Molecular characterization of *Platonia insignis* Mart. ("*Bacurizeiro*") using inter simple sequence repeat (ISSR) markers

I. G. B. Souza · V. A. B. Souza · P. S. C. Lima

Abstract *Platonia insignis* Mart. (Clusiaceae) is widespread throughout the Amazon and adjacent areas. The fruits (known locally as "*bacuri*") have significant commercial potential, but the species is under threat from agro-industrial expansion. The genetic variability within 72 genotypes of *P. insignis* belonging to ten populations collected in the Brazilian states of Maranhão and Piauí, and maintained in the germplasm collection of Embrapa Meio-Norte, has been determined, and the organization of genetic diversity within populations, between populations and among geographic groups verified. Eighteen selected inter simple sequence repeat primers allowed amplification of 236 loci of which 221 (93.64 %) were polymorphic, indicating a high level of genetic diversity. At the population level, the Shannon and Nei diversity indices ranged from 0.082 to 0.323 and from 0.120 to 0.480, respectively. The global coefficient of genetic differentiation (G_{ST}) was 0.4730 indicating that differentiation between populations was significant, a finding that was confirmed by analysis of molecular variance ($\Phi_{ST} = 0.28$). UPGMA cluster analysis revealed that the genotypes could be stratified into groups that were well defined and consistent with those identified in the dendrogram constructed using pair wise Φ_{ST} values. The high genetic diversity established in this study may facilitate the management and conservation of the germplasm of *P. insignis*.

Keywords Analysis of molecular variance · Bacuri · Genetic structure · Unweighted pair group method

Introduction

The recent growth in consumer preference for tropical fruits with interesting aromas, unusual flavours and/or enhanced nutritional properties, has provided new opportunities for the agro-industrial sector and has stimulated interest in the diversification of cultures and the transformation of traditional fruits into added-value products. In this respect, Brazil is privileged with a diversity of soil types and climatic conditions that favour the culture and economic exploration of exotic tropical fruits [1], particularly in the northern and northeastern regions of the country. One species with significant economic potential is *Platonia insignis* Mart. (Clusiaceae), a deciduous tree (known locally as "*bacurizeiro*") that is widespread throughout the Amazon and in the adjacent cerrado regions of the states of Maranhão (MA) and Piauí (PI) [2, 3]. The fruit of *P. insignis* ("*bacuri*") (Fig. 1) can be consumed *in natura* as well as in various processed forms [4] and constitutes a viable option for exploitation in the internal and external markets.

Since the early 1990 s, the valuable genetic resources of *P. insignis* have been threatened by exploitation of the wood by the timber industry, while the expansion of cattle farming has significantly reduced natural plant regeneration

V. A. B. Souza—Deceased.

Electronic supplementary material The online version of this article (doi:10.1007/s11033-012-2462-6) contains supplementary material, which is available to authorized users.

I. G. B. Souza
Programa de Pós-graduação em Genética e Melhoramento, Universidade Federal do Piauí, Campus Agrícola da Socopo, Teresina, PI 64049-550, Brazil
e-mail: isisgomesmd@hotmail.com

V. A. B. Souza · P. S. C. Lima (✉)
EMBRAPA Meio-Norte, Av. Duque de Caxias 5650, Teresina, PI 64.006-220, Brazil
e-mail: sarmanho@cpamn.embrapa.br

Fig. 1 Examples of fruits of some genotypes studied. **a** BGB11, **b** BGB 16, **c** BGB 15, **d** BGB 66, **e** BGB 7, **f** BGB 32

through damage caused by trampling and grazing of shoots that arise from the roots. Urban growth, especially of cities located along the coastlines of the states of Pará (PA) and MA, has also contributed to the destruction of genetic diversity within *P. insignis* [5]. In order to protect genetic heritage and, at the same time, to encourage the sustainable exploitation of resources, Embrapa Meio-Norte (Teresina, PI, Brazil) has conducted conservation programmes involving the collection, characterization and evaluation of germplasm of *P. insignis*, and has instituted research into the development of methods of propagation and cultivation [2].

Characterization of the genetic resources of crop plants has typically been based on long-term observations of agronomical and morphological attributes [6–8], but rapid and powerful molecular techniques are now widely available. The use of DNA markers in germplasm screening, for example, has become universal since it allows the discrimination of plants that are morphologically indistinguishable [9–11]. Inter simple sequence repeat (ISSR) markers represent an example of a genotyping technique that has been

employed successfully in the evaluation of the diversity and structuration of intraspecific populations [12–14]. ISSRs are regions of the genome located between adjacent SSR (or microsatellite) loci, and the corresponding markers may be generated by a polymerase chain reaction (PCR) using a simple primer based on a two nucleotide multi-repeat sequence, such as $(AC)_n$, with a 3′ anchor. On this basis, the technique is particularly appropriate for application to the study of plant species that have not been fully sequenced [15], as is the case for *P. insignis*. Furthermore, ISSR markers comprise long nucleotide sequences that have high annealing temperatures, and this improves the precision of amplification and reproducibility [16, 17].

