Analysis of genetic diversity of *Cucurbita moschata* (D.) germplasm accessions from Mesoamerica revealed by PCR SSCP and chloroplast sequence data

N. Barboza [a,*], F.J. Albertazzi [a,b], J.A. Sibaja-Cordero [b], F. Mora-Umaña [c], C. Astorga [d], P. Ramírez [a,b]

[a] *Centro de Investigación en Biología Celular y Molecular, Universidad de Costa Rica, Ap. 2060 San José, Costa Rica*
[b] *Escuela de Biología, Universidad de Costa Rica, San José, Costa Rica*
[c] *Servicio Fitosanitario del Estado, Ministerio de Agricultura y Ganadería, Ap. 2060 San José, Costa Rica*
[d] *Centro Agronómico Tropical de Investigación y Enseñanza Cartago, Turrialba 30501, Costa Rica*

ABSTRACT

The aim of this research was the genetic characterization of 218 accessions of *Cucurbita moschata* Duchesne, a squash, and its relationship with morphological characteristics of agronomic interest, which are part of the international collection conserved at Tropical Agricultural Research and Higher Education Center (CATIE), Costa Rica. The majority of the accessions came from Mexico and Central America; single genotypes from Curaçao, Colombia, Peru and the Russian Federation were also included. The polymerase chain reaction (PCR) and single strand conformation polymorphism analysis (SSCP) were used for the analysis of the regions amplified with ITS1–ITS2 nuclear primers and t*RNL*–F chloroplast primers. Haplotypes were constructed according to band patterns in SSCP gels. Twenty-five haplotypes were found using the ITS1–ITS2 markers, and 24 haplotypes were found with the t*RNL*–F markers. Unique haplotypes were found with both markers. Two individuals of each t*RNL*–F haplotype were sequenced. The results indicated a high level of genetic diversity in CATIE squash collection. Using ITS1–ITS2 primers, it was found that the number of haplotypes was independent of the geographical source of the accession, and haplotypes were distributed randomly throughout the study area. Mexico had the highest values of total heterozygosity (HE), genetic diversity (*H*) and Shannon index (*I*) while Panama showed the lowest values. Sequences obtained from t*RNL*–F intergenic marker showed the highest diversity index values were present in the group of additional sequences and Mexico, and lower values were observed for Nicaragua, Guatemala and Panama. PCoA based on morphological data showed three groups and by ANOSIM (*R*) all group differences were significant. Results obtained in this study suggest that high diversity is a characteristic of *C. moschata* from Mesoamerica.

1. Introduction

In the new world, the genus *Cucurbita* (pumpkin, gourd and zucchini) is composed of 12–14 species found from the United States to Argentina (Nee, 1990; Lira and Caballero, 2002). Five of these species were domesticated before European contact and represented an important source of food in the native economies of America (Sauer, 1950; Andres, 1990; Nee, 1990; Merrick, 1990; Decker-Walters and Walters, 2000). *Cucurbita moschata* (squash) is one of the most important vegetables in tropical areas. The mature and young fruits, flowers, seeds, and the tops of young stems are used for human consumption (National Research Council, 1989; León, 2000). *C. moschata* is cultivated principally in small

subsistence gardens or for sale in local markets. These farms have handled the majority of the world production of squash for decades (Gwanama et al., 2000).

Squash is among the most commonly cultivated species of the West of India and the lowlands of Central and South America. To date, a wild ancestor of *C. moschata* is not known (Nee, 1990; Sanjur et al., 2002). Colombia has been proposed as a center of origin because of the great diversity of fruits observed there (Zhiteneva, 1930). The possible hybridization of cultivars with wild species of *Cucurbita* in Colombia has been suggested (Nee, 1990). Robinson and Decker-Walkers (1997) and Decker-Walters and Walters (2000) mentioned that *C. moschata* has been cultivated in Mesoamerica since pre-Colombian times and that it was possibly domesticated in Mexico and South America independently.

Few molecular analyses to estimate genetic relationships in squash have been conducted (Ferriol et al., 2004; Gwanama et al., 2000). Characterization of regional collections of *C. moschata* in germplasm banks is necessary in order to define strategies for the

* Corresponding author. Tel.: +506 2511 3193; fax: +506 2511 3190.
E-mail addresses: natalia.barboza@ucr.ac.cr, nabargas@hotmail.com
(N. Barboza).

collection and preservation of genetic variability, and to generate information useful in genetic improvement programs. Markers based on DNA such as RAPD (Random Amplified Polymorphic DNA) have been used to analyze the genetic diversity of *C. moschata* in Korea and southern Africa (Youn and Chung, 1998; Baranek et al., 2000; Gwanama et al., 2000). The aim of this investigation was to determine the genetic variability of 218 accessions of *C. moschata* belonging to the germplasm bank of the Tropical Agricultural Research and Higher Education Center (CATIE), Turrialba, Costa Rica. Conformational variants in DNA sequences were amplified using the polymerase chain reaction (PCR) and visualized in non-denaturing polyacrylamide gels (PCR-SSCP) (Orita et al., 1989). The intergenic regions of genes 18S–5.8S–26S were amplified by nuclear primers ITS1 and ITS2 (Mullis and Faloona, 1987), and the chloroplastic intergenic region was amplified using chloroplast primers t*RN*L–F (Taberlet et al., 1991). Then DNA of selected haplotypes was sequenced.

2. Materials and methods

2.1. Plant material

Plant material used in this study was obtained from the germplasm bank at the Tropical Agricultural Research and Higher Education Center (CATIE), Turrialba, Costa Rica. The majority of the selected 218 accessions of *C. moschata* were collected in Mexico and Central America. Individual samples from Curaçao, Colombia, Peru and the Russian Federation, and an accession of *Cucurbita argyrosperma* (Nicaragua 12025, Table 5) were also included. Seeds from the 218 accessions were planted in pots using two seeds per accession. Twenty-two days after sowing, seedlings were collected, labeled and stored at −70 °C. One gram of leaf tissue was used for isolation of nucleic acid.

2.2. DNA extraction and isolation

Nucleic acids were extracted using the CTAB method (Doyle and Doyle, 1987), quantified using a Thermo Spectronic Helios γ spectrophotometer (Thermo Fisher Scientific, Inc.) and treated with ribonuclease A (100 mg/ml) (QIAGEN). The DNA was quantified again, and visualized in 1% agarose gels stained with ethidium bromide (TAE 1×, 95 V for 40 min).

2.3. Molecular markers

Total DNA was used in the polymerase chain reaction (PCR) to amplify a fragment of approximately 500 bp using rDNA primers ITS1–ITS2 (Mullis and Faloona, 1987) and a 500 bp fragment using chloroplast DNA primers t*RN*L–F (Taberlet et al., 1991). The PCR mixture (10 μl) contained 30–50 ng of total DNA, 100 ng/μl of primers, 2.5 mM MgCl$_2$, 0.25 mM of dNTPs, 1× reaction buffer solution and 0.05 units of DNA polymerase (Fermentas Hamover, MD) The mixture was initially denatured at 94 °C for 5 min, then amplified by 5 cycles at 94 °C for 1 min, 48 °C for 45 s, and 72 °C for 3 min, followed by 23 cycles at 94 °C for 1 min, 55 °C for 45 s, 72 °C for 2 min, and a final extension at 72 °C for 7 min in a PTC-100 thermal cycler (MJ Research, Inc., MI). Following amplification, 4 μl of loading buffer (15% Ficoll, 0.25% Bromophenol blue, 0.25% xylene cyanol), 2.65 μl urea (5 M) and 9.85 μl TBE 1× were added to each sample. Amplified DNA products were denatured for 5 min at 100 °C and placed on ice. The PCR fragments were analyzed in 10% non-denaturing polyacrylamide gels (PCR-SSCP 29:1) (Bio-Rad, CA) in 0.5× TBE at 200 V. Denatured products amplified by primers ITS1 and ITS2 (10 μl) and t*RN*L–F (15 μl) were separated during 65 min

and 55 min, respectively. Gels were stained with silver nitrate (Caetano-Anollés et al., 1994).

