DIVERSITY AND GENETIC STRUTURE IN CAJÁ TREE (*Spondias mombin* L.) POPULATIONS IN NORTHEASTERN BRAZIL[1]

EDSON FERREIRA DA SILVA[2,5], LUIZA SUELY SEMEN MARTINS[3], VISÊLDO RIBEIRO DE OLIVEIRA[4]

ABSTRACT – *Spondias mombin* L. is a fruit tree from the American continent from the Anacardiaceae family. In Brazil it is common in different vegetation types but is more frequent in the Atlantic and Amazonian rainforests. It is economically important because of its fruits, which are widely consumed raw or processed as fruit jellies, juices and ice creams. The leaves have great importance in the pharmaceutical industry because of their antibacterial properties. In the state of Pernambuco, cajá tree is widely distributed in the Zona da Mata region and less frequently in the Agreste and Sertão areas. In this work diversity and genetic structure were studied in four populations of cajá tree from Pernambuco's Zona da Mata, Northeast Brazil, using isozymes polymorphism analyses from electrophoreses. The result showed 100% of polymorphism (P) for nine alleles ($\hat{A}$)and the average of alleles per locus s was 2.4. The expected heterozygosity ($\hat{H}_e$)ranged from 0.530 to 0.574 and the observed heterozygosity ($\hat{H}_o$), from 0.572 to 0.735. It was not observed inbreeding and the average F_{IT} was -0.175, whereas within population inbreeding (*f*) varied from -0.08 to -0.37. The genetic divergence among the populations (F_{ST}) ranged from 0.006 to 0.028 and the average was 0.026. The average of estimated gene flow (*Nm*) was high (5.27). The CG-IPA population, corresponding to the germplasm collection of IPA, showed more than 96% of genetic similarity with other populations; therefore, it is a good representative of the existent genetic diversity in the Zona da Mata region.
Index terms: Genetic diversity, germplasm conservation, isozymes.

DIVERSIDADE E ESTRUTURA GENÉTICA DE POPULAÇÕES EM CAJAZEIRA (*Spondias mombin* L.) NO NORDESTE DO BRASIL

RESUMO - A espécie *S. mombin* L. é uma árvore frutífera de origem americana pertencente à família Anacardiaceae . No Brasil, é encontrada nos mais variados tipos de vegetação, porém é mais freqüente na Mata Atlântica e na Floresta Amazônica. Destaca-se economicamente por seus frutos serem amplamente consumidos *in natura* ou utilizados na indústria, principalmente para fabricação de polpas, geléias, sucos e sorvetes. As folhas têm grande importância para a indústria farmacêutica, por possuírem propriedades antibacterianas. No Estado de Pernambuco, a cajazeira encontra-se amplamente distribuída na Zona da Mata, ocorrendo com menos freqüência no Agreste e Sertão. Neste trabalho, estudaram-se a diversidade e a estrutura genética de populações de cajazeiras na Zona da Mata do Estado de Pernambuco, Nordeste do Brasil, por meio da análise do polimorfismo isoenzimático obtido por eletroforese. Os resultados revelaram 100 % de polimorfismo para nove alelos, sendo que o número médio de alelos por loco ($\hat{A}$) foi de 2,4. A heterozigosidade esperada ($\hat{H}_e$) variou de 0,530 a 0,574, e a heterozigosidade observada ($\hat{H}_o$)variou de 0,572 a 0,735. Não se verificou endogamia nas populações, sendo que o valor médio de F_{IT} foi de -0,175, intrapopulacional (*f*) variou de -0,08 a -0,37. A diferenciação entre as populações (F_{ST}) variou de 0,006 a 0,028, sendo que o valor médio de 0,026 e, o fluxo gênico estimado entre as populações (*Nm*) foram altos (5,27). A população CG-IPA, que constitui a Coleção de Germoplasma do IPA, apresentou acima de 96% de identidade genética com as demais populações e, portanto, tem alta representatividade na diversidade existente na região.
Termos para indexação: Isoenzimas, conservação de germoplasma, diversidade genética.

INTRODUCTION

The Cajá (*Spondias mombin* L.) is a fruit tree of the Anacardiaceae family that could reach up to 25 m of height. The genus *Spondias* has 18 species, nine of which in Asia and Oceania, and nine others in the Neotropics. The origin center of *S. mombin* is the American continent and it is widely distributed in the tropics (Airy e Forman, 1967). The species occurs in many regions of Brazil in different types of vegetation but is more frequent in the Atlantic and Amazon rainforests (Silva e Silva, 1995). In Brazil it is popularly called cajá in the Northeast, cajá-mirim in the South and tapereba in the North (Corrêa, 1984).

Economically *S. mombin* is important because of its succulent fruits which have high concentration of vitamin C and is widely consumed raw or processed as fruit jellies, juices and ice creams. The leaves have great importance in the pharmaceutical industry because they have antibacterial proprieties (Ajao *et al.*, 1985; Silva e Silva, 1995). According to Daniel (1990), all parts of the tree, such as the leaves, flowers and fruits have been used in the pharmaceutical industry.

