Genetic diversity in a cajuí (*Anacardium* spp.) germplasm bank as determined by ISSR markers

A.N.C. Borges[1], A.C.A. Lopes[2], F.B. Britto[3], L.F.L. Vasconcelos[4] and P.S.C. Lima[4]

1 Programa de Pós-Graduação em Genética e Melhoramento, Centro de Ciências Agrárias, Universidade Federal do Piauí, Ininga, Teresina, PI, Brasil
2 Departamento de Fitotecnia, Centro de Ciências Agrárias, Universidade Federal do Piauí, Teresina, PI, Brasil
3 Departamento de Biologia, Centro de Ciências da Natureza, Universidade Federal do Piauí, Teresina, PI, Brasil
4 Embrapa Meio-Norte, Teresina, PI, Brasil

Corresponding author: P.S.C. Lima
E-mail: paulo.costa-lima@embrapa.br

ABSTRACT. Some species of *Anacardium* (Anacardiaceae) produce fruits and pseudofruits that are smaller than those of the common cashew (*Anacardium occidentale*) and, for this reason, are known collectively in Brazil as "cajuí". Despite their economic value in the food market and their important environmental and ecological functions, cajuí trees remain underexploited. We employed nine inter-simple sequence repeat (ISSR) markers to characterize two presupposed populations of cajuí comprising 25 accessions maintained in the germplasm bank of Embrapa Meio-Norte (Teresina, PI, Brazil). Population structure and relationships between accessions were determined in order to generate knowledge that could contribute to genetic improvement programs and better management of this germplasm bank. A high degree of polymorphism (91.3%) was observed among the accessions. Analysis of molecular variance and Bayesian analysis demonstrated that the two presupposed populations were not genetically differentiated but constituted a single population containing highly diversified individuals including migrants,

migrant descendants and possible hybrids. Nonetheless, genetic variability within the accessions could be organized into two distinct, but linked, groups that had undergone extensive exchange of genetic material, as verified by the high gene flow index ($Nm = 13.145$). The substantial genetic variability observed could be attributed to individual differences between accessions rather than to differential spatial distribution. This report enhances our knowledge of the genus *Anacardium* and should facilitate the future improvement of cajuí culture and fruit quality. In addition, our study highlights the importance of further taxonomic studies on the species of *Anacardium* that comprise cajuí.

Key words: Indigenous Brazilian fruit tree; Population structure; Molecular characterization

INTRODUCTION

The genus *Anacardium* (Anacardiaceae) comprises 10 species of cashew trees that are indigenous to Brazil (Mitchell and Mori 1987). Some of these species, including *Anacardium nanum* A. St-Hil. and *A. humile* A. St-Hil., produce fruit (drupe or nut) and pseudofruit (inflated pedicel) that are smaller than those of the common cashew (*A. occidentale L.*) and, for this reason, are classified locally as "cajuí" rather than "caju" (Carbajal and Silva-Júnior 2003). Cajuí trees are dispersed throughout Brazil, from the Amazon to the Northeast, Center-West and Southeast regions, but their natural abundance is particularly high in coastal areas of the state of Piauí (PI) (Crespo and Souza 2014). Cajuí is of considerable socioeconomic importance to these coastal communities since the fruiting period of the plant occurs between the harvesting of seasonal crops (cereals and legumes) such as beans, rice and maize (Rufino 2004; Rufino et al. 2008; Almeida 2009; Crespo and Souza 2014). The pseudofruits are nutrient-rich, with high levels of sugar and phenolic compounds along with vitamin C, iron and phosphorus, and may be consumed *in natura* or processed as juices, jams or savory preserves, while the nuts are roasted and marketed as snacks or employed in the preparation of cakes and desserts. Alongside the economic value in the food market, cajuí trees play important environmental and ecological roles by stabilizing the sand dunes that provide habitat for plants and animals.

Despite its commercial potential, cajuí remains underexploited since, in the absence of a sustainable agricultural system, production relies almost entirely on extractivism. Moreover, little information regarding the biodiversity of cajuí is available. Considering that knowledge of the genetic variability within a species is a prerequisite for agricultural development, it is important to create and characterize active germplasm banks for the purpose of plant breeding, conservation and other research goals. In this context, inter-simple sequence repeat (ISSR) markers are among the most powerful molecular tools employed in the study of plant genetic variability, and the technique is both simple to use and cost efficient. Furthermore, ISSR-based methods allow the detection of polymorphism in microsatellite and intermicrosatellite loci without prior knowledge of the DNA sequence, and a large number of loci can be analyzed simultaneously even when working with different species in the same assay (Gupta et al. 1994; Reddy et al. 2002; Thimmappaiah et al.2009).

Considering the potential importance of cajuí, the aims of our study were: (i) molecular characterization of cajuí accessions maintained in the Active Germplasm Bank at Embrapa Meio-Norte (Teresina, PI, Brazil) using ISSR markers, and (ii) determination of population

structure and the relationships between accessions in order to generate knowledge that could contribute to future genetic improvement programs and better management of the germplasm bank.

MATERIAL AND METHODS

Plant material

The study focused on 25 accessions of cajuí originating from the municipalities of Ilha Grande (*n* = 8) and Parnaíba (*n* = 17) in Piaui state, and maintained at the Active Germplasm Bank of Embrapa Meio-Norte/Unidade de Execução de Pesquisa de Parnaíba (UEP). Considering that the two collection points were 8 km apart, it was assumed that the accessions belonged to two distinct populations. The study also included one accession of cajuí originating from municipality of Luzilândia, PI, and maintained at the Active Germplasm Bank of Embrapa Meio-Norte, and six accessions of authenticated *Anacardium* spp. originating from municipality of Pacajus, Ceará (CE), and maintained at the Active Germplasm Bank of Embrapa Agroindústria Tropical. Young leaves were collected from each of the accessions and submitted to analyses of genetic similarity, percentage polymorphism and clustering (Table 1). The cajuí accession from municipality of Luzilândia was introduced in the study as an outgroup (external group) for the purpose of comparison since this municipality is 100 km distant from municipalities of Ilha Grande and Parnaíba.