2.4. Data analysis

The Visual Basic Application for Excel (VBA) GenAlEx vrs. 6.0 (Peakall and Smouse, 2006) was used to construct a matrix of genetic distances for the haplotypes obtained from the pattern of bands observed in the PCR-SSCP gels of the intergenic ITS marker. Principal coordinate analysis (PCoA) was used to classify and detect structure in the relationships between the accessions, and a three-dimensional graph was constructed with Systat vrs. 11.0 (SYSTAT 11, 2004). The haplotype frequencies were calculated using 218 accessions of *C. moschata*. Total heterozygosity (HE), genetic diversity (*H*) (Nei, 1973) and Shannon index (*I*) (Lewontin, 1972) indices were determined with the software Popgene vrs. 1.31 (Yeh and Boyle, 1999). Geographical zones with only a single individual were eliminated.

Accessions with geographic data were used to compare genetic and linear geographic distance matrices by Mantel test (1967) (GenAlEx6) (Peakall and Smouse, 2006), using 9999 permutations and a significance level of 0.05 (33 from Costa Rica, 17 from El Salvador, 34 from Guatemala, 37 from Honduras, 26 from Mexico, eight from Nicaragua and 17 from Panama). An analysis of molecular variance (AMOVA) (Excoffier et al., 1992) (GenAlEx6) (Peakall and Smouse, 2006) was performed to estimate the percentage of variation among and within locations.

The combination of the primers t*RN*L–F (Taberlet et al., 1991), t*RN*V–M and t*RN*S–T (Demesure et al., 1995), *atp*B-*rbc*L and *atp*F intron (Weising and Gardner, 1999), t*RN*H-*psb*A (Hamilton, 1999), and *rps*19 (Skinner, 2000) was tested using PCR-SSCP gels, with some samples selected by random.

The haplotypes were selected based on the pattern of bands observed in the PCR-SSCP gels according to the chloroplast markers. The VBA GenAlEx vrs. 6.0 (Peakall and Smouse, 2006) was used to construct a matrix of genetic distances for the haplotypes obtained from the pattern of bands observed. The 218 accessions were analyzed using the UPGMA method with 300 replacements. The Mantel test (GenAlEx6) (Peakall and Smouse, 2006) was performed to compare genetic and linear geographic distance matrices, using 9999 permutations and a significance level of 0.05.

Two individuals were selected randomly from each haplotype generated by PCR-SSCP. The chloroplast t*RN*L–F region was sequenced with the 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequences were aligned using the algorithm from the CLUSTAL X tool in BioEdit vrs. 7.0 (Hall, 2005), and then improved manually. Accessions from Curaçao, Peru and the Russian Federation were placed in a single group as additional sequences. DnaSP software vrs. 4.10.9 (Rozas et al., 2003) was used to calculate polymorphic haplotypes without considering insertions or deletions. Heterozygosity and nucleotide diversity indices Theta (Hom), Theta (*S*) Theta (*k*), and Theta (Phi) were also calculated.

Analysis of molecular variance AMOVA (Excoffier et al., 1992) was performed with Arlequin vrs. 3.01 (Excoffier, 2007) using 100 permutations and a significance level of 0.05. For all analyses, accessions were grouped according to source country (Costa Rica, Guatemala, Honduras, Nicaragua, and Panama). Genetic distance was calculated according to K2P (Kimura, 1980) and a dendrogram was constructed using UPGMA with the software Treecon (Van de Peer and De Wachter, 1994) with 500 replacements. A tree was constructed considering insertions and deletions. The accession from Nicaragua 12025 (Nic27-6) (*C. argyrosperma*), with haplotype 6, was used as an outgroup. To establish the relationship between accessions, a network map was built with the polymorphic chloroplastic sequences using the program Network vrs. 4.01 (Bandelt et al., 1995, 1999).

Table 1
Number of haplotypes and unique haplotypes in each group using ITS1–ITS2 and tRNL–F markers.

	Primer	Population							
		Costa Rica	El Salvador	Guatemala	Honduras	Mexico	Nicaragua	Panama	Outgroup
Accessions analyzed		36	17	46	42	38	15	20	3
Number of haplotypes	ITS1–ITS2	10	5	12	13	10	7	4	1
Unique haplotypes		0	0	1	0	0	2	0	0
Number of haplotypes	tRNL–F	8	5	11	8	12	6	5	3
Unique haplotypes		0	1	1	1	1	1	0	1

The Genebank (http://www.ncbi.nlm.nih.gov/genbank) number for the accessions are: 2005 (EU262283), 6302 (EU262285), 7036 (EU262284), 7907 (EU262282), 8675 (EU262279), 8901 (EU262280), 8903 (EU262281), 9021 (EU262287), 20119 (EU262278), 8702 (EU262288), 11290 (EU262275), 15048A (EU262289), 15763A (EU262299), 15730B (EU262269), 12269 (EU262291), 12148E (EU262298), 12145 (EU262265), 12139A (EU262295), 8004 (EU262267), 8005 (EU262271), 9213 (EU262268), 9233 (EU262264), 18916A (EU262272), 18943A (EU262276), 18943B (EU262270), 18946A (EU262273), 18946B (EU262266), 9317A (EU262274), 12029A (EU262293), 12035A (EU262290), 12481 (EU262296), 10700 (EU262294), 10785 (EU262297), 19218 (EU262277), 12025 (EU262292), 14185 (EU262300), 5713 (EU262301).

2.5. Agronomic and morphological characterization

Seven quantitative characteristics based on morphological and agronomic data previously generated by CATIE were used in this study: rind thickness, peduncle length, fruit width, fruit length, number of seeds per fruit, weight of 100 seeds, and internode length. In addition, 47 qualitative characteristics were evaluated, including: cotyledon size and color, growth habit, stem shape, tendrils, leaf shape, color and size, marbling, lobes, pubescence of the dorsal and ventral leaf surfaces, flower color, type of sex, time to maturity, transverse peduncle shape, the union of peduncle and fruit abscission, fruit color, shape and yield, rind color intensity, flesh color, design, texture and hardness of the rind, the outer shape of the peduncle, fruit apex, fruit luster, rib shape, fruit size and fruit storage characteristics, color, intensity, moisture, flavor and texture of the flesh, amount of placenta, ease of separation of pulp and seed, seed size, proportion of seeds per pulp (class), surface of the seed and seed surface luster, seed coat color, seed margin and seed margin color.

Distances for morphological data were made with the Gower index (1967) using the MVSP (Kovach Computing Services, 2006), for this test only the accessions with sequences were considered. The accessions CR18-22, CR13-2, M3-20, M1-16, M2-17, N28-11 were excluded for lack of some morphological data. Additionally, a PCoA biplot based on the similarity matrix of the accessions (Gower index) and the morphological data was constructed to show the correlation between these two data sets. The score of the morphological variant vectors was scaled with the Pearson coefficient (for numerical variants) or Spearman rank correlation (for the nominal data) by their standard deviation between the standard deviations of the accession's scores for each principal coordinate axis (Legendre and Gallagher, 2001). Also the similarity matrix of accessions (Gower index) was analyzed using ANOSIM (R) (Clarke and Warwick, 1994), with 10,000 permutations. The value of R determines the level of difference; $R = 1$ when accession groups are completely different from each other.

Genetic distances were calculated using the K2P (Kimura, 1980) model for sequences obtained with primers for chloroplast tRNL–F. The Mantel test was used for comparison of quantitative and qualitative morphological data and chloroplast sequences. To evaluate the association between quantitative and qualitative morphological data and DNA sequences, a dendrogram was constructed with the program Mesquite vers. 1.11 (Maddison and Maddison, 2005), using the method of maximum parsimony and a maximum of 100 trees for the three matrices. Accession 12025 from Nicaragua (*C. argyrosperma*) was used as an outgroup.

Tree confrontation, or mirror analysis, was used to find an association between genetic distances of the sequences and morphological characters. 61 polymorphic positions were considered and the positions of most parsimony (402, 496, 500 and 507) were selected. In addition, three morphological characteristics (variation in fruit size, fruit storagability and the number of seeds per fruit) were chosen to represent the agronomic variation.