The application of DNA marker techniques should allow the rapid identification of genetic variability in the germplasm of *P. insignis*, knowledge of which is essential in the characterization of traits, the development of conservation strategies and the selection of appropriate genitors. A previous study involving random amplified polymorphic DNA (RAPD) analysis of *P. insignis* revealed a high diversity among the genotypes originating from the states

of MA, PI and Ceará (CE), which showed 93.5, 87.2 and 80.4 % polymorphism, respectively [18]. The objectives of the present study were: (i) to determine the genetic variability among 72 genotypes originating from ten populations of *P. insignis* maintained in the germplasm collection of Embrapa Meio-Norte, and (ii) to verify the organization of genetic diversity within the different hierarchical levels, namely, within populations, between populations of the same group and among different groups.

Materials and methods

Genetic material

Seventy-two genotypes of *P. insignis* were available in the germplasm collection of Embrapa Meio-Norte established at Teresina, PI, Brazil (latitude 5°05′S, longitude 42°49′W, altitude 72 m above sea level). These genotypes had been collected in eight different areas of MA and two areas of PI (Fig. 2), thus providing ten diverse populations. The climate of the collection region is hot and humid with annual precipitation ranging from 700 to over 2,000 mm. The flora of the region is typical of a transition zone between semi-arid, equatorial and tropical climates, and the soil presents features that favour cultivation but with some limitations such as elevated acidity. Details of the genotypes analysed and their areas of collection are presented in Table S1: genotypes bearing similar codes constitute half-sibling progenies.

Extraction of genomic DNA

Young leaves were collected from actively growing plants of each genotype and stored in ice until required for extraction. DNA was extracted from plant material with the aid of DNeasy Plant Mini Kits (Qiagen Biotecnologia Brazil, São Paulo, SP, Brazil) using the protocol suggested by the manufacturer. In order to verify the quality of the extracted DNA, aliquots were submitted to electrophoresis on 0.8 % agarose gels in 0.5× Tris–borate–EDTA (TBE) buffer and subsequently stained with 1× SYBR® Safe DNA gel stain (Invitrogen Brazil, São Paulo, SP, Brazil). The separated bands were compared with those produced by 150 ng of λ DNA. The amount of DNA in each extract was quantified using a NanoDrop 2000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA), following which

Fig. 2 Map of the Brazilian states of Maranhão (MA) and Piauí (PI) showing the approximate collection points of *Platonia insignis* genotypes. Key to locations: *1* Aldeias Altas–MA, *2* Carolina–MA, *3* Caxias–MA, *4* Matões–MA, *5* Passagem Franca–MA, *6* Santa Quitéria do Maranhão–MA, *7* Timon–MA, *8* Urbano Santos–MA, *9* Barras–PI, and *10* Palmeirais–PI
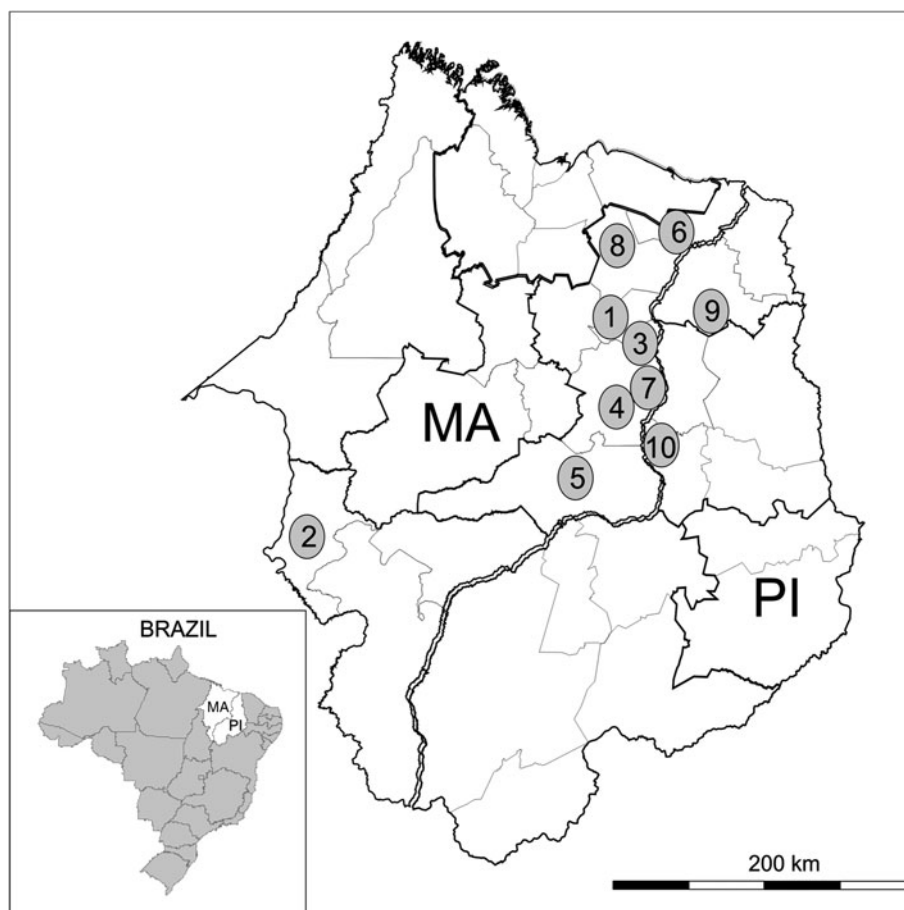
samples were diluted with water to a final concentration of 5 ng/μl and stored at −20 °C.

