	9	0	11.1	0	0	11.1
3715	8	0	0	62.5	25.0	25.0
3716	3	0	100	33.33	0	0
3746	10	10	0	20	0	50
3819	13	7.7	0	0	0	30.8
3820	12	0	8.3	53.8	0	0
3835	13	0	38.5	0	7.7	15.4
3804	3	0	33.3	66.7	0	33.3
3805	3	0	33.3	53.8	0	0
3806	3	33.3	0	100	0	33.3
3813	3	0	66.7	33.3	0	0
3815	3	0	33.3	0	0	0
3816	3	66.7	0	0	0	0
3818	3	100	0	66.7	0	33.3
3821	4	0	50.0	50.0	0	0
3847	2	0	100	50.0	0	0
3861	16	0	25	0	56.3	12.5
3880	3	0	33.3	33.3	0	0
3885	4	0	100	0	0	25
3891	3	0	100	33.3	0	0
3894	3	0	66.7	0	0	0
3906	3	0	100	0	0	0
3932	3	0	66.7	0	0	33.3
3944	3	33.3	0	0	0	0
3984	3	0	66.7	0	0	0
4005	3	0	66.7	0	0	33.3
4011	3	0	66.7	33.3	0	0
4020	3	0	33.3	66.7	0	33.3
4037	3	0	66.7	0	0	0
4052	3	0	100	0	0	0
4069	3	0	33.3	0	0	0
4082	3	0	66.7	0	0	03
4083	3	100	66.7	0	0	0
4084 A	3	0	33.3	100	0	0
4084 B	3	0	33.3	100	0	0
4085	3	0	100	100	0	0
4095	3	66.7	0	100	0	0
4105	3	66.7	0	100	0	0
4138	3	0	66.7	100	0	0
4154 A	3	0	33.3	66.7	0	0
4154 B	3	33.3	0	33.3	0	0
4161	3	0	100	100	0	0
4184	3	66.7	0	66.7	0	0
4191	3	66.7	0	66.7	0	0
4195	3	0	0	100	0	0
4212	3	0	66.7	66.7	0	0
4224	3	33.3	0	33.3	0	0
4231	3	0	50.0	0	0	0
4237	3	0	33.3	33.3	0	0
4263	nt	nt	nt	nt	nt	nt
4264	nt	nt	nt	nt	nt	nt
4265	nt	nt	nt	nt	nt	nt
4266	nt	nt	nt	nt	nt	nt
4267	nt	nt	nt	nt	nt	nt
4268	nt	nt	nt	nt	nt	nt
4269	nt	nt	nt	nt	nt	nt
4270	nt	nt	nt	nt	nt	nt
4271	nt	nt	nt	nt	nt	nt
4272	nt	nt	nt	nt	nt	nt
4273	nt	nt	nt	nt	nt	nt
4274	nt	nt	nt	nt	nt	nt
4283	nt	nt	nt	nt	nt	nt
4304	nt	nt	nt	nt	nt	nt
4353	nt	nt	nt	nt	nt	nt
4364	nt	nt	nt	nt	nt	nt

[1/](Number of infected plants/total number of plants tested in DAS-Elisa)X100. Samples were considered as positive when absorbance values were at least three times higher compared to absorbance values of healthy control.[2/]nt=not tested.

The core collections (CoreColl-1; CoreColl-2; CoreColl-3; CoreColl-4) of *C. frutescens* each consisted of 13 accessions, 12.6% from the original collection (103 accessions). These 13 accessions selected by using each approach, as well as the allelic representativeness of each core collection regarding to the original collection, are presented in Table 3.

Table 3. Lists of accessions in the core collections established according to four selection approaches and a collection representing 100% of the allelic variability of the full *Capsicum frutescens* collection composed of 103 accessions.