3. Results

3.1. ITS1–ITS2 marker

Electrophoretic haplotypes generated with the intergenic marker ITS1–ITS2 by the PCR-SSCP technique were used to determine the genetic variability among the accessions from Mexico, Central America, Colombia, Peru, Russian Federation and Curaçao. Twenty-five haplotypes were found distributed in patterns of 1–17 bands (Table 1). The number of haplotypes was independent of the geographical source of the accession, and haplotypes were distributed randomly throughout the study area. Unique haplotypes were found in three of the seven locations. Specifically, haplotypes 13 and 16 were present in Nicaragua, and haplotype 24 was found in Guatemala.

Accessions from Mexico and Honduras presented 12 of the 25 haplotypes generated for the intergenic marker ITS. The accessions from other countries showed fewer haplotypes. Accessions from all the countries with the exception of Panama shared the haplotype one. The most prevalent was the nine in accessions from the Mesoamerican area, with the exception of Honduras and El Salvador. Accessions 11919B (Costa Rica), 7769 (El Salvador), 8704 and 11290 (Guatemala), and 10970 and 10973 (Honduras) were not amplified by the ITS primers.

Principal coordinate analysis (PCoA) of the ITS1–ITS2 haplotypes explained 70.18% of the variability between accessions. The first component explained 30.04% of the variability; the second and third components explained 23.68% and 16.46% of the observed variability, respectively. The dispersion of the accessions showed three groups, $p = 0.4$. The first group was characterized by the presence of a large number of accessions of *C. moschata*, mainly from Nicaragua, El Salvador and Mexico. The second group consisted mostly of accessions from Honduras, Costa Rica and Guatemala and the third group was composed of accessions from Panama, Mexico and Guatemala (data not shown).

Haplotype frequency values of the accessions ranged from 0.023 to 0.550 (Table 2). This showed a considerable variation in the genetic material analyzed. In the specific case of Mexico accessions, a low value of 0.026 was observed for haplotype 22, six and one. Panama accessions presented the lowest value (0.05) for haplotype four, intermediate (0.10) for haplotype 10, and the highest value

Table 2
Haplotype frequencies with the ITS1–ITS2 primer using 218 accessions of *C. moschata* from Mesoamerica.

Haplotype	Costa Rica	El Salvador	Guatemala	Honduras	Mexico	Nicaragua	Panama
1	0.257	0.375	0.023	0.150	0.026	0.000	0.000
2	0.000	0.438	0.068	0.025	0.053	0.200	0.000
3	0.029	0.125	0.000	0.025	0.000	0.000	0.000
4	0.286	0.000	0.250	0.325	0.026	0.000	0.050
5	0.000	0.000	0.023	0.025	0.000	0.000	0.000
6	0.000	0.063	0.114	0.025	0.026	0.000	0.000
7	0.000	0.000	0.000	0.150	0.000	0.000	0.000
8	0.029	0.000	0.045	0.175	0.000	0.000	0.300
9	0.029	0.000	0.364	0.000	0.211	0.133	0.550
10	0.000	0.000	0.023	0.000	0.000	0.000	0.100
11	0.000	0.000	0.045	0.000	0.000	0.000	0.000
12	0.000	0.000	0.000	0.025	0.105	0.267	0.000
13	0.000	0.000	0.000	0.000	0.000	0.067	0.000
14	0.000	0.000	0.000	0.000	0.105	0.067	0.000
15	0.000	0.000	0.000	0.025	0.000	0.200	0.000
16	0.000	0.000	0.000	0.000	0.000	0.067	0.000
17	0.171	0.000	0.000	0.000	0.000	0.000	0.000
18	0.086	0.000	0.000	0.000	0.000	0.000	0.000
19	0.000	0.000	0.023	0.000	0.079	0.000	0.000
20	0.000	0.000	0.000	0.000	0.053	0.000	0.000
21	0.000	0.000	0.000	0.000	0.184	0.000	0.000
22	0.057	0.000	0.000	0.025	0.026	0.000	0.000
23	0.000	0.000	0.000	0.025	0.105	0.000	0.000
24	0.000	0.000	0.023	0.000	0.000	0.000	0.000
25	0.057	0.000	0.000	0.000	0.000	0.000	0.000

(0.55) for haplotype nine. Mexico had the highest values of total heterozygosity (HE), genetic diversity (*H*) (Nei, 1973) and Shannon index (*I*) (Lewontin, 1972), while Panama showed the lowest values (Table 3).

Comparison of the genetic distance matrix with geographical data in the Mantel test showed a value of *r* = 0.052, with a significant probability *p* = 0.010, but close to zero. The majority of the haplotypes were distributed throughout the study region. An analysis of molecular variance (AMOVA) for the ITS markers showed a 12% variation between the origin of the accessions and 88% variability among accessions, with a significant probability *p* = 0.010 (data not shown).

3.2. tRNL–F marker

A pattern of high diversity was observed with the tRNL–F (Taberlet et al., 1991) primer combination using the PCR-SSCP technique. Analysis of the chloroplastic tRNL–F intergenic region by the PCR-SSCP technique generated 24 haplotypes (Fig. 1). Mexico showed the highest number of haplotypes for this marker with a total of 12. This result was consistent with ITS data. Haplotype two was common to all countries of the Mesoamerican region, with the exception of Panama. Some accessions presented haplotypes unique to their country of origin, such as haplotype seven in Colombia, eight in Guatemala, nine in Honduras, 10 in Nicaragua, 13 in Mexico and 24 in El Salvador (Table 1).

A dendrogram of the haplotypes of the 218 selected accessions did not show defined groupings that would establish a marked difference between the evaluated materials (data not shown). A Mantel test was performed and no relationship was detected (*r* = 0.100), with a significant probability *p* = 0.010.

3.3. Sequencing

Selected accessions for sequencing the tRNL–F intergenic region were clustered in six groups according to the source of the accessions: Mexico, Costa Rica, Guatemala, Honduras, and Panama. Unique haplotypes were excluded from the analysis of sequences. Additionally, Curaçao and Peru were included because they shared haplotype 18 with some Mesoamerican accessions. These sequences and the Russian Federation accession were named as additional sequences in the analysis.

A total of 29 haplotypes were found among the 38 accessions sequenced (Table 5). Haplotype six was present in the Mexico 6-12, Costa Rica 12-2, Guatemala 23-3, Guatemala 24-4, Guatemala 25-19, Guatemala 26-11, Nicaragua 28-11, Nicaragua 27-6 and Honduras 34-15 accessions.

Diversity indices Theta (Hom), Theta (*k*), Theta (*S*) and Theta (Phi), were obtained for each population (Table 4). The highest diversity index values were present in the group of additional sequences and Mexico. Lower values were observed for Nicaragua, Guatemala and Panama, respectively. AMOVA analysis explained 7.04% of the variation between groups and 92.96% within groups. This result is consistent with the observed variation between and within accessions using the intergenic marker ITS.

The dendrogram (Fig. 2) showed no clear separation between accessions from Central America, Mexico and Peru. The Russian Federation accession F11-17 was markedly distant from the others. The sequence of the accession from Nicaragua Nic27-6 was used as an outgroup, however, it presented haplotype six, obtained with the program DnaSP, and was placed in the center of the minSpanningnetwork clustering generated by the program Network (data not shown).

Table 3
Heterozygosity (HE), genetic diversity (*H*) (Nei, 1973), Shannon index (*I*) (Lewontin, 1972) and the total obtained for every analysis using rDNA intergenic marker ITS.