Table 1. Identification and origin of the accessions of cajuí and *Anacardium* spp.

No.	Genotype code	Species	Origin [a]	Maintaining institution[a]
1	BGCA 22	*Anacardium* sp.	Povoado Labino, Ilha Grande, PI	Embrapa Meio-Norte/UEP
2	BGCA 23	*Anacardium* sp.	Povoado Baixão, Ilha Grande, PI	Embrapa Meio-Norte/UEP
3	BGCA 24	*Anacardium* sp.	Povoado Cal, Ilha Grande, PI	Embrapa Meio-Norte/UEP
4	BGCA 30	*Anacardium* sp.	Povoado Labino, Ilha Grande, PI	Embrapa Meio-Norte/UEP
5	BGCA 39	*Anacardium* sp.	Povoado Labino, Ilha Grande, PI	Embrapa Meio-Norte
6	BGCA 43	*Anacardium* sp.	Povoado Labino, Ilha Grande, PI	Embrapa Meio-Norte
7	BGCA 44	*Anacardium* sp.	Povoado Labino, Ilha Grande, PI	Embrapa Meio-Norte
8	BGCA 45	*Anacardium* sp.	Povoado Labino, Ilha Grande, PI	Embrapa Meio-Norte
9	BGCA 25	*Anacardium* sp.	Pedra do Sal, Parnaíba, PI	Embrapa Meio-Norte/UEP
10	BGCA 26	*Anacardium* sp.	Pedra do Sal, Parnaíba, PI	EmbrapaMeio-Norte/UEP
11	BGCA 27	*Anacardium* sp.	Pedra do Sal, Parnaíba, PI	Embrapa Meio-Norte/UEP
12	BGCA 28	*Anacardium* sp.	Fazenda Bom Jesus, Parnaíba, PI	Embrapa Meio-Norte/UEP
13	BGCA 29	*Anacardium* sp.	Fazenda Bom Jesus, Parnaíba, PI	Embrapa Meio-Norte/UEP
14	BGCA 31	*Anacardium* sp.	EmbrapaMeio Norte/UEP, Parnaíba, PI	Embrapa Meio-Norte/UEP
15	BGCA 32	*Anacardium* sp.	EmbrapaMeio Norte/UEP, Parnaíba, PI	Embrapa Meio-Norte/UEP
16	BGCA 33	*Anacardium* sp.	Fazenda Bom Jesus, Parnaíba, PI	Embrapa Meio-Norte/UEP
17	BGCA 34	*Anacardium* sp.	Pedra do Sal, Parnaíba, PI	Embrapa Meio-Norte/UEP
18	BGCA 35	*Anacardium* sp.	Fazenda Bom Jesus, Parnaíba, PI	Embrapa Meio-Norte/UEP
19	BGCA 36	*Anacardium* sp.	Embrapa Meio Norte/UEP, Parnaíba, PI	Embrapa Meio-Norte/UEP
20	BGCA 37	*Anacardium* sp.	Embrapa Meio Norte/UEP, Parnaíba, PI	Embrapa Meio-Norte/UEP
21	BGCA 40	*Anacardium* sp.	Pedra do Sal, Parnaíba, PI	Embrapa Meio-Norte
22	BGCA 41	*Anacardium* sp.	Pedra do Sal, Parnaíba, PI	Embrapa Meio-Norte
23	BGCA 42	*Anacardium* sp.	Fazenda Bom Jesus, Parnaíba, PI	Embrapa Meio-Norte
24	BGCA 48	*Anacardium* sp.	Fazenda Bom Jesus, Parnaíba, PI	Embrapa Meio-Norte
25	BGCA 49	*Anacardium* sp.	Fazenda Bom Jesus, Parnaíba, PI	Embrapa Meio-Norte
26	BGCA 46[b]	*Anacardium* sp.	Luzilândia, PI	Embrapa Meio-Norte
27	*A. humile*[b]	*A. humile*	Embrapa Agroindústria Tropical, Pacajus, CE	Embrapa Agroindustria Tropical
28	*A. microcarpum*[b]	*A. microcarpum*	Embrapa Agroindústria Tropical, Pacajus, CE	Embrapa Agroindustria Tropical
29	*A. othonianum*[b]	*A. othonianum*	Embrapa Agroindústria Tropical, Pacajus, CE	Embrapa Agroindustria Tropical
30	CCP-06[b]	*A. occidentale*	Embrapa Agroindústria Tropical, Pacajus, CE	Embrapa Agroindustria Tropical
31	CCP-76[b]	*A. occidentale*	Embrapa Agroindústria Tropical, Pacajus, CE	Embrapa Agroindustria Tropical
32	*A. giganteum*[b]	*A. giganteum*	Embrapa Agroindústria Tropical, Pacajus, CE	Embrapa Agroindustria Tropical

[a] UEP, Unidade experimental de Parnaíba; PI, State of Piauí; CE, State of Ceará
[b] These accessions were not submitted to the full set of analyses described

DNA extraction and amplification by polymerase chain reaction (PCR)

Genomic DNA was extracted from fresh leaves (150g) using Qiagen (Venlo, Netherlands) DNEASY Plant Mini Kit following the recommendations of the manufacturer. DNA samples were amplified using primers developed by the University of British Columbia (UBC), Vancouver, Canada, nine of which were selected from an initial set of 32 based on the quality and resolution of the bands and presentation of high levels of polymorphism (Table 2). The PCR reaction mixture contained 1.0 · buffer [20 mM Tris-HCl (pH 8.0), 0.1 mM ethylenediaminetetraacetic acid, 1 mM dithiothreitol, 50% (v/v) glycerol], 2.0 mM $MgCl_2$, 0.8 mM dNTPs, 0.8 µM primer, 1 U Taq DNA polymerase (Invitrogen, Life Technologies do Brasil, São Paulo, SP, Brazil), 1.0 µL DNA template (7.0 ng/µL) and ultrapure distilled water to a final volume of 10 µL. Amplification reactions were carried out in a Veriti 96 Well Thermal Cycler (Applied Biosystems, Foster City, CA, USA) under the following conditions: initial denaturation for 1.5 min at 94 ºC, followed by 40 cycles each comprising denaturation for 40 s at 94 ºC, annealing for 45 s at the primer-dependent temperature specified in Table 2, extension for 2 min at 72 ºC, and final extension for 7 min at 72 ºC. The resulting amplicons were separated by electrophoresis on 1.5% agarose gel in 0.5x. TBE buffer for4h at 110 V, stained with 1 ·GelRed™ (Biotium, Hayward, CA, USA), visualized under a UV transilluminator and subsequently photographed. The sizes of amplicons were estimated by comparison with a 1 kb DNA ladder (Invitrogen).