The cajá fruit is drupe with a yellow-orange color, with a succulent mesocarp (Cavalcante, 1976; Braga, 1976; Lorenzi, 1992), which presents a high variability for both size and shape. According to Bosco *et al.* (1997) the fruits are classified as big, when they are more than 15g, medium when are from 12 to 15g and small when they are less 12g. Relative to the shape, the fruits are classified as rounded when the difference between longitudinal and transversal diameters is less than 1 cm and elongated when the difference is more than 1 cm.

As to the phenology, in the northeast of Brazil the *S. mombin* flowering starts in October and ends in December, depending upon the genotypes and also could vary according to the environment. In northeast Brazil, on the average the time period from fecundation to fruit ripening lasts from 85 to 88 days (Bosco *et al.*, 2000).

The cajá flowers are hermaphrodite, diclamidious, pentamers and diplostemonous, grouped in terminal inflorescence. There are few studies about flower biology, however some reports shows that there are variations in the reproductive behavior in different areas where the specie occur. For example, Pennington and Sarukhan (1968) related dioecism in Mexico, Bawa and Opler (1975) described monoecism in Costa Rica and Croat (1978) showed that in Panama the majority of flowers were bisexual. Croat (1978) and Janzen (1985) related that in Costa Rica the pollination of *S. mombin* was predominately carried out by bees and other small insects.

The cajá tree exploration in Brazil is done basically by extractivism (Silva e Silva, 1995). In the state of Pernambuco, this species is distributed in anthropic environments such as road margins, in sugar cane plantations and backyards in the Zona da Mata region. In the Agreste and Sertão areas, the species occurs less frequently, mainly on road margins and backyards. In spite of being less studied, this species has a great potential for economic exploration, mainly in the northeastern region where the climatic conditions are adequate for commercial cultivation. The genetic resources of the cajá tree are very important and its characterization could give support for its exploration and conservation. According to Brown and Moran (1981) in order to evaluate the genetic resources of some species it is necessary to have information about the genetic structure of its populations including the genetic diversity within and among populations.

In order to study the natural populations of Brazilian tree fruits this methodology was used by Telles *et al.* (2003) with araticunzeiro (*Annona crassiflora* Mart.), by Sebbenn *et al.* (2003) studying genipapo (*Genipa americana* L.) and by Melo Junior *et al.* (2004) investigating the pequizeiro (*Caryocar brasiliense* Camb.).

Studies of isozyme polymorphism allow investigators to examine not only the genetic structure and variability of a population, but also to estimate genetic flow among populations and make inferences concerning reproductive processes (Hamrick 1983). Genetic flow (N_m) is classified as an indirect investigative method (Slatkin 1985), generating data that allow investigators to determine if genetic drift alone is sufficient to produce the genetic differentiation observed among populations (Slatkin e Barton 1989).

The objective of this work was to study the diversity and genetic structure of cajá tree (*S. mombin*) populations in the Zona da Mata of the state of Pernambuco and to compare them with a germplasm collection maintained at the experimental center of Empresa Pernambucana de Pesquisa Agropecuária (IPA) in the county of Itambe, in order to obtain information for preservation and use in future plant breeding programs.

MATERIAL AND METHODS

Plants from the germplasm collection of the Empresa Pernambucana de Pesquisa Agropecuaria (IPA) and from three regions of Zona da Mata from the state of Pernambuco constituted the four populations used in this study.

The germplasm collection of IPA (CG-IPA) is located in the experimental station of Itambé at the geographic coordinates 07° 24' 50'' S e 35° 06' 30'' O. The other three populations occur under natural conditions and were named respectively Muribeca, Itamaraca and Mata Sul. The Muribeca population is located at the district with the same name in the county of Jaboatão dos Guararapes, located in the central region of Zona da Mata; the Itamaraca population is located in the island of Itamaraca, situated in the northern region of the Zona da Mata, in Vila Velha community and near the Agricultural penal colony, and the Mata Sul population was constituted of individuals that occur in sugar cane plantations, road margins and backyards in the south of the Zona da Mata region . For each sampled individual from these populations the geographic coordinates (latitude and longitude) were registered using a Global Position System with precision of 10 m (Table 1).

In the CG-IPA population all 33 individuals that represent the collection were studied and in the Muribeca, Itamaracá and Mata Sul populations 20, 32 and 32 individuals, respectively were sampled. The sampling in Muribeca, Itamaracá and Mata Sul was by a randomized sample scheme of adult individuals in order to obtain representation of the studied area.

The samples consisted of young leaves which were placed in plastic bags and kept in ice coolers to transport to the Genetics Laboratory of the Biology Department of the Federal Rural University of Pernambuco, where they were then stored at -80°C. The enzymes were extracted from every sample in 1 ml of extraction buffer n° 1, according to the methodology developed by Alfenas *et al.* (1998).