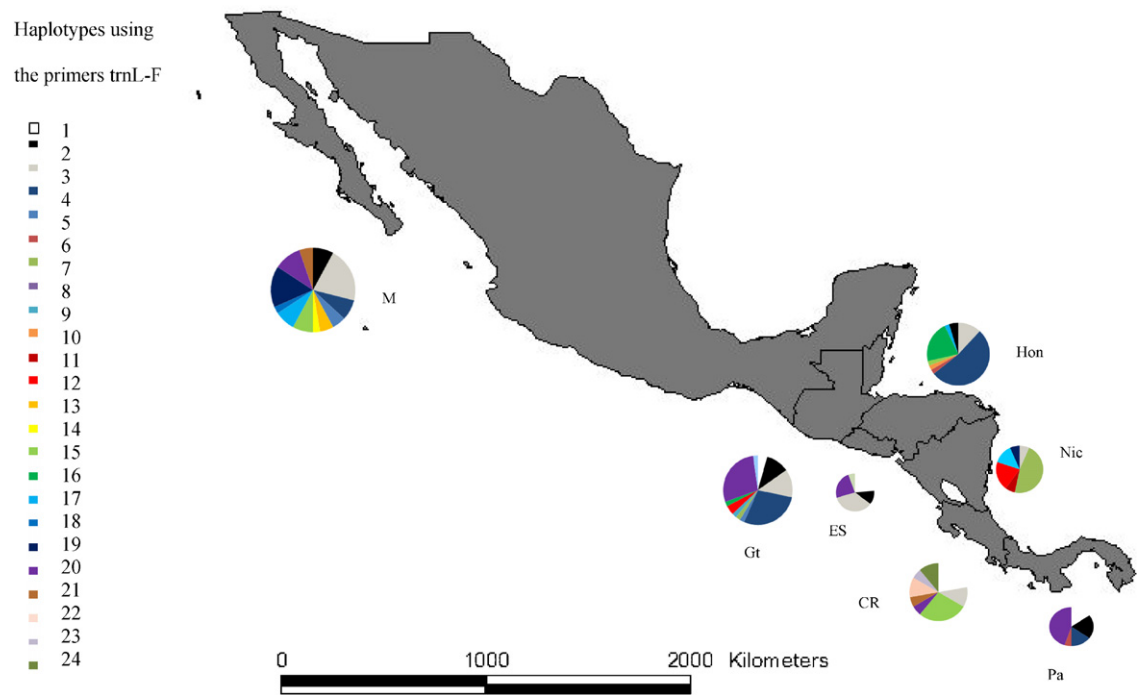| Marker	Population	Costa Rica	El Salvador	Guatemala	Honduras	Mexico	Nicaragua	Panama	Total
ITS1–ITS2	HE	0.807	0.648	0.781	0.814	0.874	0.818	0.595	0.762
	H	0.0628	0.0527	0.0614	0.0638	0.0672	0.0629	0.0458	0.0698
	I	0.1087	0.0822	0.1093	0.1147	0.1222	0.1042	0.0701	0.1422

Fig. 1. Estimation of genetic diversity of the accessions based on haplotypes obtained with the chloroplast t*RN*L–F primer.

3.4. Morphological data

The PCoA (Fig. 3) showed three groups of accessions based on morphological characteristics. In the same way that genetic data, the accessions cannot be grouped by the source sites. The analysis of similarities (ANOSIM), determinates the difference of these groups in base to their quantitative and nominal morphological characteristics in a 67% ($R = 0.675$, $p < 0.0001$). The group in quadrat II of the PCoA contained accessions CR14-21, Gt25-19, Mx4-4, Pa33-5 CR16-23, and CR21-23, and showed higher values or ranges for the following characteristics: leaf shape, leaf size, weight of 100 seeds, rind thickness and hardness of the rind. This group differed from the group in quadrat IV of the PCoA in 81% of the morphological data (pairwise comparison ANOSIM, $p < 0.001$). The group in quadrat IV had the accessions Mx8-18, Pr38-18, CR17-22, CR20-14, and CR19-14, and presented higher values or ranges for time to maturity, fruit color, flesh color, internode length, amount of placenta, and fruit. The last group was located mainly in quadrats I and III of the PCoA and contained samples Ho34-15, Gt35-15, F11-17, CR12-2, Mx9-12, Ho36-16, Mx5-20, Gt23-3, Mx10-19, Ni27-6, Ni30-6, Ho29-5, Cu37-18, Mx7-18, Gt26-11, Pa31-1, and Pa31-1. This group presented higher values or ranges for seed coat color, seed margin, flower color, pulp color, growth habit, stem shape, texture of the flesh, pubescence of the dorsal leaf surfaces, the outer shape of the peduncle, leaf lobes, peduncle length, proportion of pulp. seeds per. Morphological data for this

group showed differences with groups in quadrats II and IV in 61% and 70%, respectively (pairwise comparisons ANOSIM, $p < 0.001$).

A Mantel test was conducted to compare continuous and discrete morphological characteristics and genetic distances of the 516 bp sequences obtained with the chloroplast marker t*RN*L–F. No significant relationship between the evaluated variables and the sequences was observed ($p = 0.110$ and $r = 0.113$). The morphological continuous characteristics were grouped in two clusters. The accession from Nicaragua 12025 (*C. argyrosperma*) was used as an outgroup. The first group included accessions CR21-23, CR16-23, Gt22-3, Gt35-3 and N30-6. Accessions CR18-22, CR19-14, CR12-2, CR15-21, P33-5, M10-14, M8-18, CR17-22, H36-16 and H34-15 were clustered in the second group. Using discrete morphological characteristics, accessions N30-6, CR19-14, CR17-22, CR18-22, Pr38-18, Cu37-18, M8-15 were grouped together. The other group included accessions M9-12, M5-20, Gt26-19, CR12-2, P33-5, H36-16, H34-15, and M6-12. The dendrogram did not show a defined group, but accessions H36-16 and H34-15 were always grouped together.

Groups generated with the chloroplast t*RN*L–F sequences were used to perform a tree confrontation analysis, in which all polymorphic positions and morphological features were considered. Sequence positions 402 and 496 showed variation between accessions and the discontinuous morphological feature of variability in fruit size, which is of agronomic interest. At position 402, all four substitutions were present.

Table 4
Theta indices, Theta (S), Theta (Phi), Theta (Hom), Theta (k), using sequences obtained with the primer t*RN*L–F, for accessions of *C. moschata* from Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, and additional sequences. NA, not available.

Marker	Population	Costa Rica	El Salvador	Guatemala	Honduras	Mexico	Nicaragua	Panama	Additional sequences
t*RN*L–F	S	7.0	NA	2.6	5.3	8.1	2.0	4.0	17.3
	Phi	5.4	NA	2.0	5.3	5.7	2.0	4.3	18.3
	Hom	NA	NA	0.3	NA	NA	NA	NA	NA
	k	NA	NA	0.6	NA	NA	NA	NA	NA

Table 5
List of accessions, source countries, observed haplotypes with primers ITS1–ITS2 and t*RNL*–F, abbreviations cited in the text, and sequence code.

Accession	Country	Haplotype ITS	Haplotype LF	Sequence code
2005	Costa Rica	4	22	CR18-22
11919B	Costa Rica	0	0	
6302	Costa Rica	4	14	CR20-14
6372	Costa Rica	18	0	
7036	Costa Rica	9	14	CR19-14
7432	Costa Rica	1	2	
7856	Costa Rica	1	0	
7905	Costa Rica	1	20	
7907	Costa Rica	17	22	CR17-22
8172	Costa Rica	4	0	
8254	Costa Rica	4	14	
8412	Costa Rica	4	14	
8414	Costa Rica	1	23	
8544	Costa Rica	17	0	
8575	Costa Rica	4	14	
8674	Costa Rica	4	0	
8675	Costa Rica	22	2	CR13-2
8676	Costa Rica	25	23	
8677	Costa Rica	17	23	CR16-23
8900	Costa Rica	1	14	
8901	Costa Rica	18	21	CR14-21
8903	Costa Rica	1	21	CR15-21
8904	Costa Rica	25	14	
8980	Costa Rica	8	0	
9021	Costa Rica	1	23	CR21-23
9908	Costa Rica	17	20	
9954	Costa Rica	18	0	
9956	Costa Rica	17	21	
10036	Costa Rica	4	14	
10810	Costa Rica	17	19	
11919A	Costa Rica	1	19	
20032	Costa Rica	4	14	
20119	Costa Rica	3	2	CR12-2
20122	Costa Rica	22	14	
20124	Costa Rica	4	21	
9907	Costa Rica	1	2	
7764	El Salvador	3	2	
7767	El Salvador	2	0	
7768	El Salvador	2	2	
7769	El Salvador	0	0	
7771	El Salvador	2	2	
7774	El Salvador	2	0	
9055	El Salvador	1	1	
9060	El Salvador	3	2	
9061	El Salvador	1	2	
9063	El Salvador	1	19	
9065	El Salvador	2	1	
9076	El Salvador	2	0	
9077	El Salvador	2	19	
9082	El Salvador	1	24	
9090	El Salvador	6	19	
9093	El Salvador	1	19	
9100	El Salvador	1	2	
8034	Guatemala	9	1	
5985	Guatemala	6	19	
6288	Guatemala	24	2	
8688	Guatemala	4	19	
5986	Guatemala	2	2	
5992	Guatemala	9	2	
5994	Guatemala	9	19	
5996	Guatemala	5	3	
6001	Guatemala	9	1	
6301	Guatemala	9	3	
7753	Guatemala	9	19	
7757	Guatemala	9	3	
7758	Guatemala	9	2	
7759	Guatemala	2	11	
7761	Guatemala	6	19	
7783	Guatemala	9	3	
8026	Guatemala	6	19	
8028	Guatemala	9	0	
8324	Guatemala	4	19	
8325	Guatemala	4	3	
8702	Guatemala	9	3	Gt22-3
8704	Guatemala	0	3	
9420B	Guatemala	4	3	