Table 2. ISSR markers used in the amplification of genomic DNA from the accessions of cajuí and *Anacardium* spp.

Primer	Sequences [a] 5'-3'	Ta °C	Total loci	Polymorphic loci	
				n	%
808	(AG)₈C	54	11	8	72.7
825	(AC)₇A	56	7	7	100
836	(AG)₈YA	52	17	14	82.3
841	(GA)₈YC	51	15	15	100
845	(CT)₈RG	50	11	11	100
855	(AC)₈YT	56	12	12	100
856	(AC)₈YA	55	14	14	100
886	VDV(CT)₇	52	11	8	72.7
889	DBD(AC)₇	52	6	6	100
Total			104	95	91.3

[a] Y=C/T; R=A/G; B=C/G/T; D=A/G/T; H=A/C/T; V=A/C/G; N=A/G/C/T
Ta: annealing temperature

Statistical analyses

Analyses were performed with the aid of the PAST program version 1.34 (Hammer et al., 2001). Assuming that each amplicon generated by the same primer and occupying the same position in the gel represented a single locus, a binary matrix was created indicating the presence (1) or absence (0) of specific bands. From this matrix, the genetic similarities between the accessions of *Anacardium* were estimated using Sørensen–Dice indices and a dendrogram was constructed using the unweighted pair group method with arithmetic average (UPGMA) clustering technique to demonstrate the genetic relationships between accessions. The cophenetic correlation coefficient (*r*) and the bootstrap confidence index were calculated from the binary matrix of amplified fragments and the dendrogram after 1000 permutations.

Principal coordinates analysis (PCoA) was applied to visualize similarities or dissimilarities of data using GenAlEx version 6 software (Peakll and Smouse2006) and a bidimensional plot was constructed. Genetic variabilities were determined using: (i) Shannon diversity index *I* (Shannon and Weaver 1949) - calculated using POPGENE version 1.31 software (Yeh et al. 1999): (ii) Nei's genetic diversity index *h* considering the Taylor expansion (Lynch and Milligan 1994) - calculated using Tools for Population Genetic Analyses (TFPGA) software (Miller 1997): and (iii) mean heterozygosity value *H*according to the Bayesian method proposed by Zhivotovsky (1999) - calculated using AFLP-SURV version 1.0 software (Vekemans2002).

Analysis of molecular variance (AMOVA) was employed to estimate intra- and inter-populational genetic differentiation (Excoffier et al. 2007) using ARLEQUIN software version 3.1. The magnitude of genetic differentiation between the two presupposed populations was expressed by the coefficients of interpopulation genetic differentiation (G_{ST}) and gene flow index (N_m), the latter being the number of migrant alleles at independently inherited loci per generation, with both analyses being performed using POPGENE version 1.3.1.

Bayesian analysis (admixture model) was employed to determine the population structure of the accessions of cajuí using STRUCTURE software version 2.3.4 (Falush et al. 2007). Analyses were performed five times for each hypothetical number of subpopulations (K = 20) with 500,000 iterations in Markov chain Monte Carlo (MCMC) and 20,000 burn-in phases for each value of K. STRUCTURE HARVESTER version 0.6.94 (Earl and von Holdt2012) was used to estimate the most likely number of groups (K) using the delta K method (Evanno et al. 2005). STRUCTURE software functions were set as follows: (i) USEPOPINFO = 1: calculates the probability that each individual originated from the assumed population (the individual with a low probability is considered a possible migrant or hybrid according to Pritchard et al. 2000); (ii) GENSBACK = 3: calculates the ancestry of individuals up to three generations back; and (iii) MIGRPRIOR = 0.05: calculates the prior probability that an individual is a migrant.

RESULTS

ISSR analyses of the 32 accessions of cajuí and *Anacardium* spp. using the nine selected UBC primers generated 104 loci, of which 95 (91.3%) were polymorphic (Table 2). The mean number of loci per primer was 11.55, with primer 836 affording the highest number ($n = 17$) and primer 889 the lowest ($n = 6$). Six of the primers generated loci that were 100% polymorphic.

The mean Sorensen-Dice coefficient ranged between 0.2647 and 0.9176, with a mean value of 0.5508 ($\sigma = 0.1195$), demonstrating the existence of large variability among the 32 accessions and verifying that ISSR markers were highly suitable for detecting genetic diversity within the genus *Anacardium*. The cophenetic correlation coefficient ($r = 0.8591$) showed a good fit between the original data points of the genetic similarity matrix and the UPGMA dendrogram.