Amplification of DNA by PCR

In order to select the most appropriate primers for the molecular characterization of the genotypes, 100 semi-arbitrary primers of 15–18 nucleotides each (University of British Columbia, Vancouver, Canada) were tested by PCR and the optimal annealing temperatures determined. Primers that generated the largest numbers of polymorphic bands were employed in the subsequent PCR analysis (Table S2).

PCR experiments were carried out in a Veriti^{TM} 96-Well Fast Thermal Cycler (Applied Biosystems, Life Technologies, Carlsbad, CA, USA). Reaction mixtures were prepared in 0.2 ml microtubes and contained buffer solution [13.5 mM Tris–HCl (pH 8.3) with 67.5 mM KCl and 2.0 mM MgCl$_2$], 200 μM of each dNTP, 0.7 μM primer, 1 U Taq DNA polymerase (New England BioLabs, Ipswich, MA, USA), 5 ng DNA template and distilled water to a final volume of 20 μl. The reaction programme involved an initial denaturation step of 10 min at 95 °C, followed by 35 cycles comprising denaturation for 1 min at 95 °C, annealing for 45 s at 50–62 °C (depending on the primer) and extension for 2 min at 72 °C, with a final extension step of 5 min at 72 °C. Amplicons were separated by horizontal electrophoresis on 1.5 % agarose gel in 0.5× TBE buffer at 80 V for 3.5 h, together with 100 bp and 1 kb DNA ladders (Invitrogen) as molecular weight markers. Gels were treated with 12.5× Gel Red (Biotium, Hayward, CA, USA) and bands were visualised and photographed under a MiniBIS Pro® transilluminator (Bio-America Inc., Miami, FL, USA).

Statistical analysis

The amplification products were designated as a single character, the presence of which was represented by a "1" and the absence by a "0", and the ISSR markers were thus transformed into a binomial matrix. Since the ISSR marker is dominant, each band was assumed to represent the phenotype in a biallelic locus [19]. Genetic similarities (sg_{ij}) between P. insignis genotypes were estimated as Jaccard coefficients obtained from the expression

$sg_{ij} = a/a + b + c$

and a similarity matrix was generated. A dendrogram was constructed by applying unweighted pair group method with arithmetic average (UPGMA) cluster analysis using a cut-off point based on the mean genetic similarity of all studied pairs. The fit between the similarity matrix and the dendrogram was estimated from the cophenetic correlation coefficient (r) [20], while the stability of groups was computed

by bootstrapping with 1,000 replicates. The aforementioned analyses were performed with the aid of PAST software version 1.34 [21].

A Bayesian analysis of the ISSR population was performed on the entire data set using the program STRUC-TURE (version 2.3) [22] to detect population structure, to estimate the number of populations in a sample and to assign individuals to one or more of these populations. The number of genetically distinct clusters (K) was set to vary from 1 to 10 (the total number of populations), and the model was run for 100,000 independent simulations for each K, with a burn-in length of 5,000 and a run length of 50,000 iterations. Following the dominant marker settings available in the program, the "no admixture" model was used, and "uncorrelated allele frequencies among populations" were assumed. The most likely number of clusters was estimated according to the procedure described by Evanno et al. [23] using the ΔK statistic based on the rate of change with respect to K in the log probability of data.

Genetic diversity within P. insignis populations was estimated using the following parameters: (i) the percentage of polymorphic loci [$P \% = $ (polymorphic loci/total loci) × 100], (ii) the observed (N_a) and the effective (N_e) numbers of alleles [24], where N_e in the number of equally frequent alleles that would be necessary to achieve the same the degree of genetic diversity produced by N_a, (iii) the Shannon diversity index H' calculated from

$$H' = \left(- \sum p_i \log_2 p_i \right) \Big/ L$$

where p_i is the frequency of ISSR fragments amplified in the population and L is the total number of fragments [25], and (iv) the Nei [26] genetic diversity index H_e. For each locus, the Nei index produces values between 0 and 0.5, while the Shannon index varies from 0 to 0.73 according to a natural log scale [27].