Table 5 (*Continued*)

Accession	Country	Haplotype ITS	Haplotype LF	Sequence code
9427	Guatemala	4	19	
10594	Guatemala	4	19	
10599	Guatemala	9	2	
11173	Guatemala	9	1	
11217	Guatemala	6	1	
11290	Guatemala	0	19	Gt25-19
11427	Guatemala	6	19	
11440	Guatemala	8	3	
11543B	Guatemala	9	3	
14912A	Guatemala	4	3	
15048A	Guatemala	4	3	Gt23-3
15048B	Guatemala	4	1	
14971	Guatemala	10	19	
15163	Guatemala	4	19	
16373	Guatemala	9	2	
9421B	Guatemala	1	0	
15763A	Guatemala	4	15	Gt35-15
15747	Guatemala	2	8	
12270A	Guatemala	11	3	
15730B	Guatemala	19	4	Gt24-4
9435	Guatemala	11	20	
9434A	Guatemala	9	6	
12269	Guatemala	8	11	Gt26-11
6498	Honduras	8	2	
6499	Honduras	8	1	
6504	Honduras	4	2	
6505	Honduras	8	2	
6507	Honduras	7	14	
6508B	Honduras	4	2	
6509	Honduras	1	2	
10863	Honduras	4	2	
10887	Honduras	1	1	
10890	Honduras	8	2	
10893	Honduras	7	14	
10915	Honduras	7	14	
10926	Honduras	8	2	
10970	Honduras	0	2	
10979	Honduras	4	2	
10983	Honduras	5	2	
10988	Honduras	2	1	
11026	Honduras	4	2	
11038	Honduras	1	2	
11048	Honduras	1	14	
11060	Honduras	6	14	
11072	Honduras	7	14	
11078	Honduras	4	2	
11088	Honduras	1	2	
11100	Honduras	4	2	
11116	Honduras	7	14	
11117	Honduras	3	14	
11118	Honduras	4	16	
11122	Honduras	4	2	
11779A	Honduras	8	2	
12148D	Honduras	1	2	
12148E	Honduras	4	15	H34-15
14914A	Honduras	7	14	
15124A	Honduras	4	1	
15124B	Honduras	4	3	
12140	Honduras	8	2	
12187	Honduras	4	2	
10973	Honduras	0	2	
12088A	Honduras	22	1	
12145	Honduras	23	16	H36-16
12123B	Honduras	15	9	
12139A	Honduras	12	5	H29-5
15030A	Mexico	14	2	
7997	Mexico	9	2	
8003	Mexico	9	2	
8004	Mexico	2	20	M3-20
8005	Mexico	19	18	M8-18
8016	Mexico	9	2	
9152	Mexico	21	19	
9207	Mexico	23	18	
9213	Mexico	20	4	M4-4
9227	Mexico	21	14	
9233	Mexico	23	16	M1-16
9284A	Mexico	12	3	

Table 5 (*Continued*)

Accession	Country	Haplotype ITS	Haplotype LF	Sequence code
9317C	Mexico	19	1	
14902A	Mexico	12	16	
15001A	Mexico	12	2	
18846A	Mexico	14	3	
18859B	Mexico	9	19	
18868B	Mexico	9	2	
18874A	Mexico	21	4	
18916A	Mexico	9	12	M6-12
18918A	Mexico	23	18	
18927A	Mexico	19	16	
18929A	Mexico	22	1	
18934A	Mexico	21	3	
18934C	Mexico	14	14	
18943A	Mexico	14	17	M2-17
18943B	Mexico	20	18	M7-18
18946A	Mexico	9	12	M9-12
18946B	Mexico	21	20	M5-20
14900	Mexico	9	2	
14968	Mexico	21	18	
15021	Mexico	23	18	
18879	Mexico	21	13	
9294	Mexico	12	14	
9351	Mexico	2	19	
9285A	Mexico	1	1	
18927B	Mexico	6	2	
9317A	Mexico	4	19	M10-19
12011C	Nicaragua	12	10	
12029A	Nicaragua	2	6	N30-6
12029C	Nicaragua	15	16	
12035A	Nicaragua	15	11	N28-11
12035C	Nicaragua	2	6	
12036A	Nicaragua	9	6	
12036C	Nicaragua	13	6	
12058B	Nicaragua	12	6	
12060B	Nicaragua	16	11	
12061B	Nicaragua	9	6	
11992	Nicaragua	14	16	
11994	Nicaragua	12	18	
12063	Nicaragua	2	6	
12066	Nicaragua	12	11	
15052	Nicaragua	15	2	
10789A	Panama	9	3	
12444A	Panama	8	19	
12488B	Panama	9	0	
10743	Panama	8	1	
12465	Panama	8	0	
12478	Panama	10	19	
12481	Panama	8	1	P31-1
7223	Panama	9	19	
7269	Panama	9	19	
7270	Panama	10	19	
7312	Panama	8	19	
10697	Panama	9	3	
10700	Panama	8	5	P33-5
10724	Panama	9	19	
10740	Panama	9	19	
10750	Panama	9	1	
10777	Panama	9	3	
10789B	Panama	9	19	
10785	Panama	9	1	P32-1
10778	Panama	4	0	
19218	Russian Federation	1	17	F11-17
14185	Curaçao	1	18	Cu37-18
5713	Peru	1	18	P38-18
11682B	Colombia	1	7	
12025	Nicaragua	15	6	Nic27-6