Examination of the UPGMA dendrogram (Figure 1) and the PCoA plot (Figure 2) revealed an interfusion between cajuí accessions of diverse origins and the different species of *Anacardium*, with accession BGCA36 being the most differentiated of all. The clustering pattern also showed that genetic differentiation between the two presupposed populations was either absent or minimal.
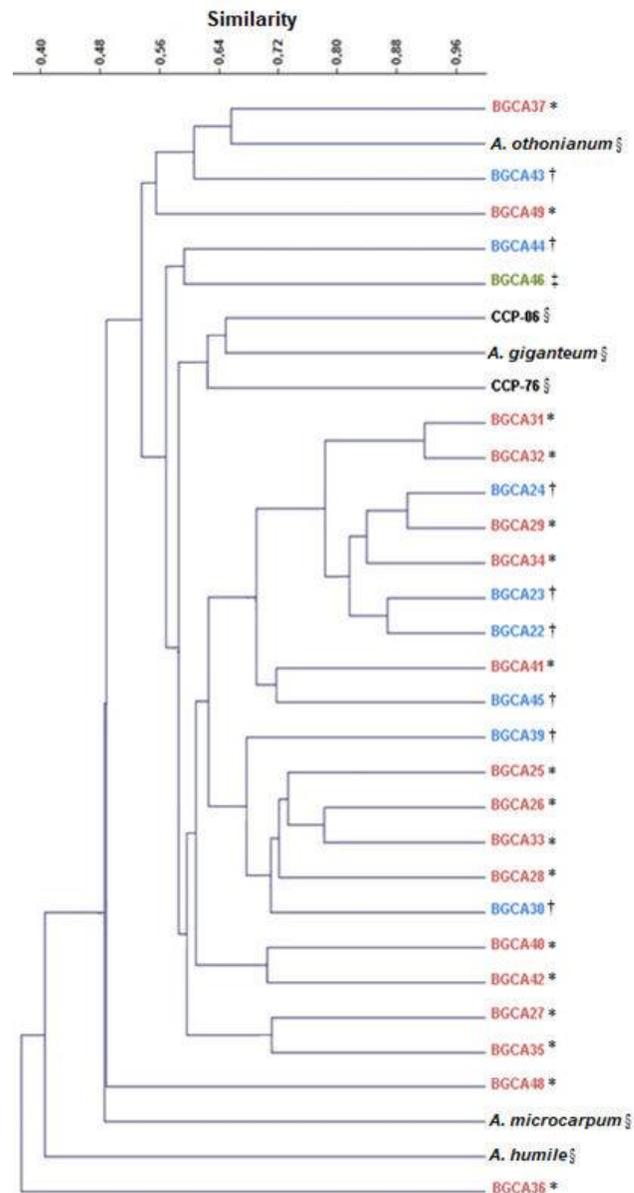
Figure 1. Dendrogram obtained from unweighted pair group method with arithmetic mean(UPGMA) cluster analysis and coefficients of similarity of Sørensen–Dice indices representing the genetic relationships among 32 accessions of cajuí and *Anacardium* spp. based on ISSR data. The origins of the accessions are denoted by the following symbols: * Parnaíba, PI; † Ilha Grande, PI; ‡ Luzilândia, PI; and § Pacajus, CE
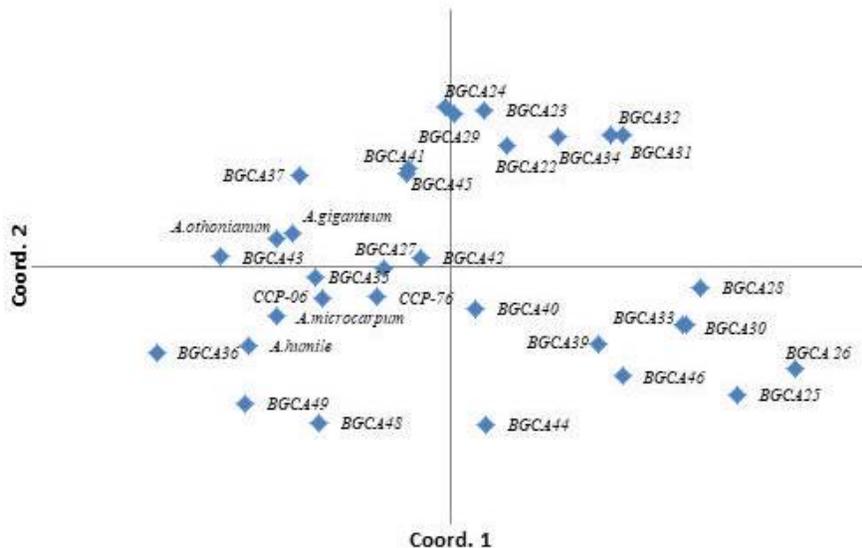
Figure 2. Scatter plot of the principal coordinate analysis (PCoA) of 32 accessions of cajuí and *Anacardium* spp. based on ISSR data

The high degree of polymorphism was confirmed by the Nei index (which can assume values between 0 and 0.5) and the Shannon index, thereby demonstrating high genetic diversity among the 25 cajuí accessions (Table 3). Although genetic diversity appeared to be higher in the population from Parnaíba than within the Ilha Grande population, the mean heterozygosity values (*H*) of the two presupposed populations were similar, suggesting that the difference in sample sizes (17 from Parnaíba and 8 from Ilha Grande) masked the true diversity.

Table 3. Genetic diversity among the accessions of cajuí and *Anacardium* spp. studied based on 104 ISSR loci

Populations	Polymorphic loci		Genetic diversity indices		Average heterozygosity
	n	%	Shannon (*I*)	Nei (*h*)	*H*
Ilha Grande accessions	75	72.12	0.371	0.289	0.314
Parnaíba accessions	94	90.38	0.416	0.365	0.309
All accessions	97	93.27	0.418		

AMOVA showed that intrapopulation variability was responsible for 100% of the genetic diversity, and this finding was confirmed by the fixation index ($\Phi_{ST} = 0$) and by the UPGMA and PCoA plots, which indicated that the cajuí accessions could not be grouped according to geographical origin. The presence of little or no genetic differentiation between the two presupposed populations was verified by the low value of G_{ST} (0.0366), implying a high crossover rate as confirmed by the high value of N_m (13.145). However, the delta K method of estimating population structure suggested K = 2 as the most likely number of groups (Figure 3). Furthermore, population structure inferred by the STRUCTURE software at K = 2 showed that each of the groups incorporated accessions from the two locations (Figure 4). Detailed analysis of the plot revealed that all of the accessions were likely to belong to either of the two groups, indicating a tendency towards