The mean diversity in all ten P. insignis populations, derived from

$$H_{POP} = (1/n) \sum H'$$

where n is the number of populations, and the mean diversity at the species level, given by

$$H_{SP} = \sum (p_S \log p_S) \Big/ L$$

where p_s is the frequency of the presence or absence of ISSR loci in all populations, were calculated with the aid of the POPGENE program version 1.31 [28], assuming that the populations were in Hardy–Weinberg equilibrium. The global coefficient of genetic differentiation (G_{ST}) of the ten populations was calculated from the expression

$$[G_{ST} = (H_{SP} - H_{POP})/H_{SP}]$$

[29] using the POPGENE program [28]. The G_{ST} coefficient indicates the level of differentiation according to the scale low (0.00–0.05), moderate (0.06–0.15), high (0.16–0.25) or very high (>0.25) [30].

Genetic variation within and between *P. insignis* populations, as well as within and between groups, was estimated using analysis of molecular variance (AMOVA), the results of which were applied to the determination of a fixation index Φ_{ST} (analogous to the F_{ST} index of Wright [31]) with the aid of Arlequin software version 3.1 [32]. In order to verify the occurrence of genetic structuration, AMOVA was performed for all loci considering two different hierarchical levels: (i) all locations were included in a single group and the differences between the populations (Φ_{ST}) estimated; and (ii) the locations were grouped according to their geographical origin, namely, MA or PI, and compared in terms of the occurrence of genetic structuration between the population groups (Φ_{CT}), and between the populations present within each group (Φ_{SC}).

The pair wise Φ_{ST} values (distances between each pair of populations) were examined in a matrix and, based on these values, an UPGMA dendrogram was constructed to illustrate the relationships between the ten populations. The significances of the components of variation and the Φ statistics were evaluated using 1,000 permutations in each analysis.

Results

Genetic diversity and cluster analysis

PCR assays using the 18 primers selected in the initial tests allowed 236 ISSR loci to be amplified from the DNA extracts derived from 72 genotypes of *P. insignis*. The average number of loci per primer was 13, with the largest number of loci ($n = 21$) being detected with primer UBC810 and the smallest number ($n = 9$) with primers UBC842 and UBC890. Of the 236 amplified loci, 221 (93.64 %) showed polymorphism at the species level (all of the 72 genotypes) (Table S3), while the size of these loci ranged between 100 and 2,072 bp. The ISSR amplification profile produced by primer UBC 888 is shown in Fig. 3.

The highest percentage of polymorphic loci (*P* %) was established in the Barras–PI population (89.41 %), while the lowest was observed in the population from Matões–MA (19.92 %). Assuming that the populations were in Hardy–Weinberg equilibrium, the observed numbers of alleles (N_a) varied between 0.894 (Barras–PI) and 1.593 (Carolina–MA) with a mean value of 1.38, while the effective number of alleles (N_e) ranged from 1.140 (Matões–MA) to 1.558 (Barras–PI) with a mean value of 1.30. At the species level, the values of N_a and N_e were 1.93 and 1.57, respectively. The Shannon (H') and Nei (H_e) diversity indices revealed significant differences between the populations. H' values of the populations varied between 0.120 and 0.480 with a value of 0.499 at the species level, while H_e values of the populations varied between 0.082 and 0.323 with a value of 0.335 at the species level (Table S4). The values of H_{POP} and H_{SP}, which express the mean genetic diversities within the populations and the species, respectively, were 0.324 and 0.171.

The coefficients of genetic similarity among the 72 genotypes of *P. insignis* were in the range 0.35 to 0.93 with an average value of 0.52 (data not shown). The dendrogram constructed using the UPGMA method (based on a 52 % similarity cut-off point) revealed the genetic interrelationships amongst the genotypes and allowed nine genotypic groups to be identified (Fig. 4). These results demonstrated the existence of considerable genetic variability in the populations of *P. insignis* studied. Furthermore, the dendrogram showed that the genotypes originating from MA ($n = 41$) were well separated from those from PI ($n = 31$), indicating that such genotypes can be grouped according to geographical area even though some populations, such as BGB5 and BGB78 from Matões–MA, were distant from others collected in the same state. The cophenetic correlation coefficient ($r = 0.79$; $P = 0.001$) indicated a good fit between the original similarity matrix and the results from cluster analysis. Analysis of the ISSR population using the program STRUCTURE program revealed that the highest probability of the data was attained when samples were clustered into four groups (K = 4; Fig. 5; Table S5). These groupings were consistent with those derived from the UPGMA clustering results.