The isozymes were separated by horizontal electrophoresis on 13% starch gel (Penetrose 30). The gel/electrode buffer systems used were TC (Tris Citrate, pH 7.5), TCB (Tris Citrate Borate, pH 7.5), and LB (Lithium Borate, pH 8.5). After the run, the electrophoresis plates were removed and sliced, the slices were subjected to staining for specific enzymes examined according to the methodology established by Alfenas *et al.* (1998).

A total of 12 enzyme systems were initially tested and, of these, nine were selected for detailed examination because they presented loci and alleles with resolutions that facilitated interpretation: peroxidase (PO), acid phosphatase (ACP), alpha-esterase (α-EST), glutamate oxaloacetate transaminase (GOT), glucose dehydrogenase (GLUDH), alcohol dehydrogenase (ADH), alkaline phosphatase (ALP), malate dehydrogenase (MDH) and superoxide dismutase (SOD). The interpretations of each enzymatic system were performed according to techniques described in detail in the literature (Alfenas *et al.* 1998; Oliveira *et al.* 2006).

The genetic variability was obtained using the allele frequency estimative and diversity indexes (expected $\left(\hat{H}_e\right)$ and observed $\left(\hat{H}_o\right)$ heterozygosis, percentage polymorphic loci (P) and average number of allele per locus $(\hat{A})$). For these estimatives the BIOSYS 1 software program was used (Swofford and Selander 1989), which gives also the inbreeding index according to *F*-statistics (Wright 1965).

In order to measure the genetic flow (N_m), the model proposed by Crow and Aoki (1984) was used indirectly according to the following equation:

$$Nm = (\frac{1}{4\alpha})\left[\left(\frac{1}{\hat{F}_{ST}}\right) - 1\right] \quad \text{where:} \quad \alpha = \left[\frac{n}{(n-1)}\right]^2$$

and where, N_m = number of migrants per generation

n - the numbers of populations

F_{ST} = genetic divergence among populations, which was calculated by the combinations of pairs of populations (using the BIOSYS-2 software program).

RESULTS AND DISCUSSION

Twenty two alleles of nine loci from nine isozyme systems Po, Acp, Est, Got, Gludh, Adh, Alp, Mdh and Sod were analyzed. The allelic frequency were from a high of (0.789) for A allele of Got locus in Itamaraca population to very low (0.069) as B allele of Glu locus in Muribeca population (Table 2).

In none of the studied populations allele loss was observed. This is probably because there was homogenic spreading of the species during human occupation of the region and also the extractivist collectors action could had favored gene flow by seed dissemination preventing allele loss. It must be considered also that it being a tree species of long maturation period, the time from the initial species spreading was not sufficient for the loss of alleles by the action of dispersive mechanisms such as genetic drift and natural selection.

According to Souza *et al.* (2004), the biggest

probability of allele's loss occurs in rare ones or of low frequency, while common or regular ones have a higher chance of fixation. According to Crow and Kimura, (1970) and Young *et al*. (1996), the intensity of genetic drift is inversely proportional to the population size, where small populations are more vulnerable. Such factor leads to non random fluctuation in the allelic frequencies and, consequently, result in allelic fixation and/or loss when selection occurs, which was not observed in the studied populations of *S. mombin*.

From the chi-square (χ^2) test for the nine loci according to the Hardy-Weinberg equilibrium, only two loci (Est e Adh) were in equilibrium in the four studied populations (Table 3). In the same Table it is showed that for the CG-IPA, Muribeca and Itamaraca populations, the majority of loci are not in equilibrium. In the Mata Sul population there were five loci in equilibrium and four were not in equilibrium. These results indicate that there are no random effects in allelic frequency, as genetic drift or natural selection for all populations, and the Mata Sul population was less affected because there are only four loci that are not in equilibrium. The higher number of loci in equilibrium observed in the Mata Sul population also could be due to the sampling which was made wider area which normally reduces the probability of collecting related individuals.

It was observed 100% of polymorphic loci for all populations (Table 4), showing that there is a high polymorphic isozyme index in *S. mombin*. Melo Júnior *et al*. (2004) obtained 100% of isozyme polymorphism studying four populations of pequizeiro (*Caryocar brasiliense* Camb.), a fruit tree species of extractive exploitation as cajá tree. Kageyama *et al*. (2003) also obtained 100% of isozyme polymorphism studying four species of tropical tree. In Table 4, it is also showed that the average of allele per locus was 2.4, common value for tropical tree species. Melo Júnior *et al*. (2004) obtained values between 2.6 e 3.0 for natural populations of *C. brasiliense* and Kageyama *et al*. (2003) obtained values from 2.1 to 3.4 for five tropical tree species.