the mixture of genomes. These results suggest that, although the accessions could be classified in two distinct genetic groups, a high frequency of allele exchange had taken place between them such that the two presupposed populations could not be structured on this basis.
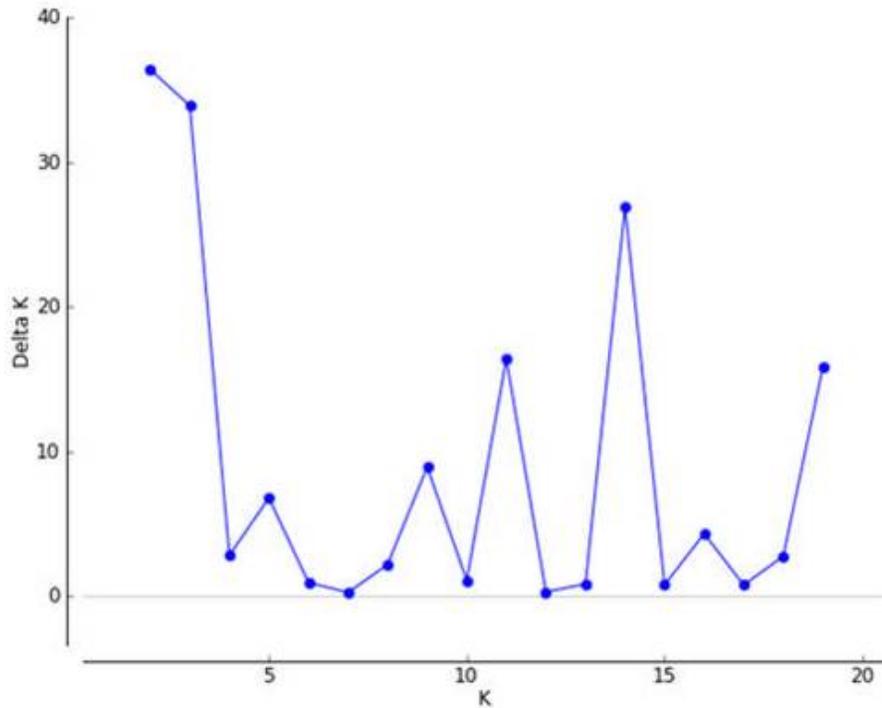

Figure 3. Graphical representation of delta K values for 25 accessions of cajuí calculated according to the method of Evanno et al. (2005). The most probable value of K (= 2) corresponds to the highest value of delta K shown on the plot
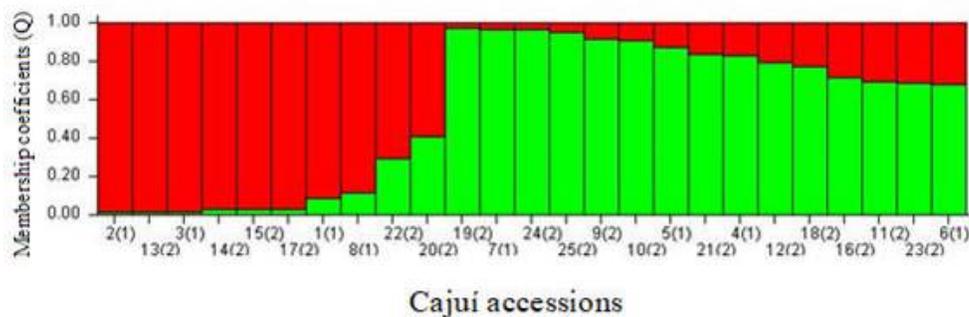

Figure 4. Population structure of 25 accessions of cajuí originating from coastal areas of Piauí (PI) inferred by STRUCTURE software at K = 2. Each color represents a different group, so that group 1 is indicated by red, whereas group 2 is indicated by green. Bars with different colors indicate the percentage of shared genome between the two groups. The accession numbers correspond with those shown in Table 1, while the accompanying numbers in parenthesis indicate the geographic populations: Ilha Grande (1) and Parnaíba (2), PI, Brazil.

Together with population structure, we investigated the evidence of possible migrants or descendents of recent migrants between the groups by using the USEPOPINFO, GENSBACK and MIGRPRIOR functions of the STRUCTURE software. Figure 5 shows overlapping of individuals from different populations in both groups such that cluster 1 contained 51% of Ilha Grande accessions and 24.5% of Parnaíba accessions while cluster 2 encompassed the remaining accessions.
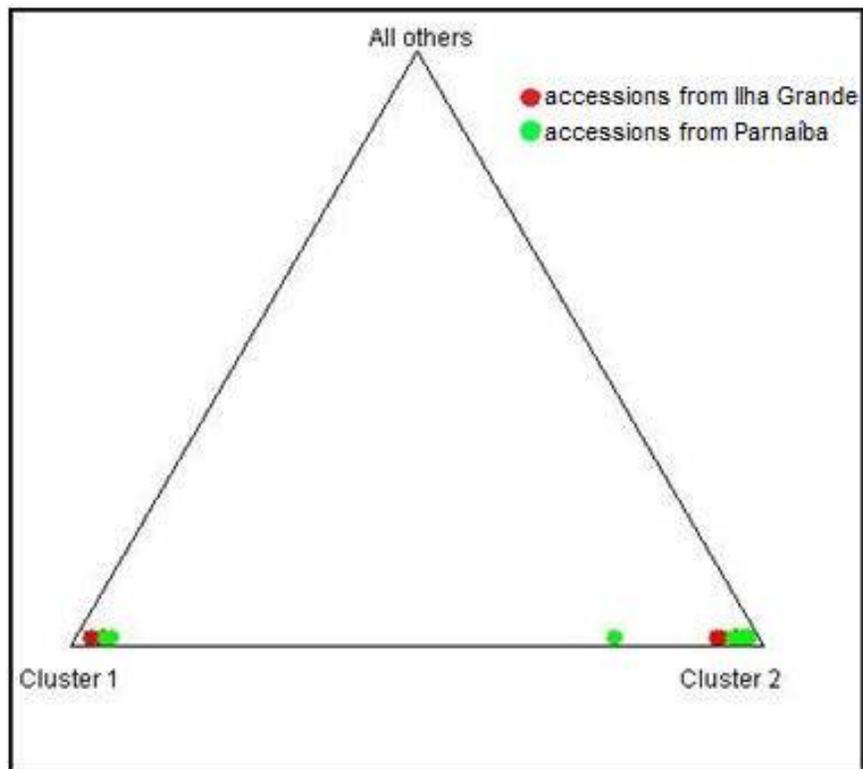