Genetic structure

The global coefficient of genetic differentiation (G_{ST}) was 0.4730 indicating a high level of differentiation among the populations of *P. Insignis*, while AMOVA revealed significant genetic differentiation ($\Phi_{ST} = 0.28$; $P < 0.001$) between the ten populations studied. The distribution of the total genetic diversity was represented by 28.18 % between the populations and 71.82 % within the populations (Table S6). Although differentiation between the two geographical groups (MA and PI) was not significant ($\Phi_{CT} = -0.02$), that between the populations within each of the groups was significant ($\Phi_{SC} = 0.29$). After grouping the populations according to the state of collection, the results showed that most of the genetic variability observed in *P. insignis* was due to differences within the populations (72.36 %; Table S6).

Pair wise Φ_{ST} values indicated significant levels of differentiation between the populations of *P. insignis*. The distance between the populations from Barras–PI and
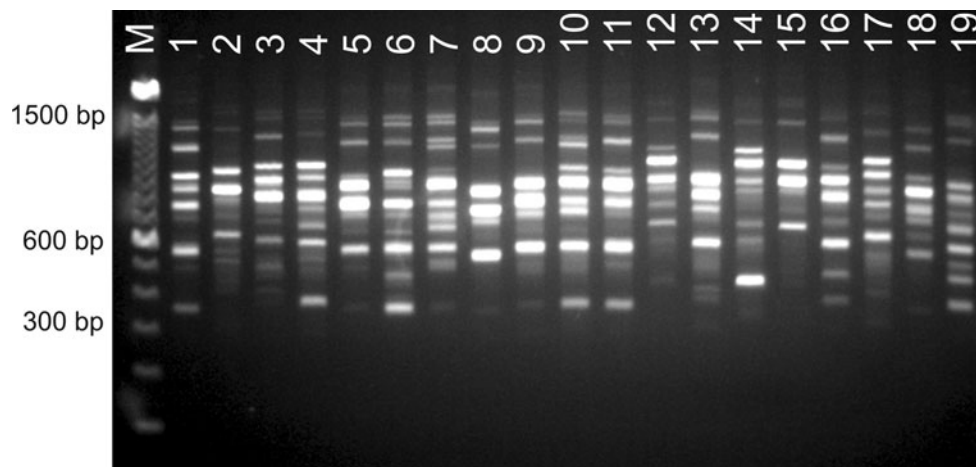
Fig. 3 ISSR amplification profiles of different *Platonia insignis* genotypes using primer UBC 888

Palmeirais–PI was small (0.14046), while the distance between the populations from Aldeias Altas–MA and Palmeirais–PI was large (0.51133) (Table S7). The dendrogram based on pair wise Φ_{ST} values (Fig. 6) clearly showed the phylogenetic relationships between the 10 populations studied.

Discussion

The principal value of plant genetic resources resides in their utilization in the production of new cultivars, the domestication of new species and the development of novel agro-industrial products. In this context, a germplasm bank should comprise sufficient genetic variants to supply breeders with the materials required for the efficient and economical production of high-quality products demanded by the consumer market [14]. The present paper describes the characterization by ISSR markers of 72 genotypes of *P. insignis*, a species with significant commercial potential, maintained in the germplasm collection of Embrapa Meio-Norte.

The quantification of polymorphic loci is a parameter frequently used in the evaluation of the genetic diversity of a population. Genetic polymorphism has been defined as the occurrence in the same population of two or more alleles at a given locus with a frequency of up to 95 % [33]. Molecular markers can provide important information regarding genetic polymorphism, knowledge of which is essential for the development and improvement of plant populations [34]. In the case of the ten populations of *P. insignis* studied, 221 (93.64 %) out of the 236 ISSR loci were polymorphic, a value considered to be very high and one that reveals the large genetic variability

of this species. A similar high level of diversity (87.15 %) has been reported previously by Almeida et al. [18] following studies involving the characterization of *P. insignis* populations collected from CE, PI and MA using RAPD markers.

In the ISSR analysis reported herein, the number of observed alleles (N_a) was 1.93, while the number of alleles (N_e) that effectively contributed to the genetic diversity at the species-level was 1.57 (Table S4). These values were close to those obtained for the perennial species *Pogostemon cablin* using ISSR markers, namely, $N_a = 1.92$ and $N_e = 1.46$ [35].

The Shannon (H') and Nei (H_e) diversity indices are often employed in evaluating the abundance and uniformity of alleles present at one locus [26, 36]. In the present study, these indices were found to be very high at the species level ($H' = 0.499$ and $H_e = 0.335$), while the level of variation also differed according to the population. The greatest diversity was detected in the Barras–PI population ($H' = 0.480$ and $H_e = 0.323$), indicating that this population may contain parental genotypes that could be used in hybridization programmes aimed at greater heterosis. In contrast, the least diversity was found in the Matões–MA population, which also presented a much lower level of polymorphism (19.92 %). However, this population comprised only two genotypes and its true variability may not have been fully expressed. There is clearly a need to collect more samples of *P. insignis* from this location in order to improve its representation in the germplasm collection of Embrapa Meio-Norte.