Selection	Random (CoreColl-1)	Morphological descriptors (CoreColl-2)	SSR (CoreColl-3)	SSR + Viruses (CoreColl-4)	SSR + Viruses***
Accessions selection	CNPH 287	CNPH 63	CNPH 63	CNPH 597	CNPH 597
	CNPH 2744	CNPH 287	CNPH 287	CNPH 3470	CNPH 3470
	CNPH 3286	CNPH 575	CNPH 1386	CNPH 3484	CNPH 3484
	CNPH 3399	CNPH 3446	CNPH 2841	CNPH 3698	CNPH 3698
	CNPH 3440	CNPH 3550	CNPH 3286	CNPH 3715	CNPH 3715
	CNPH 3462	CNPH 3630	CNPH 3399	CNPH 3820	CNPH 3820
	CNPH 3470	CNPH 3645	CNPH 3630	CNPH 3821	CNPH 3821
	CNPH 3539	CNPH 3649	CNPH 3716	CNPH 3861	CNPH 3861
	CNPH 3697	CNPH 3821	CNPH 3835	CNPH 3885	CNPH 3885
	CNPH 3932	CNPH 3944	CNPH 3894	CNPH 3891	CNPH 3891
	CNPH 4020	CNPH 4020	CNPH 4037	CNPH 3894	CNPH 3894
	CNPH 4283	CNPH 4263	CNPH 4161	CNPH 3944	CNPH 3944
	CNPH 4364	CNPH 4304	CNPH 4264	CNPH 4304	CNPH 4304
	-	-	-	-	CNPH 3374*
	-	-	-	-	CNPH 3606*
	-	-	-	-	CNPH 3646*
	-	-	-	-	CNPH 4005*
	-	-	-	-	CNPH 4084*
	-	-	-	-	CNPH 3716**
	-	-	-	-	CNPH 3805**
	-	-	-	-	CNPH 3813**
Allelic Representativeness (%)	29.3	73	75.8	77	100

* Segregating accessions. ** Accessions identified through morphological characterization as *C. chinense* and molecular characterization as *C. frutescens*. ***Minimal collection with the smallest number of accessions that represent 100% allelic representativeness of the full collection.

The allelic representativeness of the core collections ranged from 29.3% (random selection) to 77% (selection based on molecular markers and viruses). The representativeness of the CoreColl-2 established according to morphological descriptors (73%) was very close to that based on SSR *loci* (75.8%). The positive correlation value of 0.66 among genetic distances calculated according to morphological descriptors and distances based on SSR explains these results. The satisfactory representativeness of the genetic variability of the core collections established in our study can also be seen in the scatterplots (Figure 3).

The CoreColl-4 based on molecular markers data and virus incidence information gave the highest representativeness (77%) of the allelic variability of the original collection. The 13 accessions represented three Brazilian geographic regions: North (n = 7), Northeast (n = 4) and Center-West (n = 2). In addition, accessions showed uniform distribution in the scatter plot figure (Figure 3b), representing part of the genetic variability of the full *C. frutescens* collection.

The minimal collection with 100% allelic representativeness was achieved by including additional eight accessions (Table 3) to the initial group of 13 accessions selected according to molecular markers and incidence of viruses. To reach the maximum allelic variability five of these accessions were segregating and resulted from interspecific crossings, and three accessions that were identified by morphological characterization as *C. chinense*, which presented characteristics that discriminate this species from others, such as a calyx annular constriction (Carvalho et al., 2014).

However, they have characteristics similar to *C. frutescens* (exclusively erect flowers, exclusively erect, elongated, red and spicy fruits with a very thin outer wall).