Figure 5. Schematic representation of group membership inferred by STRUCTURE software at K = 2 using USEPOPINFO function with MIGRPRIOR = 0.05 and GENSBACK = 3

DISCUSSION

The high degree of polymorphism detected among the studied accessions demonstrates the efficiency of ISSR markers selected in the determination of genetic diversity in the active germplasm bank of cajuí maintained at Embrapa Meio-Norte. Our findings are consistent with those of previous studies on the use of ISSR markers in assessing the genetic diversity of *Anacardium* (Archak et al. 2003; Pessoni2007; Thimmappaiah et al. 2009). Indeed, ISSR markers are very valuable in plant breeding research since they allow the characterization and discrimination of genetic diversity within a genus, with low cost and efficient differentiation between individuals, even among

samples from different species (Gupta et al. 1994; Bornet and Branchard2001; Reddy et al. 2002).

Genetic diversity among accessions of cajuí, as determined by the Nei and Shannon indices (Table 3), was compatible with the high percentage of polymorphism detected by the ISSR markers. These results indicate the existence of substantial genetic variability among the studied germplasm. Moreover, mean heterozygosity values, estimated using the Bayesian approach in order to eliminate bias caused by population size, were similar for both presupposed populations and confirmed the high genetic diversity in the overall population. High genetic diversity has also been found for other species *of Anacardium*, such as the natural populations of *A. humile* from the Cerrado (Cota et al. 2017).

According to the UPGMA dendrogram and the PCoA plot there was no correspondence between genetic variability and geographic origin of the cajuí accessions since genotypes from the same location emerged in distinct clusters and those from different locations displayed genetic similarity. Patterns similar to those presented in our study were suggested for eight natural *A. humile* populations, for whichno significant correlation between genetic and geographic distances was observed by the Mantel test (Cota et al. 2017). The absence of spatial relationship patterns can be attributed to several processes, especially those resulting from foraging by dispersal agents and pollinators, which affects the genetic diversity.

Findings relating to the organization of genetic diversity obtained by AMOVA were consistent with the low estimate of G_{ST} (0.0366) and suggested that the high genetic variability observed was not structured among the populations but rather the result of individual differences between the accessions. The absence of significant differentiation was likely the result of various factors that favored intense gene flow among the populations (N_m = 13.145) such as geographic proximity (only 8 km) and a reproductive system that allows the extensive distribution of pollen and seeds.

Since gene flow in plants occurs preferentially through the movement of pollen and seeds (Martins 1987), factors that facilitate their dispersion over long distances may have significant implications on the population structure of a species (Sork and Smouse2006). In the case of *Anacardium*, the most important seed dispersers are frugivorous bats (Mitchell and Mori 1987) that are attracted to the fleshy and juicy pseudofruits and can transport seeds in their guts over long distances. Four of the 13 species of bats that inhabit the coast of Piauí are frugivorous, while two of these, namely *Artibeus lituratus* and *Carollia perspicillata*, exhibit foraging behavior (Rocha and Portella 2012) and are able to cover distances of up to 8 and 13 km, respectively, in a single night while visiting several feeding sites (Galindo-González 1998). Other long-distance seed dispersers that can intensify gene flow among *Anacardium* are humans, through intentional or accidental actions, and water, in which the fruit and associated peduncle has the ability to float and be carried along a watercourse. Moreover, *Anacardium* species present significant changes in the different periods of the year, as a strategy for improved seed dispersal (Sousa and Cunha 2018). Alongside the efficiency of the dispersing agents, efficient long-distance gene flow requires superior seed viability and, in this context, cajuí seeds reportedly exhibit high germination rates (Corrêa et al. 2002).

The main pollen dispersers in *Anacardium* are honeybees, especially *Apis mellifera* (Holanda-Neto2008), although some indigenous bees, including *Centris tarsata*, are also important dispersal agents (Freitas and Paxton 1998). Honeybees exhibit foraging behavior

that favors pollination in that they visit flowers at a time when the pollen is viable and the stigma is receptive, they touch the anther and the stigma with the same portion of the body (the mesothorax), they exhibit pollinator constancy by not alternating between *Anacardium* flowers and those of other plants (Freitas and Paxton 1998; Freitas et al. 2002), and they cover long distances searching for desirable floral resources (Hagler et al. 2011). Moreover, the flowering period of cajuí coincides with the dry season (Rufino 2004) in which food sources are scarce, so that visitation of *Anacardium* flowers is most important for the survival of the bees. In this sense, pollen dispersion by bees would certainly contribute to efficient long-distance gene flow in *Anacardium.*

Detailed taxonomic studies are required to elucidate which species of *Anacardium* constitute cajuí in Piauí, since the existing literature refers only to cashew species (Rufino et al. 2008; Gomes et al. 2009, 2013) or to *A. microcarpum* Duck (Barros et al. 1999; Vieira et al. 2004; Andrade et al. 2012), although it has been reported that some cajuí specimens grown in Piauí are genetically similar to *A. othonianum* Rizzini and *A. occidentale* (Pessoni 2007). Insofar as taxonomic studies relating to the species of cajuí are very limited, it must be assumed that the accessions studied herein may comprise different species. Thus, an alternative hypothesis that could explain the genetic diversity between the studied accessions is the possibility of intense gene flow between *A. occidentale* and cajuí species facilitated by the overlap of geographic distribution. Furthermore, interspecific hybridization appears to be a common phenomenon within the genus under natural conditions, as exemplified by the three indigenous sympatric species of the Brazilian central plateau, namely *A. occidentale*, *A. humile* and *A. nanum* (Mitchell and Mori 1987), and by artificial crosses between, for example, *A. occidentale* and *A. microcarpum* (Crisóstomo et al. 2002). Along with weak reproductive barriers within *Anacardium* and the possibility of successful interspecific hybridization, the absence of population structure is supported by leaf morphology. According to Vieira et al. (2014), the leaves of specimens of cajuí growing along the coast of Piauí are so similar to those of *A. occidentale* occurring in the same area that it is impossible to differentiate between the groups on the basis of this characteristic.