Sequence position 496 (data not shown) presented a G in accessions M9-12, CR16-23 and Gt22-3; F11-17 showed a G, and only P31-1 showed an A. The rest of the accessions had a deletion, which suggests the possibility that the alternation of bases (A and G) could be due to a common origin. The accession of the Russian Federation (F11-17) presented a T at position 500, in contrast to Costa Rica (CR16-23) which showed an A. The rest of the accessions were characterized by having a G in this position, with the exception of Mexico (M9-12) and Guatemala (Gt22-3), which had a C. The morphological character named fruit storage was classified in three stages (first stage one week, second stage six weeks and the third 12 weeks). Costa Rica 17-22 accession presented the first
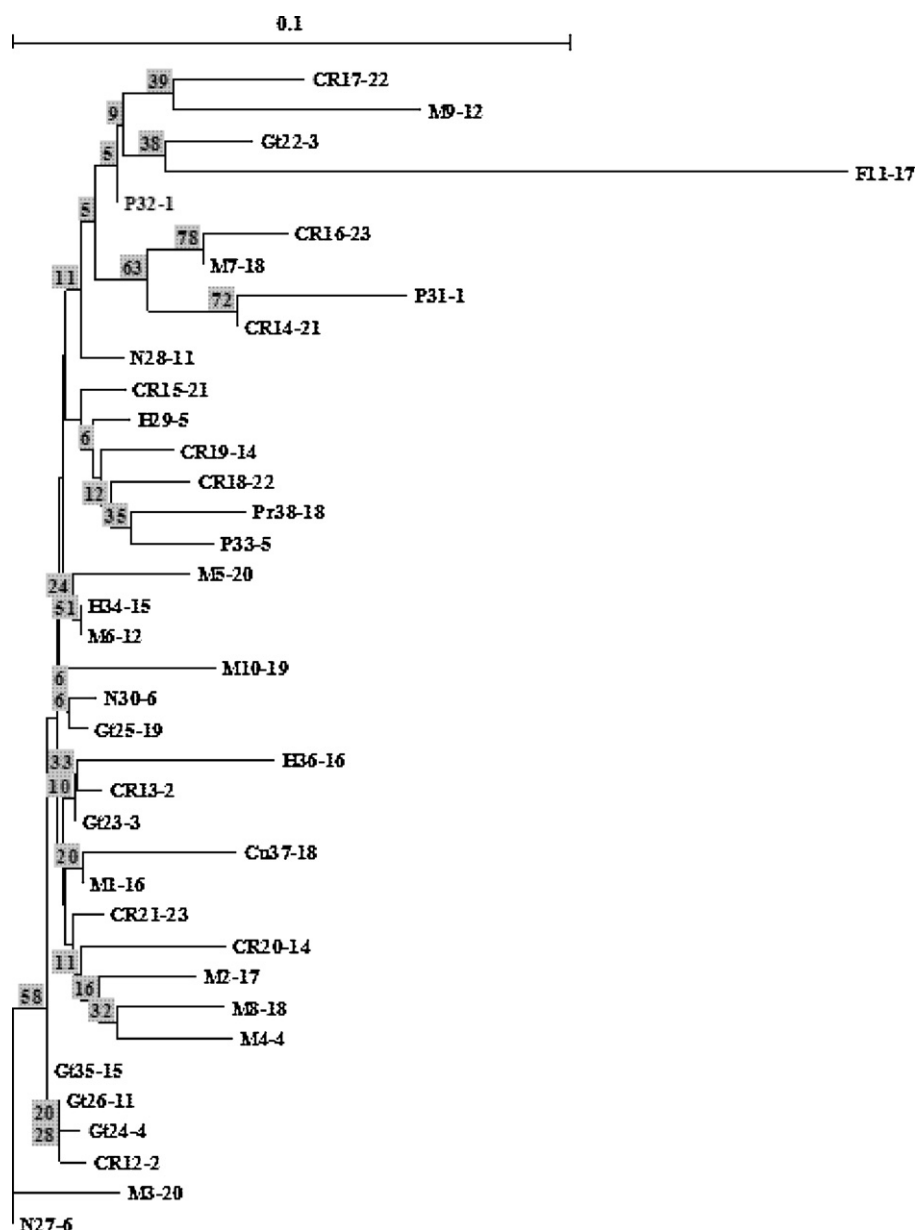
Fig. 2. Cluster with the genetic distance of K2P (Kimura, 1980), using the methodology of UPGMA and 500 replacements with sequences obtained with the marker for chloroplast t*RN*L–F, considering 51 nt. Only insertions and deletions were considered. Shaded numbers represent the number of repetitions.

stage. Five accessions from different countries (Costa Rica 19-4, Peru 38-18, Honduras 29-5, Nicaragua 30-6 and Curaζao 39-18), share the second stage. All the rest of accessions presented the third stage.
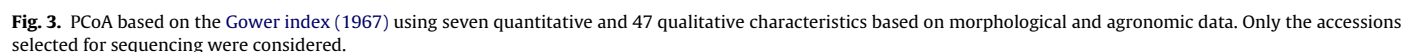
In mirror tree analyses of position 402 of the chloroplast sequences and the morphological quantitative character of number of seeds per fruit, the dendrogram on the left shows that accession CR20-14 from Costa Rica had a T, in contrast to most of the accessions, in which there was an A. The highest value for the morphological characteristic of number of seeds per fruit was found for accession H29-5 from Honduras; accessions from Costa Rica (CR21-23), the Russian Federation (11-17), Nicaragua (N30-6), Curaζao (Cu37-18) and Panama (P31-1) showed the lowest values. No association between position 402 and the number of seeds per fruit was observed (data not shown).

4. Discussion

4.1. ITS and tRNL–F markers

Germplasm banks are used for the conservation and management and use of plant genetic resources. This study is the first genetic characterization of *C. moschata* accessions from Mesoamerica maintained in the germplasm collection of CATIE, Costa Rica. The genetic variability and relationships between accessions of squash from different locations in Mesoamerica were studied using the intergenic markers ITS1–ITS2 (Fig. 2), the chloroplast intergenic region t*RN*L–F and sequencing.

A wide distribution of haplotypes in the region of study was found. In most cases, the same haplotypes were present in accessions from various countries. A clear separation according to origin was not found. The open-pollination system in squash may provide

Fig. 3. PCoA based on the Gower index (1967) using seven quantitative and 47 qualitative characteristics based on morphological and agronomic data. Only the accessions selected for sequencing were considered.

an explanation for the observed variability (León, 2000). On the other hand, Chen et al. (2005) in their work with *Camellia sinensis* (L.) using RAPD markers, explained that evolutionary history and continuous interspecific hybridization allow for the presence of plant genetic resources with overlapping morphological, physiological, chemical and genetic characteristics.

Ferriol et al. (2004) used AFLP and SRAP to group squash cultivars from the Institute for the Preservation and Improvement of Valencian Agro-diversity (COMAV) of the Polytechnic University of Valencia in Spain and samples from America. Using both markers, most of the accessions were grouped according to source location, and 12 samples from Central and South America were differentiated as well. One possible explanation is that most of the production of this crop has been maintained by farmers in the same region for hundreds of years, and possibly only a few accessions were brought to Europe from America in the 16th century (Paris et al., 2006). This may have caused a bottleneck effect, resulting in the observed differences in accessions from different locations.

Studies in southern Africa revealed four groups among landraces of *C. moschata*, where groups I, II and III consisted of individuals found only in Malawi. Samples from Zambia were grouped apart in group IV. These dissimilarities revealed a low variability in accessions of squash grown in South Central Africa, which may be due to the high frequency of seed exchange among farmers from the same ethnic and geographical areas of study. The authors recommended using genotypes from America to expand the molecular diversity of their accessions (Gwanama et al., 2000).

One explanation for the differences between earlier works with other markers and the present research may be the fact that genetic material from Mesoamerica was used in this study. Nee (1990), Sanjur et al. (2002) and Piperno and Stohert (2003) suggested in previous studies that the center of origin and domestication of *C. moschata* is located in northern South America. Decker-Walters and Walters (2000) and Ferriol et al. (2004) postulated the existence of two sites of independent domestication for squash, one in Mexico and another in South America, but this idea still is uncertain (Pickersgill, 2007). Therefore, greater diversity in samples from these areas of America was expected. It is also important to consider

the type of markers used, as AFLP and RAPD provide different views of genome variation compared with ITS1, ITS2 and t*RNL*–F primers.

In this study, the lowest heterozygosity values were present in accessions from Panama (HE: 0.595), while those of Mexico (HE: 0.874) showed the highest. The values obtained for heterozygosity (HE), Nei genetic diversity (*H*) and the Shannon diversity index (*I*) were higher in this study (Table 3) than in other work with the genus *Cucurbita* using other molecular markers and different source locations.

Isoenzymes were used to estimate the structure and gene flow in *C. argyrosperma* spp. *sosoria*, *C. argyrosperma* spp. *argyrosperma* and *C. moschata* in Mexico (Montes-Eguiarte and Hernandez, 2002). HE values of 0.416 and 0.407 were obtained for *C. argyrosperma* ssp. *sosoria* and HE: 0.391 for *C. argyrosperma* spp. *argyrosperma*. Decker-Walters et al. (1990) in Mexico, indicated a value of $M = 0.052$ for *C. moschata*. In Spain, Ferriol et al. (2003) worked with AFLP in *Cucurbita pepo* and found diversity indices ranging from 0.09 to 0.48, with an average of 0.25 per primer combination. These values confirm the high genetic diversity of the accessions used in this work. Although different markers were used in the present study, the results are comparable (see Nybom, 2004) and were similar to or greater than those described previously. Results obtained in this study help to explain the number of haplotypes from accessions of squash from Mesoamerica, and suggest that high diversity is a characteristic of *C. moschata*.