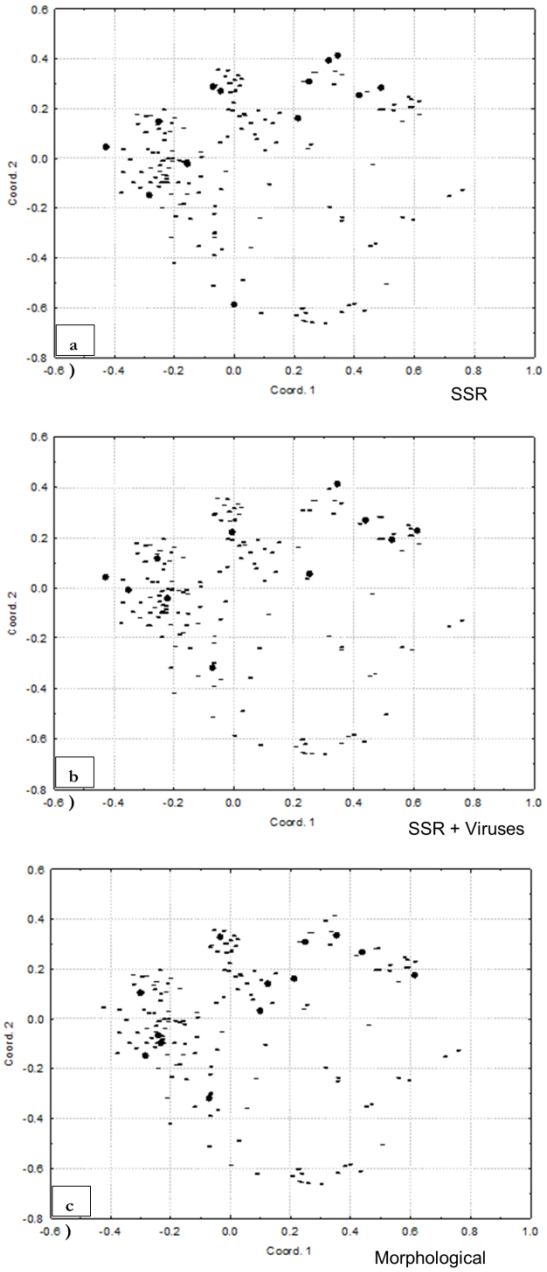

Figure 3. Graphic dispersion of 103 accessions of *Capsicum frutescens* based on a matrix of genetic distances generated using 239 alleles of 24 SSR *loci*. The 13 highlighted accessions were selected based on the representativeness of the similarity groups defined by: a) SSR; b) SSR and incidence of viruses and c) Morphological descriptors.

DISCUSSION

Organizing a germplasm collection from a core collection is an approach aimed primarily at improving conservation, accessibility and encourages the employment of genetic resources by breeding programs. Using genetic variability will provide, through genetic combinations, the emergence of new, better adapted, productive and disease resistant genotypes, among other characteristics of interest (Martinez et al., 2017).

Core collections of *C. annuum*, *C. baccatum* and *C. chinense* species were established based on morphological descriptors for GB maintained by the Southern Plant Genetic Resource Conservation Unit, Griffin, GA, USA (Zewdie et al., 2004). Species stratification before implementing the clustering process was strategically important because the genes or alleles found in different *Capsicum* species may differ. In addition, interspecific hybridization is difficult and laborious to transfer characteristics among *Capsicum* species (Zewdie et al., 2004).

Due to high genetic proximity existing between *C. frutescens* and *C. chinense*, some accessions can be easily confused, when relying on morphological traits. In general, both *C. frutescens* and *C. chinense* have a set of morphological characteristics that enable discriminating from each other (position of the flower, presence or absence of calyx annular constriction and fruit shape). However, there are intermediate accessions, which present characteristics of both species, leading to a misidentification when only morphological descriptors are employed. In our study, some accessions presented intermediate phenotypes, hampering their identification when the classification relied exclusively on morphological data.

In the morphological characterization analysis, the accessions CNPH 3804, CNPH 3805, CNPH 3806, CNPH 3813, CNPH 3815, CNPH 3816 and CNPH 3818 were initially classified as *C. chinense*. However, when molecular characterization was performed, these accessions grouped in Group 2 (Figure 2) of *C. frutescens*, representing the most divergent accessions within this group. In fact, passport data of these seven accessions suggested that they came possibly from cross-fertilization, resulting from proximity between plantings of these two species.

Similar results were verified in studies reported by Baral and Bosland (2004); they morphologically characterized 301 genotypes of *C. frutescens* and *C. chinense* and found that 8% of the accessions had an intermediate phenotype, hampering species classification. When performing molecular analysis through RAPD markers, Baral and Bosland (2004) detected no integration between accessions of *C. frutescens* and *C. chinense*, because probably no interspecific hybrids were included in the research. According to these authors, accessions showing intermediate phenotypes can be explained by introgressive hybridization; genes from one species move to another through interspecific hybridization process followed by successive backcrossing to one of the parents.