The appropriate selection of parental combinations depends on the understanding of relationships between genotypes [37, 38]. Within the germplasm collection under study, genotypes BGB61 (Carolina–MA) and BGB1 (Palmeirais–PI) were the most divergent and might, therefore, present a greater
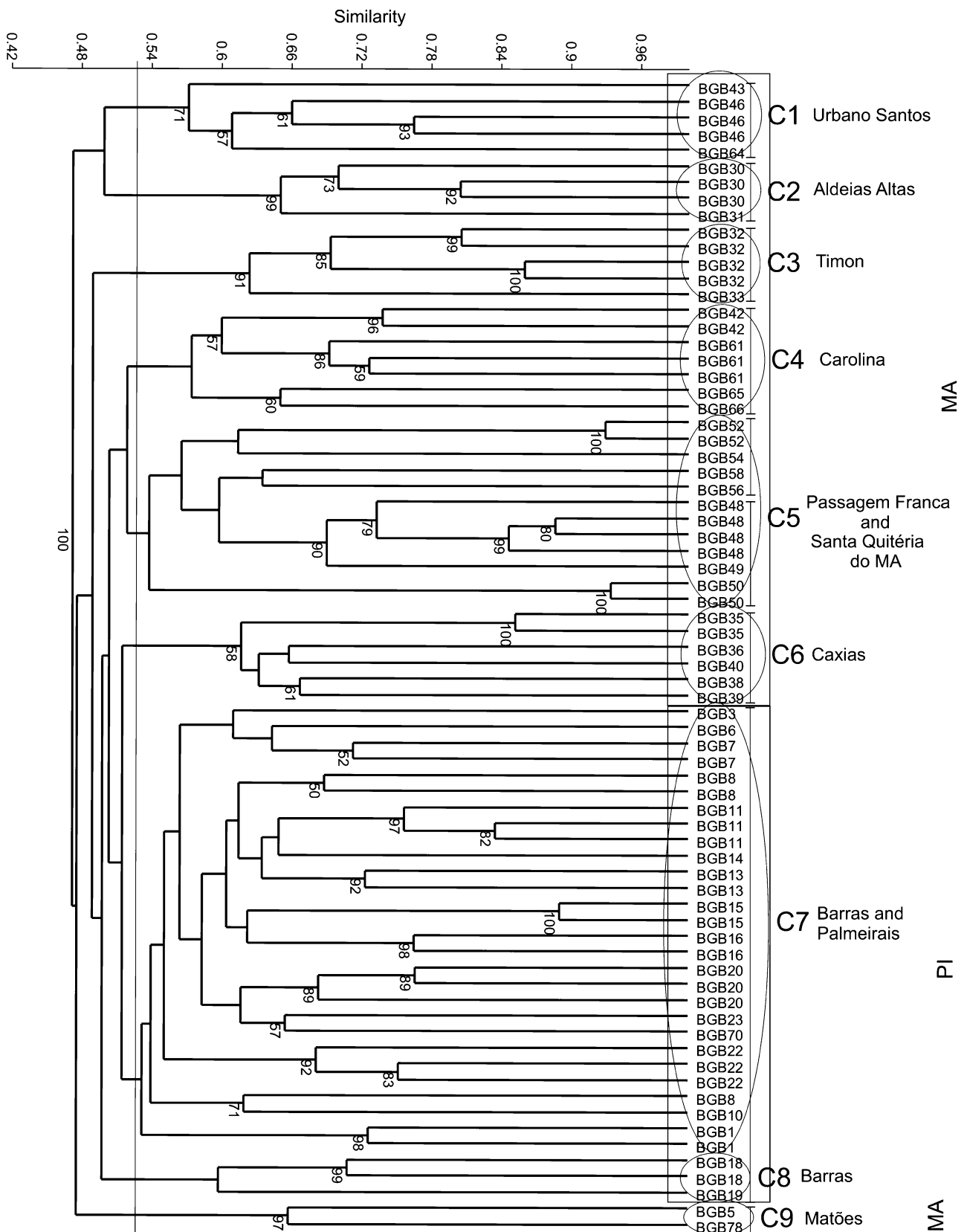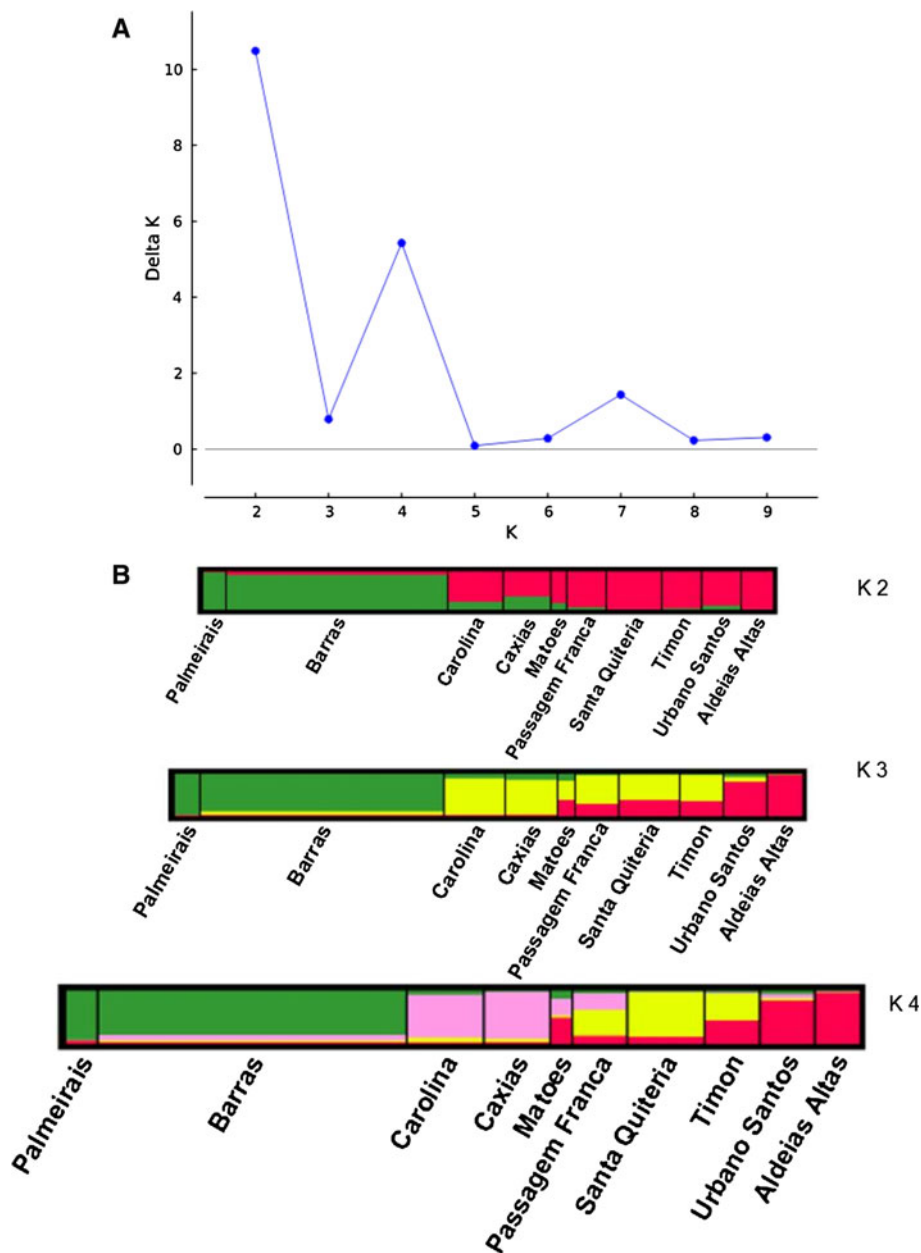
Fig. 4 UPGMA dendrogram based on Jaccard coefficients of similarity of 72 genotypes of *Platonia insignis* maintained in the germplasm collection of Embrapa Meio-Norte (Teresina, PI, Brazil). Bootstrap values ≥50 % are shown