In this work, the observed average heterozygosity $\left(\hat{H}_o\right)$ ranged from 0.572 to 0.735 and the expected heterozygosity $\left(\hat{H}_e\right)$, which corresponds to genetic diversity, varied from 0.530 to 0.574 (Table 4). These values are higher than the ones obtained by Hamrick e Godt (1990), in working with tropical trees and are similar to the value obtained by Kageyama *et al*. (2003) for six tropical

tree species.

The highest heterozygosity $\left(\hat{H}_o\right)$ was observed in the Itamaracá population (0.735) and the lower value was observed in CG-IPA (0.572). The lower value observed in CG-IPA in relation to the others (Muribeca, Itamaracá and Mata Sul), probably is a consequence of used plant sampling strategy for assemblage of this collection. According to Ledman (2006)[*], during the plant collection, the representative of genetic diversity was considered, but preference was given for individuals with good commercial fruit characteristics, which could reduce genetic diversity.

In studying natural populations of a tropical tree with extractive exploitation (*C. brasiliense*), Melo Júnior *et al*. (2004) obtained values of genetic diversity (0.450 a 0.530) similar to the present study, while lower average values were obtained in *A. crassiflora* (0.285) by Telles *et al.* (2003) and in *G. americana* (0.195 for adults and 0.105 for progeny) studied by Sebbenn *et al.* (2003). The studied populations of *S. mombin* are not in protected areas and they are maintained by the local communities mainly due to fruit collection purposes.

Average indices for allele fixation (*f*) were negative for the majority of alleles in all populations. The Got locus was fixed for all populations and the Adh and Mdh loci in the CG-IPA population had positive values 0.247 and 0.457, respectively (Table 5). However, average fixation indices were negative for all populations, from -0.37 in Itamaracá population to -0.08 in CG-IPA population (Table 4). Negative values for fixation indicate inbreeding absence in the population.

The CG-IPA population showed three inbreeding loci, despite the negative average value. That result probably was influenced for non random methodology used for germplasm collection, as previously discussed.

The estimated average of fixation index within (F_{IS}) and among populations (F_{IT}), according to Nei (1978), revealed that there is no allele fixation in the populations because the averages of F_{IS} and F_{IT} were negative -0.191 and -0.175, respectively, and only Got locus showed fixation (Table 6). The absence of inbreeding and the high observed diversity indices agree with the statement that *S. mombin* in the Zona da Mata of Pernambuco is predominantly alogamous.

The average of genetic divergence between populations (F_{ST}) was 0.026, higher than average F_{IS} (-0,243). However, genetic differentiation among populations was 2.6%. Studying four populations of *C. brasiliense*, a tree of extractive exploitation as

cajá tree, Melo Junior (2004) obtained also low genetic differentiation (F_{ST}=0.020).

The average of F_{IS} was -0,243 and the negative value confirms that the populations are not in inbreeding process. The highest part of genetic diversity was observed within the populations (1-F_{ST} = 0.97), while only 0.03 were among populations. In this context, we can consider that the genetic diversity in cajá tree in the Zona da Mata has a satisfactory distribution, because the differentiation among population was low (2.6). This result reveals that, as in other populations, the preserved CG-IPA holds a high representation of the existent genetic variability in the region. Thus, it is not necessary to improve collection efforts in order to preserve the local gene pool.

Regarding genetic distances among populations, it could be observed on Table 7 that they are low and that the CG-IPA and Muribeca populations are the most genetically divergent (4%), whereas dissimilarity among Muribeca and Itamaracá populations were not observed. These results indicate that the populations had common ancestry during the species disseminations in the Zona da Mata in Pernambuco. These results agree with estimated gene flow for the four populations (5.27).

According to Slatkin and Barton (1989), Pinto and Carvalho (2004) and Souza et al. (2004) when the estimated gene flow is higher than 1.0, it is sufficient for avoiding differentiation among populations, and thus populations are not genetically isolated.

The estimated gene flow for populations by combination varied from 2.17 to 10.35 (Table 7). However, we must be careful when analyzing estimated Nm with CG-IPA, because the values represent gene flow before the establishment of the collection and we do not know exactly where the seeds were collected, maybe they came from the other three populations. Considering the other three natural populations by combination, Nm indices were higher than 1.0, showing that alleles are exchanged with high frequency among populations avoiding genetic differentiation. Population combination that showed higher gene flow was Muribeca and Itamaracá (10.35), indicating that both probably had a common origin or that fruits exchange is more frequent because they are closer. Even though *S. mombin* could be considered an autogamous species, high *Nm* index could be mainly due to seed dispersal as commercial fruits, because populations are far from each other.

The gene flow is estimated using F_{ST} and depends on the period of population separation, on reproductive system and on the period needed for species reproduction. According to Oliveira *et al.* (2002), the calculated value for migrant number in each generation represents the gene flow in the last generation, giving the actual pattern of population genetic structure.