A wide distribution of haplotypes of the chloroplast marker was also observed in the region of study. Most of the haplotypes were present in different Mesoamerican countries (Fig. 1). These findings are consistent with those identified by AMOVA, where 92.96% variability within accessions was observed. No relationship was found between genetic and geographical distances using the Mantel test. These results suggest that gene flow between individuals of *C. moschata*, and other processes such as migration and out-crossing in squash plants, are likely throughout the region.

Unique ITS1 haplotypes were present in accessions from Costa Rica, Guatemala, Mexico and Nicaragua, while unique cpDNA haplotypes were found in Colombia, Guatemala, Honduras, Nicaragua, Mexico, Costa Rica and El Salvador (Table 1). Similar examples of

unique haplotypes were found in accessions of *Spondias purpuera* collected in Central America and characterized from chloroplast sequences amplified in the intergenic region trnG–trnS. Miller and Schaal (2005) suggested that the presence of unique haplotypes can be the result of an incomplete sampling system or that new alleles have emerged.

Genetic diversity of the accessions was estimated based on sequences obtained with the chloroplast tRNL–F. The highest diversity values were detected in samples from Curaζao, Peru and the Russian Federation, although fewer individuals in this group were sequenced. The ancestor of the Russian Federation sample probably presented a high number of polymorphic sites. Geographic isolation and differences in environmental conditions between the two continents could be factors that explain the results observed between accessions.

Although Nicaragua had the diversity lowest values, these results were greater than those reported in the literature for other cultivated species (Mohanty et al., 2002; Trapnell and Hamrick, 2004). According to Hamrick et al. (1991), large values of genetic variation may be associated with broad geographic ranges, interspecific hybridization and wind pollination, factors which are present for accessions of *C. moschata* from different sites in Mesoamerica.

A dendrogram was constructed using the sequences obtained with the chloroplast primers. It was constructed using the method of highest similarity (Fig. 2). This method grouped the accessions in a similar manner, and did not separate them by country.

The sequence of the *C. argyrosperma* accession (Nic27-6) was used as an outgroup in the cluster analysis since it was part of the germplasm bank of CATIE. It is worth mentioning that this accession (Nic27-6) shared haplotype six with eight other accessions. Moreover, this haplotype was located within the core group of accessions obtained by network analysis. Decker-Walters et al. (1990) state that squash was grown adjacent to *C. argyrosperma* Huber and *C. pepo* for decades in North America in the region between northern Mexico and the southern United States, and report a spontaneous hybridization between *C. mochata* and *C. argyrosperma*. Montes-Eguiarte and Hernandez (2002) have identified gene flow between squash, C. *argyrosperma* ssp. *argyrosperma* and *C. argyrosperma* ssp. sosoria in traditional agroecosystems in Mexico, where wild species often grow close to cultivated crops. Sanjur et al. (2002) found that accessions of *C. moschata* grouped with *C. sosoria* and *C. argyrosperma* accessions from America, indicating a close evolutionary relationship between these species. Hybridization may be one of the reasons why *C. argyrosperma* was observed to share haplotype 6 and was located within the circle of the main network sequences in this study (data not shown).

4.2. Agronomical characterization

A wide genetic diversity was found with markers ITS1 and ITS2, tRNL–F, sequences amplified with the chloroplast primers. These results were related to the morphological variability observed between accessions of *C. moschata*. No significant value was found in the Mantel test conducted among continuous and discrete agronomic traits and sequences obtained with the chloroplast tRNL–F primers. In addition, there was no association between qualitative and quantitative morphological data and sequences. A similar level of morphological characterization was found in the work of Ferriol et al. (2004) in cultivars of squash in Spain. Wide variability was found among the accessions, mostly in fruit size and shape, and rind color. These accessions presented a prostrate growth habit, and most could not be classified into the three types proposed (Castetter, 1925). Similar results were obtained by Wessel-Beaver (1998), where a large morphological diversity was found in *C. moschata* from Puerto Rico. These data are in contrast to those

obtained in Africa. Gwanama et al. (2001) worked with fruit characteristics of tropical squash from Zambia. However, none of the landrace accessions presented more than three desirable characteristics. The dissimilarities found in his study showed a low variability in the current crop of *C. moschata*, which is consistent with the molecular characterization using RAPD.

Fruit color and flavor were characteristics considered in this study, and there was a cluster of accessions for these characteristics. As explained by Gwanama et al. (2001), squash is grown in the tropics primarily for the consumption of ripe fruit. Fruit size, color and hardness, among other characteristics, are qualities subject to consumer preference. However, those characters showed high variation but were not grouped by PCoA, only rind hardness and thickness were relatively close. It would be interesting to extend more analyses based on those characteristics with accessions Gt25-19 and CR14-21. Correlation between morphological data and geographical origin is variable and depends on the crop; a correlation was described in *Artocarpus altilis* (Skeekumar et al., 2007) but was absent in *Capsicum annuum* (Geleta et al., 2005) and in *Solanum melongena* (Tuembelin et al., 2011). It is probable that intense exchange of materials in Mesoamerica and agricultural practices by small farmers contribute to maintaining such morphological variability. In most cases, squash is grown for subsistence or for sale in local markets (Ferriol et al., 2004). This could help to explain the morphological and genetic diversity found in this study.

Fruit color is one of the best-known features in *C. pepo* because of its usefulness for breeders (Paris, 2000). To date, 11 loci that affect the development of color have been identified. In the case of *Cucumis melo*, López-Sése et al. (2003) mentioned that there are distinctive morphological characteristics such as fruit texture and flavor that have avoided the introgression of genes from other germplasm from different sources, even in the absence of geographic isolation.

It is important to stress that the knowledge of the relationships between genotypes is essential for the proper use of germplasm, especially for species that are not well characterized (Gwanama et al., 2001). In addition, the ability to recognize available genotypes and their genetic diversity should be considered when starting a breeding program.

Acknowledgements

References

Andres, T.C., 1990. Biosystematics, theories on the origin and breeding potential of *Cucurbita ficifolia*. In: Bates, D.M., Robinson, R.W., Jeffrey, C. (Eds.), Biology and Utilization of the Cucurbitaceae. Cornell University Press, Ithaca, New York, pp. 102–119.

Bandelt, H.J., Forster, P., Sykes, B.C., Richards, M.B., 1995. Mitochondrial traits of human populations using median networks. Genetics 141, 743–753.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Baranek, M., Stift, G., Vollmann, J., Lelley, T., 2000. Genetic diversity within and between the species *Cucurbita pepo*, *C. moschata* and *C. maxima* as revealed by RAPD markers. Cucurbit Genet. Coop. Rep. 23, 73–77.

Caetano-Anollés, G., Bassam, B.J., Gresshoff, P.M., 1994. Buffer components tailor DNA amplification with arbitrary primers. Genome Res. 4, 59–61.

Castetter, E.F., 1925. Horticultural groups of cucurbits. Proc. Am. Soc. Hortic. Sci. 22, 338–340.

Chen, L., Gao, G.K., Chen, D.M., Xu, C.J., 2005. The use of RAPD markers for detecting genetic diversity, relationship and molecular identification of Chinese elite tea genetic resources *Camellia sinensis* (L.) O. Kuntze preserved in a tea germplasm repository. Biodivers. Conserv. 14, 1433–1444.

Clarke, K.R., Warwick, R.M., 1994. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. Bourne, England, p. 144.

Decker-Walters, D.S., Walters, T.W., Posluszny, U., Kevan, P.G., 1990. Genealogy and gene flow among annual domesticated species of *Cucurbita*. Can. J. Bot. 68, 782–789.

Decker-Walters, D.S., Walters, T.W., 2000. Squash. In: Klipe, K.F., Ornelas, K.C. (Eds.), The Cambridge World History of Food. Cambridge University Press, Cambridge, pp. 335–351.

Demesure, B., Sodzi, N., Petit, R.J., 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol. Ecol. 4, 129–131.

Doyle, J.J., Doyle, J.D., 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. Phytochem. Bull. 19, 11–15.

Excoffier, L., Smouse, P., Quaitro, J., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479–491.

Excoffier, L., 2007. Arlequin. An Integrated Software for Population Genetics Data Analysis., http://cmpg.unibe.ch/software/arlequin3.