An excellent example of introgression is the cultivar Greenleaf Tabasco developed by interspecific hybridization between *C. frutescens* and *C. chinense* and then repeated backcrossing with *C. frutescens*. Greenleaf Tabasco resembles *C. frutescens*, but also has some morphological characteristics of *C. chinense* and results of

molecular analysis revealed that it actually has *C. chinense* alleles (Baral and Bosland, 2004). Thus, the identification of *Capsicum* accessions in the present study with intermediate phenotype (morphological characteristics between *C. chinense* and *C. frutescens*) may be a result of introgressive natural hybridization and no intraspecific variation.

The analysis of genetic variability among *C. frutescens* accessions herein was performed with genotypes originated from different ecological conditions of North, Northeast, Center-West, Southeast and South Regions of Brazil, relevant to breeding programs of that hot pepper species. The 13 accessions of CoreColl-4 (12.6% in size of the full *C. frutescens* collection with 103 accessions) selected based on SSR and virus incidence (Figure 3a), have a high allelic representativeness (77%) and occupied almost all the scatter graphic, representing a large part of the genetic variability of the full collection prevenient mainly from three Brazilian regions (53.8% North, 30.8% Northeast and 15.4% Center-West). The highest percentage of accessions was from the North region, considered rich in genetic diversity of Malagueta pepper and indeed, spontaneous populations of *C. frutescens* are found (Bianchetti and Carvalho, 2005). In addition, the present study represent an important contribution for the knowledge of the genetic diversity of *C. frutescens*, especially for the most popular types found in Brazil, such as 'Malagueta' and 'Malaguetinha' pepper.

In the development of a core collection Brown (1989) suggests a 10% sampling in order to retain at least 70% of alleles present in the original collection, to represent total genetic diversity. Samples of the core and the whole collection can be compared to determine whether they have broadly similar molecular marker alleles (van Hintum et al., 2000).

The core collection based on molecular markers and virus incidence (CoreColl-4) proposed in the present study retained the highest representativeness of the genetic variability (77%) among the four strategies studied. It maximizes genetic possibilities and increase the chances of success in developing cultivars adapted to distinct regions and ecosystems. The inclusion of eight accessions - five intermediate between *C. frutescens* and *C. chinense* and three identified as *C. chinense* in the morphological characterization - to the CoreColl-4 made it possible to reach 100% of the allelic representativeness of the full *C. frutescens* collection (Table 3).

Core collections established for *Stylosanthes* species based on SSR variability reached the total allelic representativeness of the *S. macrocephala* and *S. capitata* individual collections (Santo-Garcia et al., 2012). The *S. macrocephala* core collection consisting of 23 accessions (17% of the original collection with 134 accessions) represented 100% of the allelic variability of the original collection; however, for *S. capitata* it was composed only by 13 accessions (7% of the original collection with 192 accessions). These results show the excellent potential of using SSR molecular markers to establish core collections and thus improve the management and utilization of the germplasm.

The number of clusters obtained using morphological descriptors and molecular markers was similar in the present research; however, comparing the two methodologies revealed differences in accession subdivisions, indicating the importance

and complementarity of using different analyses to estimate genetic diversity and composition of core collections.

There is much discussion on which descriptors data are suitable in developing a core collection - morphological, genetic, molecular, ecogeographic, etc. A total of 96 accessions morphologically identified as *C. frutescens* evaluated in this work presented high variability for several morphological traits such as fruit size, fruit weight, fruit persistence and incidence of viruses. Likely, much of that diversity was captured in the accessions selected for the CoreColl-2 based on morphological descriptors that can be effectively used by breeding programs. It is noteworthy that the morphological characterization data were relevant for identifying the genetic variability in *C. frutescens*, since the CoreColl-2 retained 73% of the original allelic variability (Table 3). In addition, the correlation between molecular and morpho-agronomic analyses performed in the present study was about 66% and statistically significant.