According to Seoane *et al.* (2000), gene flow is an evolutionary factor that favors homogenization among populations, reducing the selection and genetic drift, consequently avoiding the genetic differentiation among populations. About the mechanisms involved in the gene flow of the studied cajá tree populations, fruit exchange was the most important because of the great distances between populations.

TABLE 1 – Location (latitude and longitude) of the sampled individuals in Muribeca, Itamaracá and Mata Sul populations registered using a Global Position System

Muribeca population			Itamaracá population			Mata Sul population		
Individual's number	Latitude (S)	Longitude (W)	Individual's number	Latitude (S)	Longitude (W)	Individual's number	Latitude (S)	Longitude (W)
1	08°10'03.8"	34°58'04.7"	1	07° 47'10.8''	34° 52'03.6''	1	08°40'12.1"	35°32'49.3"
2	08°10'04.1"	34°58'04.2"	2	07° 47'11.6''	34° 52'04.6''	2	08°40'28.9"	35°32'43.7"
3	08°10'04.2"	34°58'03.6"	3	07° 47'10.9''	34° 52'05.2''	3	08°40'44.1"	35°32'19.5"
4	08°10'05.9"	34°58'02.4"	4	07° 47'07.8''	34° 52'05.3''	4	08°40'42.9"	35°32'19.4"
5	08°10'07.0"	34°58'01.3"	5	07° 47'08.2''	34° 52'06.3''	5	08°41'03.0"	35°31'36.8"
6	08°10'07.0"	34°58'00.2"	6	07° 47'07.2''	34° 52'05.7''	6	08°41'50.1"	35°31'41.5"
7	08°10'04.6"	34°57'58.4"	7	07° 47'06.9''	34° 52'04.9''	7	08°41'30.3"	35°32'54.6"
8	08°10'04.0"	34°57'58.5"	8	07° 47'05.5''	34° 52'05.4''	8	08°40'58.7"	35°34'42.3"
9	08°10'01.7"	34°57'57.7"	08'	07° 47'02.8''	34° 52'07.7''	9	08°41'00.1"	35°34'39.3"
10	08°10'02.0"	34°57'59.2"	9	07° 47'10.8''	34° 52'03.6''	10	08°40'57.8"	35°34'42.5"
11	08°10'06.5"	34°57'59.2"	10	07° 47'06.3''	34° 52'00.6''	11	08°40'57.9"	35°34'39.4"
12	08°10'08.8"	34°57'59.3"	11	07° 47'09.4''	34° 51'59.5''	12	08°35'10.8"	35°41'04.2"
13	08°10'08.9"	34°57'59.1"	12	07° 47'09.9''	34° 51'59.2''	13	08°35'10.3"	35°41'03.4"
14	08°10'09.2"	34°57'59.1"	13	07° 47'09.8''	34° 51'59.2''	14	08°35'11.0"	35°40'30.3"
15	08°10'10.3"	34°57'58.2"	14	07° 47'11.4''	34° 51'59.4''	15	08°35'05.8"	35°40'07.5"
16	08°10'10.7"	34°57'59.7"	15	07° 47'11.3''	34° 52'02.5''	16	08°35'29.4"	35°39'43.8"
17	08°10'10.4"	34°58'01.8"	16	07° 47'11.1''	34° 52'02.6''	17	08°36'00.0"	35°38'15.4"
18	08°10'10.0"	34°58'02.0"	17	07° 46'04.0''	34° 52'09.6''	18	08°36'13.1"	35°38'05.0"
19	08°10'13.7"	34°58'03.5"	18	07° 46'04.0''	34° 52'09.6''	19	08°38'10.6"	35°36'50.3"
20	08°10'13.9"	34°58'04.6"	19	07° 46'01.2''	34° 52'15.5''	20	08°38'32.4"	35°36'41.4"
21	08°10'16.1"	34°58'04.7"	20	07° 46'01.4''	34° 52'15.7''	21	08°41'24.9"	35°36'14.5"
22	08°10'08.8"	34°57'55.1"				22	08°31'07.2"	35°05'33.8"
23	08°10'10.2"	34°57'56.2"				23	08°31'06.7"	35°05'33.6"
24	08°10'11.9"	34°57'56.7"				24	08°31'00.4"	35°05'35.7"
25	08°10'11.8"	34°57'56.7"				25	08°30'34.6"	35°05'43.7"
26	08°10'11.5"	34°57'54.7"				26	08°30'33.8"	35°05'46.0"
27	08°10'13.5"	34°57'53.7"				27	08°30'26.5"	35°05'48.4"
28	08°10'15.0"	34°57'55.1"				28	08°29'42.0"	35°05'32.5"
29	08°10'14.2"	34°57'56.6"				29	08°28'10.7"	35°05'40.0"
30	08°10'14.7"	34°57'57.3"				30	08°27'33.2"	35°05'51.7"
31	08°10'14.7"	34°57'52.8"				31	08°26'48.6"	35°05'34.6"
32	08°10'09.4"	34°57'52.4"				32	08°26'03.1"	35°04'49.6"

TABLE 2 - Estimated allele frequencies at nine polymorphic loci in the four populations of *Spondias mombin* in the NE Brazil.