probability of heterosis in a breeding programme aimed at improving desirable agronomical traits in *P. insignis*.

UPGMA cluster analysis (Fig. 4) proved to be an efficient method of grouping genotypes of *P. insignis* since it revealed similarities and/or differences between half-sibling progenies as well as between genotypes from unrelated sources. Most of the nine genotypic groups comprised subgroups, and the half siblings were generally grouped into one (or a small number) of these subgroups. The genotypes collected from different states (MA and PI) could be clearly separated, indicating that the genetic material may be classified according to geographical origin.

Bayesian analysis revealed that the germplasm collection of Embrapa Meio-Norte could be represented by two

(K = 2) or four (K = 4) large genotypic groups (Fig. 5). According to the separation obtained through simulation models, for K = 2, the two groups comprised the gene pools from Maranhão and Piauí. However, subdivision into a larger number of groups (K = 4), where one group encompassed the genotypes from Piauí (shaded green) and the other three groups encompassed the genotypes from Maranhão (shaded pink, yellow and red) was well supported by the results. The analysis of the standard deviations of probabilities obtained in different simulations also helps to define the number of groups and their genotypic structure [23]. As can be seen in Table S5, the separation of genotypes into four groups (K = 4) coincided with one of the lowest values of standard deviation (97.840794) of