Martins et al. (2015) established and compared in terms of representativeness, core collections obtained from 67 tomato accessions of the Vegetable Germplasm Bank of Viçosa Federal University (BGH-UFV), using 19 quantitative morphological characters, 30 multi-categorical characters, 52 ISSR *loci*, reaction to three pathogens and also one core collection that contemplated all this information simultaneously. They concluded that when data of different natures are available, priority should be given to the establishment of core collections based on integration of the whole data, as these were more representative.

According to Nicolaï et al. (2013), compilation of the *Capsicum* phenotypic data (6 primary traits: flowering date, stem length, number of leaves, fruit length, fruit width, fruit wall thickness) and genotypic data (genotyping with 28 SSR *loci*) for genetic association studies, allowed the establishment of a core collection of *C. annuum* with 332 accessions, having 97% of the genetic and phenotypic diversity of the complete collection (908 accessions). In addition, several core collections were established using SSR alleles, ranging in size from 8 to 128 accessions and representing 37 to 90% of the allelic variability of *C. annuum* and its wild relatives (var. *glabriusculum*).

Sampling strategies with different clustering methods were employed by Lee et al. (2016) in the development of a *Capsicum* core collection. The authors performed a population structure analysis in a large *Capsicum* germplasm collection consisting of 3,821 accessions by applying 48 genome-wide SNPs, and selected a core set using the SNP data together with data for 32 morphological traits. Use of either genotypic or phenotypic information, only, for selection of core collection entries were not efficient for capturing genetic diversity of the entire germplasm collection. When only genotypic data were used, they demonstrated insufficient coverage of the phenotypic variation of the entire collection. Nevertheless, when phenotypic data for 32 traits were included for selection of the core sets, the representativeness of the phenotypic variation slightly increased. The selection of the core set using genotypic and phenotypic data together after clustering analysis showed to be the best methodology. Different from Lee et al. (2016), Gu et al. (2019) stated that a core collection based on genotype is more representative than that based on phenotypic data. In the present research, the most

promising core collection considering allelic representativeness – CoreColl-4 – was based on SSR and incidence of viruses.

Viruses are among the most important and complex disease for species in the genus *Capsicum* and may result in significant yield losses (Lima et al., 2017). Serological test results of accessions of CoreColl-4, revealed that plants of CNPH 597, CNPH 3698, CNPH 3885, CNPH 3894 and CNPH 3944 were infected only with tospoviruses (e.g. GRSV), while for CNPH 3715, solely potyviruses (PepYMV; PVY) were identified. The natural occurrence of PMMoV was verified in 25% of the samples, suggesting that virus dissemination is expressive in *C. frutescens*. PMMoV transmission in the field is mainly due to planting of contaminated seeds and spreading from plant to plant occurs during handling of plants.

According to Bhattacharjee et al. (2007), the millet core collection can most effectively be used as a starting point for breeding programs, involving research into screening the germplasm collection for sources of desirable characteristics, as well as photoperiod sensitivity, disease resistance, drought tolerance and adaptation to saline or alkaline environments due to conservation of genetic variability in these accessions for most traits. It also provides a guideline for the curator when purchasing new accessions for the collection.

The establishment of core collections should be considered a dynamic process, with continuous evaluation of new accessions and incorporation of additional information when available. The CoreColl-4 proposed for *C. frutescens* is expected to increase emphasis on genetic resources exploration and therefore the efficiency of the *C. frutescens* breeding program carried out at Embrapa Vegetables, contributing effectively to the development of new cultivars that meet consumer's demand.

C. frutescens core collections were established using SSR and morphological characterization, as well as resistance to viruses. The best approaches for establishment of a core collection (CoreColl-4, consisting of 13 accessions) was found to be SSR and incidence of viruses, which included 77% of the genetic variability found in the full collection of *C. frutescens*. In addition, CoreColl-4 presented significative geographic representativeness (53.8% North, 30.8% Northeast and 15.4% Center-West Brazilian Regions). CoreColl-4 is the first *Capsicum* core collection available from the Embrapa breeding program.