Locus/Allele	Populations			
	CG-IPA	Muribeca	Itamaracá	Mata Sul
Po				
(N)	32	29	18	32
A	0.500	0.345	0.444	0.391
B	0.109	0.293	0.083	0.281
C	0.391	0.362	0.472	0.328
Acp				
(N)	32	28	20	32
A	0.469	0.429	0.350	0.375
B	0.109	0.071	0.150	0.266
C	0.422	0.500	0.500	0.359
Est				
(N)	32	30	20	31
A	0.469	0.483	0.475	0.387
B	0.531	0.517	0.525	0.613
Got				
(N)	32	29	19	28
A	0.406	0.724	0.789	0.464
B	0.594	0.276	0.211	0.536
Glu				
(N)	32	29	19	32
A	0.344	0.466	0.447	0.328
B	0.125	0.069	0.158	0.203
C	0.531	0.466	0.395	0.469
Adh				
(N)	29	28	19	31
A	0.241	0.500	0.500	0.484
B	0.759	0.500	0.500	0.516
Alp				
(N)	31	28	19	30
A	0.484	0.446	0.447	0.383
B	0.145	0.107	0.105	0.250
C	0.371	0.446	0.447	0.367
Mdh				
(N)	30	29	19	30
A	0.567	0.483	0.474	0.483
B	0.433	0.517	0.526	0.517
Sod				
(N)	29	29	20	28
A	0.431	0.448	0.525	0.482
B	0.569	0.552	0.475	0.518

(N) = Number of individuals.

TABLE 3 - Chi-square (c^2) and degrees of freedom (DF) for Hardy-Weinberg equilibrium in nine studied loci in four populations of *Spondias mombin* in the NE Brazil.

Locus	Populations							
	CG-IPA		Muribeca		Itamaracá		Mata Sul	
	DF	χ^2	DF	χ^2	DF	χ^2	DF	χ^2
Po	3	16.021[*]	3	9.713[*]	3	12.090[*]	3	13.273[*]
Acp	3	8.699[*]	3	23.233[*]	3	14.940[*]	3	13.144[*]
Est	1	0.000[ns]	1	0.000[ns]	1	1.052[ns]	1	0.049[ns]
Got	1	28.940[*]	1	25.145[*]	1	14.341[*]	1	25.070[*]
Glu	3	16.682[*]	3	21.974[*]	3	10.156[*]	3	4.162[ns]
Adh	1	0.851[ns]	1	0.232[ns]	1	1.605[ns]	1	0.000[ns]
Alp	3	13.077[*]	3	18.829[*]	3	11.987[*]	3	10.424[*]
Mdh	1	4.949[*]	1	1.067[ns]	1	0.681[*]	1	0.223[ns]
Sod	1	0.005[ns]	1	6.723[*]	1	1.052[ns]	1	0.019[ns]

* = P < 0,05, ns = not significant

TABLE 4 - Estimates of genetic variability at nine loci in four populations of *Spondias mombin* (standard errors in parentheses) and fixation indexes in the NE Brazil.

Populations	Average number of alleles per locus ($\hat{A}$)	% of polymorphic loci (P)	Average heterozygosity		Average of fixation index ($\hat{f}$)
			Observed ($\hat{H}_o$)	Expected ($\hat{H}_e$)	
CG-IPA	2.4	100.0	0.572	0.530	-0.08
	(0.2)		(0.115)	(0.026)	
Muribeca	2.4	100.0	0.701	0.539	-0.30
	(0.2)		(0.109)	(0.025)	
Itamaracá	2.4	100.0	0.735	0.538	-0.37
	(0.2)		(0.109)	(0.030)	
Mata Sul	2.4	100.0	0.650	0.574	-0.32
	(0.2)		(0.108)	(0.028)	

TABLE 5 – Fixation index (*f*) for nine analyzed loci in four *Spondias mombin* populations studied in the NE Brazil.

Locus	CG-IPA	Muribeca	Itamaracá	Mata Sul
Po	-0.601	-0.298	-0.747	-0.514
Acp	-0.429	-0.782	-0.653	-0.516
Est	-0.004	-0.001	-0.303	-0.088
Got	1.000	1.000	1.000	1.000
Glu	-0.498	-0.780	-0.615	-0.188
Adh	0.247	-0.143	-0.368	-0.033
Alp	-0.541	-0.695	-0.699	-0.473
Mdh	0.457	-0.243	-0.267	-0.135
Sod	-0.055	-0.534	-0.303	-0.073

TABLE 6 – F-statistical calculated for nine loci in four populations of *Spondias mombin* in the NE Brazil.