The number of chromosomes in *Anacardium* is highly variable as exemplified by *A. occidentale,* (2n = 24, 40 or 42; Goldblatt 1981, 1984; Thankamma-Pillai and Nambiar 1985) and *A. othonianum* (2n = 16; Leão et al. 2001).Moreover, Santos et al. (2015) reported variations in the chromosome number of specimens of *A. occidentale* (2n = 20 or 28) occurring in sympatry with cajuí (2n = 26). This variability may be due to several factors including polyploidy in *A. occidentale* (Leão et al. 2001) and natural hybridization with the generation of intraspecific hybrids.

The results shown herein indicate that high gene flow between the two presupposed populations contributed to the homogenization of the accessions by promoting intense exchange of alleles between them. The results of Bayesian analysis performed using the STRUCTURE software were in agreement with those derived from UPGMA, PCoA, AMOVA and G_{ST}, demonstrating that the genetic variability was not spatially organized and that the studied accessions actually constituted a single population. Although delta K analysis suggested that the accessions could be structured in two genetically distinct groups (K = 2), a closer observation of the pattern represented in Figure 4 reveals that each accession could be placed in either of the two clusters. This means that there had been intense exchange of alleles between the accessions and that the two apparently distinct

groups are significantly linked by a number of attributes resulting in a mixture of genomes. Owing to their proximity, the Parnaíba and Ilha Grande populations contain recent *Anacardium* migrants or migrant descendents, a situation that was clearly exposed using the USEPOPINFO, GENSBACK and MIGRPRIOR functions of the STRUCTURE software.

CONCLUSIONS

The results obtained from ISSR analysis with nine markers demonstrated that the two presupposed populations of cajuí were not genetically differentiated but constituted a single population containing highly diversified individuals including migrants, migrant descendants and possible hybrids. Nonetheless, the genetic variability could be organized into two distinct, but linked, groups that had undergone extensive exchange of genetic material as demonstrated by the high gene flow index. The high genetic variability among the accessions was attributed to individual differences rather than to differential spatial distribution. This report enhances our knowledge of the genus *Anacardium* and should facilitate future research on the improvement of cajuí culture and fruit quality. Moreover, our study highlights the importance of further taxonomic studies on *Anacardium*.

ACKNOWLEDGMENTS

REFERENCES

Almeida AS (2009). Qualidade, compostos bioativos e atividade antioxidante total de pendúculos de cajuizeiros e frutos de umbuzeiros nativos do semiárido do Piauí. PhD Thesis. Universidade Federal Rural do Semiárido, Mossoró. Available in [http://www.alice.cnptia.embrapa.br/alice/handle/doc/662479].

Andrade IM, Silva MFS, Mayo SJ, et al. (2012). Diversidade de fanerógamas do delta do Parnaíba – litoral piauiense. In: (Guzzi A ed.). Biodiversidade do delta do Parnaíba: litoral piauiense. *EDUFPI*, Parnaíba.

Archak S, Gaikwad AB, Gautam D, et al. (2003). DNA fingerprinting of Indian cashew (*Anacardium occidentale* L.) varieties using RAPD and ISSR techniques. *Euphytica* 230:397-404.

Barros LM, Paiva JR, Cavalcanti JJV (1999). Recursos genéticos de cajueiro: situação atual e estratégias para o futuro. In: (Queiroz MA, Goedert CO, Ramos SRR eds.). Recursos genéticos e melhoramento de plantas para o Nordeste brasileiro. Embrapa Semiárido, Petrolina.

Bornet B, Branchard M. (2001). Nonanchored inter simple sequence repeat (ISSR) markers: reproducible and specific tools for genome fingerprinting. *Plant Mol Biol Rep*. 19:209-215.

Carbajal ACR, Silva-Júnior N (2003). Castanha de caju: recomendações práticas para a melhoria da qualidade. 1st edn. Embrapa Agoindústria Tropical, Fortaleza.

Corrêa MPF, Rufino MSM, Vasconcelos LFL, et al (2002). Germinação e vigor de sementes de genótipos de cajuí (*Anacardium* spp.). In: Anais do Congresso Brasileiro de Fruticultura – Os novos desafios da fruticultura brasileira. Sociedade Brasileira de Fruticultura, Belém, pp 1-5.

Cota LG, Moreira PA, Brandão MM, et al. (2017). Structure and genetic diversity of *Anacardium humile* (Anacardiaceae): a tropical shrub. *Genet Mol Res*. 16(3): gmr16039778

Crespo MFV, Souza LI (2014). Cajuí: boas práticas e manejo sustentável. 1st edn. Embrapa Meio-Norte, Teresina.

Crisóstomo JR, Cavalcante JJV, Barros LM, et al. (2002). Melhoramento do cajueiro a não-precoce: avaliação da qualidade do pedúnculo e a heterose dos seus híbridos. *Rev Bras Frutic*. 24:477-480.

Earl DA, Von Holdt BM. (2012). Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Conservation Genet Resour*. 4:359-361.

Evanno G, Regnaut S, Goudet J. (2005). Detecting the number of clusters of individuals using the software Structure: a simulation study. *Mol Ecol*. 14:2611–2620.

Excoffier L, Laval G, Schneider SL .(2007). Arlequin version 3.11: a software for population genetic data analysis. University of Geneva, Geneva.

Falush D, Stephens M, Pritchard JK. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes*. 7:574–578.

Freitas BM, Paxton RJ. (1998). A comparison of two pollinators: the introduced honey bee *Apis mellifera* and an indigenous bee *Centristarsata* on cashew *Anacardium occidentale* in its native range of NE Brazil. *J Appl Ecol*. 35:109-121.