Ferriol, M., Picó, B., Fernández, P., Nuez, F., 2003. Genetic diversity of a germplasm collection of Cucurbita pepo using SRAP and AFLP markers. Theor. Appl. Genet. 107, 271–282.

Ferriol, M., Picó, B., Fernández, P., Nuez, F., 2004. Molecular diversity of a germplasm collection of squash (Cucurbita moschata) determined by SRAP and AFLP markers. Crop Sci. 44, 653–664.

Geleta, L.F., Labuschagne, M.T., Viljoen, C.D., 2005. Genetic variability in pepper (Capsicum anuum L.) estimated by morphological data and amplified fragment length polymorphism markers. Biodivers. Conserv. 14, 2361–2375.

Gwanama, C., Labuschagne, M.T., Botha, A.M., 2000. Analysis of genetic variation in Cucurbita moschata by random amplified polymorphic DNA (RAPD) markers. Euphytica 113, 19–24.

Gwanama, C., Botha, A., Labuschagne, M., 2001. Genetic effects and heterosis of flowering and fruit characteristics of tropical pumpkin. Plant Breed. 120, 271–272.

Gower, J.C., 1967. A comparison of some methods of cluster analysis. Biometrics 23, 623–637.

Hall, T., 2005. BioEdit vrs 7.0.5. Biological Sequence Alignment Editor., http://www.mbio.ncsu.edu/BioEdit/bioedit.html.

Hamilton, M.B., 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol. Ecol. 8, 521–523.

Hamrick, J.L., Godt, M.J.W., Murawski, D.A., Loveless, M.D., 1991. Correlations between species traits and allozyme diversity: implications for conservation biology. In: Falk, D.A., Holsinger, K.E. (Eds.), Genetic and Conservation of Rare Plants. Oxford University Press, New York, pp. 75–86.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Kovach Computing Services, 2006. Multi Variate Statistical Package. Provalis Research, and Data Description Inc., Anglesey, Wales, All Rights Reserved. Portions Copyright Addinsoft.

León, J., 2000. Botánica de los cultivos tropicales, third ed. Editorial Agroamérica, San José, Costa Rica.

Legendre, P., Gallagher, E.D., 2001. Ecologically meaningful transformations for ordinations of species data. Oecologia 129, 271–280.

Lewontin, R.C., 1972. The apportionment of human diversity. Evol. Biol. 6, 381–398.

Lira, R., Caballero, J., 2002. Ethnobotany of the wild Mexican Cucurbitaceae. Econ. Bot. 56, 380–398.

López-Sése, A.I., Staub, J.E., Gómez-Guillamón, M.L., 2003. Genetic analysis of Spanish melon (Cucumis melo L.) germoplasm using a standardized molecular-marker array and geographically diverse reference accessions. Theor. Appl. Genet. 108, 41–52.

Mantel, N., 1967. The detection of disease clustering and generalized regression approach. Cancer Res. 27, 209–220.

Maddison, W., Maddison, D., 2005. Mesquite. A Modular System for Evolutionary Analysis., http://mesquiteproject.org/mesquite/mesquite.html.

Merrick, L.C., 1990. Systematics and evolution of a domestication squash, Cucurbita argyrosperma, and its wild and weedy relatives. In: Bates, D.M., Robinson, R.W., Jeffey, C. (Eds.), Biology and Utilization of the Cucurbitaceae. Cornell University Press, Ithaca, New York, pp. 77–95.

Miller, A., Schaal, B., 2005. Domestication of a Mesoamerican cultivated fruit tree, Spondias purpurea. Proc. Natl. Acad. U. S. A. 102, 12801–12806.

Mohanty, A.J., Martín, P., Aguinaldo, I., 2002. Population genetic analysis of European Prunus spinosa (Rosaceae) using chloroplast DNA markers. Am. J. Bot. 89, 1223–1228.

Montes-Eguiarte, S., Hernandez, L., 2002. Genetic structure and indirect estimates of gene flow in three taxa of Cucurbita (Cucurbitaceae) in western Mexico. Am. J. Bot. 89, 1156–1163.

Mullis, K., Faloona, F.A., 1987. Specific synthesis of DNA in vitro via polymerase-catalyzed chain reaction. Methods Companion Methods Enzymol. 155, 335–350.

National Research Council, 1989. Lost Crops of the Incas: Little Known Plants of the Andes with Promise for Worldwide Cultivation. National Academy Press, Washington, DC.

Nee, M., 1990. The domestication of Cucurbita (Cucurbitaceae). Econ. Bot. 44, 56–68.

Nei, M., 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U.S.A. 70, 3321–3323.

Nybom, H., 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol. Ecol. 13, 1143–1155.

Orita, M., Suzuki, Y., Sekiya, T., Hayashi, K., 1989. A rapid and sensitive detection of point mutations and genetic polymorphisms using polymerase chain reaction. Genomics 5, 874–879.

Paris, H., 2000. Quiescent Intense (qi): a gene that affects young but not mature fruit color intensity in Cucurbita pepo. J. Hered. 4, 333–339.

Paris, H., Daunay, M.C., Pitrat, M., Janick, J., 2006. First known image of Cucurbita in Europe. Ann. Bot. 98, 41–47.

Peakall, R., Smouse, E.P., 2006. GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6, 288–295.

Pickersgill, B., 2007. Domestication of plants in the Americas: insights from Mendelian and molecular genetics. Ann. Bot. 100, 925–940.

Piperno, D.R., Stohert, K.E., 2003. Phytolith evidence for early Holocene Cucurbita domestication in southwest Ecuador. Science 229, 1054–1057.

Robinson, R., Decker-Walkers, D., 1997. Cucurbits. CAB International, 226 pp.

Rozas, J., Sánchez-Del Barrio, J.C., Messegyer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Sanjur, O., Piperno, D., Andres, T., Wessel-Beaver, L., 2002. Phylogenetic relationships among domesticated and wild species of Cucurbita (Cucurbitaceae) inferred from a mitochondrial gene: implications for crop plant evolution and areas of origin. Proc. Natl. Acad. Sci. U. S. A. 99, 535–540.

Sauer, J.D., 1950. The grain amaranthus. A survey of their history and classification. Ann. Mo. Bot. Gard. 37, 561–632.

Skeekumar, V.B., Binoy, A.M., George, S.T., 2007. Genetic and morphological variation in breadfruit (Artocarpus altilis (Park.) Fosberg) in the Western Ghats on India using AFLP markers. Genet. Resour. Crop Evol. 54, 1659–1665.

Skinner, D.Z., 2000. Non random chloroplast DNA hypervariability in Medicago sativa. Theor. Appl. Genet. 101, 1242–1249.

SYSTAT, 2004. SYSTAT Software, Inc©Copyright.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of tree non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.

Trapnell, D., Hamrick, J., 2004. Partitioning nuclear and chloroplast variation at multiple spatial scales in the neotropical epiphytic orchid, Laelia rubescens. Mol. Ecol. 13, 2655–2666.

Tuembelin, Y., Frary, A., Mutlu, S., Doganlar, S., 2011. Genetic diversity in Turkish eggplant (Solanum melongena) varieties as determined by morphological and molecular analyses. Int. Res. J. Biotechnol. 2, 16–25.

Van de Peer, Y., De Wachter, R., 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft windows environment. Comput. Appl. Biosci. 10, 569–570.

Weising, K., Gardner, R.C., 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42, 9–19.

Wessel-Beaver, L., 1998. Pumpkin breeding with a tropical twist. Cucurbit Network News 5, 2–3.

Yeh, F.R., Boyle, T., 1999. POPGENE Version 1.32. Microsoft Window-Based Freeware for Population Genetic Analysis. University of Alberta, Edmonton, Canada.

Youn, S.J., Chung, H.D., 1998. Genetic relationship among the local varieties of Korean native squashes (Cucurbita moschata) using RAPD technique. J. Korean Soc. Hortic. Sci 39, 517–521 (In Korean, with English abstrac.).

Zhiteneva, N.E., 1930. The world's assortment of pumpkins. Trudi Prkl. Bot. Genet. Selek. 23, 157–207.