Locus	F_{IS}	F_{IT}	F_{ST}
Pó	-0.530*	-0.493*	0.024 [ns]
Acp	-0.591*	-0.563*	0.018 [ns]
Est	-0.099	-0.092 [ns]	0.006 [ns]
Got	1.000*	1.000*	0.111*
Glu	-0.513*	-0.492*	0.014 [ns]
Adh	-0.098 [ns]	-0.044 [ns]	0.049 [ns]
Alp	-0.598*	-0.582*	0.010 [ns]
Mdh	-0.049 [ns]	-0.043 [ns]	0.006 [ns]
Sod	-0.241 [ns]	-0.235*	0.005 [ns]
Average	-0.191	-0.175	0.026

* $p < 0.05$, ns not significant

TABLE 7 – Genetic identity (*GI*), estimated genetic differentiation (F_{ST}) average Neil's (1978) and gene flow (*Nm*) at the combined populations of *Spondias mombin* in the NE Brazil.

Combinations	GI	F_{ST}	Nm
CG-IPA x Muribeca	0.965	0.023	2.65*
CG-IPA x Itamaracá	0.960	0.028	2.17*
CG-IPA x Mata Sul	0.988	0.014	4.40*
Muribeca x Itamaracá	1.000	0.006	10.35
Muribeca x Mata Sul	0.986	0.014	4.40
Itamaracá x Mata Sul	0.989	0.019	3.23

* Represent gene flow of relatives before establishing of CG-IPA collection.

CONCLUSIONS

1-There is a high genetic diversity in cajá tree populations in the Zona da Mata in Pernambuco with great values within than among populations, probably favored by higher gene flow.

2-Although alteration in allele frequencies had been observed according to Hardy-Weinberg equilibrium, *S. mombin* populations are not on inbreeding or vulnerable to gene loss.

3-The CG-IPA population (IPA germplasm collection) showed little differentiation compared to other populations and therefore is a good representative of the genetic diversity of cajá tree in the Zona da Mata in the state of Pernambuco.

REFERENCES

AIRY SHAW, H.K; FORMAN, L.L. The genus *Spondias* L. (Anacardiaceae) in tropical Asia. **Kew Bulletin**, London, v.21, n.1, p.1-20. 1967.

AJAO, A.; SHONUKAN, O.; FEMI-ONADEKO, B. Antibacterial effect of aqueous and alcohol extratcts of *Spondias mombin* and *Alchornea cordifolia* - two local antimicrobial remedies. **International Journal of Crude Drug Research,** Leisse, v.23, p.67-72, 1985.

ALFENAS, A.C.; PETERS, I.; BRUNE, W.; PASSADOR, G.C. **Eletroforese de proteínas e isoenzimas de fungos e essências florestais**. Viçosa: Universidade Federal de Viçosa, 1998. 242p.

BAWA, K.; OPLER, P.A. Dioecism in tropical forest trees. **Evolution**, Boulder, v.29, p.167-179, 1975.

BOSCO, J.; SOARES, K.T.; AGUIAR FILHO, S.P.; BARROS, R.V. **A cultura da cajazeira.** João Pessoa:

EMEPA, 2000. 29p.

BOSCO, J.; SOARES, K.T.; AGUIAR FILHO, S.P.; BARROS, R.V. Banco ativo de germoplasma de cajá no Estado da Paraíba. In: WORKSHOP PARA CURADORES DE BANCOS DE GERMOPLASMA DE ESPÉCIES FRUTÍFERAS, 1997, Brasília. **Anais...** p.80-85.

BRAGA, R. Cajá. In: **Plantas do Nordeste especialmente do Ceará**. 3. Mossoró: ESAM. 1976. p. 103. (Coleção Mossoroense, 42).

BROWN, A.H.D.; MORAN, G.F. Isozymes and the genetic resources of forest trees. In: CONCKLE, M. T. **Isozymes of North American forest insects**. Bekeley: Department Agriculture, 1981. p.1-10.

CAVALCANTE, P.B. **Frutas comestíveis da Amazonas**. Belém: INPA, 1976. 166p.

CORRÊA, M.P. **Dicionário das plantas úteis do Brasil e das exóticas cultivadas**. Rio de Janeiro: Ministério da Agricultura/IBDF, 1984. 6 v.

CROAT, T.B, **Flora of Barro Colorado Island**. Stanford: Stanford University Press, 1978. 943p.

CROW, J.E.; KIMURA, M.A. **An introduction to population genetics theory**. London: Harper Row, 1970. 591p

CROW, J.F.; AOKI, K. Group selection for polygenic behavioral trait: estimating the dgree of population subdivision. **Proceeding of Natural Academy of Sciences USA**, v.81, p.6073-6077, 1984.

DANIEL, J.N. **Useful plants of Chana.** Kiev: The Royal Botanic Garden, 1990.

HAMRICK, A.D.H.; GODT, M.J.W. Allozyme diversity in plant species. In: BROWN, A.D.H.; CLEG, M.T.; KAHLER, A.L.; WEIR, B.S. **Plant population genetics, breeding and genetic resources**. Massachusetts: Sinauer, 1990. p.43-63.