Freitas BM, Paxton RJ, Holanda-Neto JP (2002). Identifying pollinators among an array of flower visitors, and the case of inadequate cashew pollination in NE Brazil. In: (Kevan PG, Imperatriz-Fonseca VL, eds.). Pollinating bees - The conservation link between agriculture and nature. Ministério do Meio Ambiente, Brasília.

Galindo-González J. (1998). Dispersion de semillas por murcielagos: su importância en la conservacion y regeneracion del bosque tropical. *Acta Zool Mex* .73:57-74.

Goldblatt P (ed) (1981) Index to plant chromosome numbers 1975 -1978. Missouri Botanical Garden, St. Louis.

Goldblatt P (ed) (1984) Index to plant chromosome numbers 1979 -1981. Missouri Botanical Garden, St. Louis.

Gomes SO, Souza VAB, Costa MPSD, et al. (2009) .Características físicas e químicas de frutos de cajuí (*Anacardium* ssp.). In: Anais do Congresso Brasileiro de Melhoramento de Plantas – O melhoramento e os novos cenários da agricultura. INCAPER, Guarapari. https://ainfo.cnptia.embrapa.br/digital/bitstream/item/117074/1/CBMPVALDOMIRO3.pdf

Gomes SO, Souza VABS, Costa MPSD, et al. (2013). Avaliação da qualidade física e química de cajuí (*Anacardium* spp.) na região Meio-Norte. *Geintec*. 3:139-145.

Gupta M, Chyi Y-S, Romero-Severson J, et al. (1994). Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theor Appl Genet*. 89:998-1006.

Hagler JR, Mueller S, Teuber LR, et al. (2011). Foraging range of honey bees, *Apis mellifera*, in alfalfa seed production fields. *J Insect Sci*.11: 144.

Hammer Ø, Harper DAT, Ryan PD. (2001).Past: Paleontological statistics software package for education and data analysis. *Palaeontol Electron*. 4: 4.

Holanda-Neto JP (2008).The pollination of cashew (*Anacardium occidentale*) in northeast Brazil. PhD Thesis. Queen's University Belfast, Belfast.

Leão ACM, Rodrigues WA, Chaves LJ, et al. (2001). Análise citogenética da espécie do caju-do-campo *Anacardium othonianum*. In: Anais do Congresso Nacional de Genética. Sociedade Brasileira de Genética, Águas de Lindóia.

Lynch M, Milligan BG (1994). Analysis of population genetic structure with RAPD markers. *Mol Ecol*. 3:91-99.

Martins PS (1987) Estrutura populacional, fluxogênico e conservação "in situ". IPEF 35:71-78. http://www.ipef.br/publicacoes/scientia/nr35/cap05.pdf

Miller MP (1997) Tools for population genetic analysis (TFPGA) 1.3: A Windows™ program for the analysis of allozyme and molecular population genetic data. Northern Arizona University, Flagstaff

Mitchell JD, Mori SA (1987). The cashew and its relatives (*Anacardium*: Anacardiaceae), Memoirs of the New York Botanical Garden. 42 stedn. NYBG Press, New York.

Pessoni LA (2007). Estratégias de análise da diversidade em germoplama de cajueiro (*Anacardium* spp. L.). PhD Thesis. Universidade Federal de Viçosa, Viçosa. Available in [http://www.locus.ufv.br/bitstream/handle/123456789/1362/texto%20completo.pdf?sequence=1&isAllowed=y]

Peakll R, Smouse P (2006) GenAlEx version 6. The Australian National University, Canberra.

Pritchard JK, Stephens M, Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics*. 155:945-959.

Reddy MP, Sarla N, Siddiq EA. (2002). Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica* 128: 9–17.

Rocha CR, Portella AS (2012). Morcegos do delta do Parnaíba, litoral piauiense. In: (Guzzi A ed.). Biodiversidade do delta do Parnaíba: litoral piauiense. EDUFPI, Parnaíba.

Rufino MSM (2004). Qualidade e potencial de utilização de cajuís (*Anacardium* spp.) oriundos da vegetação litorânea do Piauí. MSc Dissertation. Universidade Federal do Piauí, Teresina. Available in [http://livros01. livrosgratis.com.br/cp109622.pdf].

Rufino MSM, Corrêa MPF, Alves RE, et al. (2008). Utilização atual do cajuí nativo da vegetação litorânea do Piauí, Brasil. *Proc Interamer Soc Trop Hort*.52:147-149.

Santos VMA, Carneiro SMG, Nascimento JDO, et al. (2015). Estudo cariotípico de dois morfotipos de *Anacardium* L. encontrados no litoral piauiense. *Rev Bras Biodiver Biotecnol*.1:270-271.

Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana

Sork VL, Smouse PE. (2006) .Genetic analysis of landscape connectivity in tree populations. *Landscape Ecol*. 21: 821-836. Sousa DG, Cunha HF (2018). Population structure, spatial distribution and phenology of *Anacardium humile* A. St.-Hil. (Anacardiaceae) in cerrados trictosensu. *Hoehnea*. 45: 450-467.

Thankamma-Pillai PK, Nambiar MC. (1985). Study of microsporogenesis in cashew. *Acta Hort*.108: 55

Thimmappaiah SWG, Shobha D, Melwyn GS. (2009). Assessment of genetic diversity in cashew germplasm using RAPD and ISSR Markers. *Sci Hort.* 120:411-417.

Vekemans X. (2002) AFLP-SURV version 1.0. Laboratoire de Génétique et EcologieVégétale, Université Libre de Bruxelles, Brussels

Vieira M, Mayo SJ, Andrade IM. (2014) Geometric morphometrics of leaves of *Anacardium microcarpum* Ducke and *A. occidentale* L. (Anacardiaceae) from the coastal region of Piauí, Brazil. *Braz J Bot.* 37:315-327.

Yeh FC, Yang R-C, Boyle TJB (1999) POPGENE version 1.31. Microsoft Windows-based freeware for population analysis. University of Alberta, Edmonton / Centre for International Forestry Research, Bogor

Zhivotovsky LA. (1999). Estimating population structure in diploids with multilocus dominant DNA markers. *Mol Ecol.* 8:907-913.