ACKNOWLEDGMENTS

This research was supported by Universidade de Brasília (UnB), Embrapa (Brazilian Agricultural Research Corporation), and by the Brazilian National Council for Science and Technology Development (CNPq).

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

AVGRIS (AVRDC Vegetable Genetic Resources Information System). Genetic Resources and Seed Unit. Database: AVGRIS [internet]. Accessed at [http://seed.worldveg.org/] October 1, 2019.

Baral JB and Bosland PW (2004). Unraveling the species dilemma in *Capsicum frutescens* and *C. chinense* (Solanaceae): a multiple evidence approach using morphology, molecular analysis, and sexual compatibility. *J. Am. Soc. Hortic. Sc*i. 129: 826-832.

Bhattacharjee R, Khairwal IS, Bramel PJ and Reddy KN (2007). Establishment of a pearl millet [*Pennisetum glaucum* (L.) R. Br.] core collection based on geographical distribution and quantitative traits. *Euphytica.* 155: 35-45. https://doi.org/10.1007/s10681-006-9298-x.

Bianchetti LB and Carvalho SIC (2005). Subsídios à coleta de germoplasma de espécies de pimentas e pimentões de gênero *Capsicum* (Solanaceae) In: Fundamentos para a coleta de germoplasma vegetal (Walter BMT and Cavalcanti TB, eds.). Embrapa Recursos Genéticos e Biotecnologia, Brasília.

Brown AHD (1989). Core collections: a practical approach to genetic resources management. *Genome*. 31: 818-824.

Carvalho SIC, Ragassi CF, Bianchetti LB, Reifschneider FJB, et al. (2014). Morphological and genetic relationships between wild and domesticated forms of peppers (*Capsicum frutescens* L. and *C. chinense* Jacquin). *Genet. Mol. Res.* 13: 7447-7464. http://dx.doi.org/10.4238/2014.September.12.11.

Carvalho SIC, Bianchetti LB, Ragassi CF, Ribeiro CSC, et al. (2017). Genetic variability of a Brazilian *Capsicum frutescens* germplasm collection using morphological characteristics and SSR markers. *Genet. Mol. Res.* 16 (3): gmr16039689. http://dx.doi.org/10.4238/gmr16039689.

Clark MF and Adams AN (1977). Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *J. Gen. Virol*. 34: 475-483. https://doi.org/10.1099/0022-1317-34-3-475.

DeWitt D and Bosland PW (1997). Pepper of the world: an identification guide. 1st edn. Ten Speed Press, Berkeley.

Eshbaugh WH (1979). Biosystematic and evolutionary study of the *Capsicum pubescens* complex. *Natn. Geogr. Soc. Res. Rep.* 1970. Projects: 143-162.

Faleiro FG (2007). Marcadores genético-moleculares aplicados aos programas de conservação e uso de recursos genéticos. Embrapa Cerrados, Planaltina.

Ferreira ME, Moretzsohn MC and Buso GSC (2007). Fundamentos da caracterização molecular de germoplasma vegetal. In: Recursos genéticos vegetais (Nass LL, ed.). Embrapa Recursos Genéticos e Biotecnologia, Brasília.

Frankel OH (1984). Genetic perspective of germplasm conservation. IN: Genetic manipulation impact on Man and Society (Arber WK, Llimensee K, Peacock WJ and Starlinger P, eds.). Cambridge University Press, Cambridge.

GRIN-USDA ARS, National Genetic Resources Program. Germplasm Resources Information Network. Database: GRIN [internet]. Accessed at [http://www.ars-grin.gov/ npgs/] October 1, 2019.

Gu X, Cao Y, Zhang Z, Zhang B, et al. (2019). Genetic diversity and population structure analysis of *Capsicum* germplasm accessions. *J. Integr. Agric.* 18: 1312-1320. https://doi.org/10.1016/s2095-3119(18)62132-X.