HAMRICK, J.L. The distribution of genetic variation within and among natural plant populations In: SCHONEWALD- COX, C.M.; CHAMBERS, S.M.; MACBRYDE, B.; THOMAS, W.L. (Ed.). **Genetics and conservation:** a reference for managing wild animal and plant populations. London: Benjamin-Cummings, 1983. p.335-348.

JANZEN, D.H. *Spondias mombin* is culturally deprived in megafauna-free forest. **Journal Tropical of Ecology,** Cambridge, v.1, p.131-155, 1985.

KAGEYAMA, P.Y.; SEBBENN A.M.; RIBAS, L.A.; GANDARA, F.B.; CASTALLEN, M.; PERECIM, M.B.; VENCOVSKY R. Diversidade genética em espécies arbóreas tropicais de diferentes estagiosucessionais por marcadores genéticos. **Scientia Forestalis**, Piracicaba, v.64, p.93-107, 2003.

LORENZI, H. **Árvores brasileiras:** manual de identificação e cultivo de plantas arbóreas nativas do Brasil. Nova Odessa: Plantarum, 1992. 368p.

MELO Jr., A.F.; CARVALHO, D.; POVOA, J.S.R.; BEARZOTI, E. Estrutura genética de populações naturais de pequizeiro (*Caryocar brasiliense* Camb.) **Scientia Forestalis**, Piracicaba, v.66, p.56-65, 2004.

NEI, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. **Genetics,** Chapel Hill, v.89, p.583-590, 1978.

OLIVEIRA, A.F.; CARVALHO, D.; ROSADO, S.C.S. Taxa de cruzamento e sistema reprodutivo de uma população natural de *Copaifera langsdorffii* Desf. na região de Lavras (MG) por meio de isoenzimas. **Revista Brasileira de Botânica**, São Paulo, v.25, p.331-338, 2002.

OLIVEIRA, C.A.M.; SILVA, E. F.; MOLINA, S.G.; FERREIRA R.; LIRA, D.A.S.; JUNIOR, J.A.B.B. Diversidade e estrutura genética em populações de *Caesalpinia echinata* (Lam.) na Estação Ecológica do Tapacurá-PE. **Scientia Forestalis**, Piracicaba, v.70, p.77-83, 2006.

PENNINGTON, T.D.; SARUKHÁN K. J. **Manual para la identificación de campo de los principales árboles tropicales de México.** México: INIF-FAO-SAG, 1968. 413p.

SEBBENN, A.M.; KAGEYAMA, P.Y.; VENCOVSKY, R. Conservação genética *in situ* e número de matrizes para a coleta de sementes em população de *Genipa americana* L. **Scientia Forestalis**, Piracicaba, v.63, p.13-22 2003.

SEOANE, C.E.S.; KAGEYAMA, P.Y.; SEBBENN, A.M. Efeitos da fragmentação florestal na estrutura genética de populações de *Esenbeckia leiocarpa* Engl. (Guarantã). **Scientia Forestalis**, Piracicaba, v.57, p.23-139, 2000.

SILVA, A.Q.; SILVA, H. Cajá, uma frutífera tropical, **Informativo da Sociedade Brasileira de Fruticultura**, Caçador, v.14, n.4, p.12-13,1995.

SLATKIN, M. Rare alleles as indicators of gene flow. **Evolution**, Boulder, v.39, p.53-65, 1985 1985

SLATKIN, M.; BARTON, N.H. A comparison of three indirect methods for estimating average levels of gene flow. **Evolution**, Boulder, v.43, p.1349-1368, 1989.

SOUZA, L.M.F.I.; KAGEYAMA, P.Y.; SEBBENN, A.M. Estrutura genética em populações fragmentadas de *Chorisia speciosa* St. Hil (Bombacaceae). **Scientia Forestalis,** Piracicaba, v.65, p.70-79, 2004.

SWOFFORD, D.L.; SELANDER, R.B. **Biosys-1:** a computer program analysis of allelic variantion in population genetics and biochemical systematics. Release 1,7. Champaign: Illinois Natural History Survey, 1989. 43p.

TELLES, M.P.C.; VALVA, F.D.; BANDEIRA, L.F.; COELHO, A.S.G. Caracterização genética de populações naturais de araticunzeiro (*Annona crassiflora* Mart.-Annonaceae) no Estado de Goiás. **Revista Brasileira de Botânica**, São Paulo, v.25, n.1, p.123-129, 2003.

WRIGHT, S. The interpretation of population estructure by F-statistic with special regard to systems of mating. **Evolution**, Boulder, v.19, p.395-420, 1965.

YOUNG, A.; BOYLE, T.; BROWN, T. The population genetic consequences of habitat fragmentation for plants. **Trend in Ecology and Evolution**, Amsterdan, v.11, n.10, p.413-418, 1996.