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

Jarret RL, Baldwin E, Perkins B, Bushway R, et al. (2007). Diversity of fruit quality characteristics in *Capsicum frutescens*. *Hort. Science.* 42: 16-19.

Lee HY, Ro NY, Jeong HJ, Kwon JK, et al. (2016). Genetic diversity and population structure analysis to construct a core collection from a large *Capsicum* germplasm. *BMC Genet.* 17: 142. https://doi.org/10.1186/s12863-016-0452-8.

Lima MF, Inoue-Nagata AK, Reifschneider FJB, Ferraz RM, et al. (2011a). Virus incidence in domesticated and semidomesticated field-grown hot peppers (*Capsicum* spp.). *Acta Hortic.* 917: 285-297. https://doi.org/10.17660/ *ActaHortic*.2011.917.40.

Lima MF, Inoue-Nagata AK, Reifschneider FJB, Souza KRR, et al. (2011b). Detection, occurrence and natural incidence of Pepper mild mottle virus (PMMoV) in hot peppers in Brazil. *Acta Hortic.* 917: 269-273. https://doi.org/10.17660/ ActaHortic.2011.917.37.

Lima MF, Carvalho SIC, Ragassi, CF, Bianchetti LB, et al. (2017). Characterization of a pepper collection (Capsicum frutescens L.) from Brazil. *Genet. Mol. Res.* 16: gmr16039704.http://dx.doi.org/10.4238/gmr16039704.

Martins KC, Pereira TNS, Souza SAM, Rodrigues R, et al. (2015). Crossability and evaluation of incompatibility barriers in crosses between *Capsicum* species. *Crop Breed. Appl. Biotechnol*. 15: 139-145.

Martinez ALA, Araújo JSP, Ragassi CF, Buso GSC, et al. (2017). Variability among *Capsicum baccatum* accessions from Goiás, Brazil, assessed by morphological traits and molecular markers. *Genet. Mol. Res*. 16: gmr16039074.

Mongkolporn O and Taylor PWJ (2011). *Capsicum*. In: Wild Crop Relatives: Genomic and Breeding Resources (Kole C, ed.). Springer, New York.

Mongkolporn O, Hanyong S, Chunwongse J and Wasee S (2015). Establishment of a core of chilli germoplasm using microsatellite analysis. *Plant Genet. Resour.* 13(2): 104-110. https://doi.org/10.1017/S1479262114000768.

Nicolaï M, Cantet M, Lefebvre V, Sage-Palloix AM, et al. (2013). Genotyping a large collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for the wild origin of cultivated *C. annuum* and the structuring of genetic

diversity by human selection of cultivar types. *Genet. Resour. Crop. Evol*. 60: 2375-2390. https://doi.org/10.1007/s10722-013-0006-0.

Reifschneider FJB, Nass LL and Henz GP (2015). Uma pitada de biodiversidade na mesa dos brasileiros. 1st edn. Brasília. Available: https://issuu.com/cica/ docs/uma_pitada_de_biodiversidade.

Reifschneider FJB, Lopes CA and Ribeiro CSC (2016). Continuity, focus and impact: a commented historical perspective on Embrapa Vegetables' extended *Capsicum* breeding program. *Hortic. Bras.* 34: 155-160. http://dx.doi.org/10.1590/S0102-053620160000200002.

Santo-Garcia MO, Silva GT, Rodrigo Sassaki RP, Ferreira TH, et al. (2012). Using genetic diversity information to establish core collections of *Stylosanthes capitata* and *Stylosanthes macrocephala*. *Genet. Mol. Biol*. 35 (4): 847-861. https://dx.doi.org/10.1590%2FS1415-47572012005000076.

van Hintum TJL, Brown AHDC, Spillane C and Hodgkin T (2000). Core collections of plant genetic resources. *IPGRI Tech. Bull. Int. Plant Genetic Res. Inst.* 3: 1-51.

Zewdie Y, Tong N and Bosland P (2004). Establishing a core collection of *Capsicum* using a cluster analysis with enlightened selection of accessions. *Genet. Resour. Crop Evol*. 51: 147-151.