Fig. 6 UPGMA dendrogram based on the pair wise Φ_{ST} distances derived from the analysis of molecular variance in 10 populations of *Platonia insignis* maintained in the germplasm collection of Embrapa Meio-Norte (Teresina, PI, Brazil)
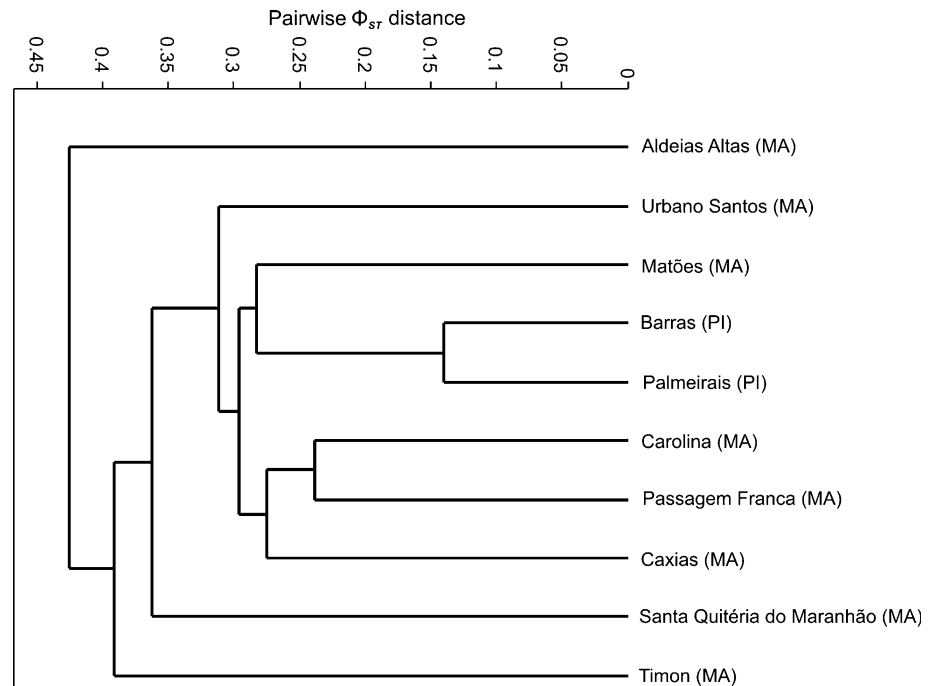
probability. Hence, the genetic structure of the studied genotypes is better represented by four groups.

The results of the present study demonstrate that there is a very high degree of genetic differentiation between *P. insignis* populations as indicated by the global index ($G_{ST} = 0.4730$) calculated over all polymorphic loci. This result was confirmed by AMOVA, which indicated that 71.82 % of the variability occurred within the populations while only 28.18 % ($\Phi_{ST} = 0.28$) occurred between the populations.

The levels of genetic diversity and the distribution of variability within and between plant populations have generally been interpreted as the result of the balanced combination of the reproductive system and the history of the species [40]. Perennial and allogamous species typically exhibit higher levels of genetic diversity within, rather than between, populations indicating the influence of biological characteristics on these parameters [41]. Other studies have also shown that perennial plants generally maintain the largest genetic variability within populations [39, 42].

In earlier studies, Oliveira and Silva [43] employed both RAPD and SSR markers to study the distribution of genetic variability in *Euterpe oleracea* Mart., a native Amazon fruit tree known locally as the *açai* palm. These authors reported genetic variability of 75.85 % determined by RAPD and 68.88 % using SSR, indicating that both markers were equally efficient in discriminating genetic variability. Since codominant markers are unavailable for *P. insignis*, the dominant ISSR marker was selected for the present study and was able to reveal the distribution of

variability both within and between the populations. Such findings confirm that this approach is reliable and efficient for the characterization of germplasm collections.

It is noteworthy that the values of Φ_{CT} recorded in the present study did not alter significantly when the genotypes were distributed according to location of collection (i.e. MA or PI). Indeed, ISSR markers are neutral and mainly amplify non-coding DNA sequences that are subject to weak selection pressures. Thus, the variations observed when using ISSR markers do not necessarily reflect alterations in the profile of adaptive genes [39]. This could explain the results reported herein, since the ISSR markers may not have amplified DNA regions containing adaptive genes that might have evolved separately at the different locations.

Statistical analysis of pair wise Φ_{ST} values revealed that all were significant, thereby indicating that the populations of *P. insignis* were genetically different, a finding that is in accord with results obtained for populations of fruit trees using other markers including SSR [42]. The dendrogram based on pair wise Φ_{ST} values (Fig. 6) showed clearly the relationships between the ten populations of *P. insignis*, and the grouping of the two populations from PI, namely, those from Barras and Palmeirais. The groupings revealed by this dendrogram are consistent with the results obtained from cluster analysis of the 72 genotypes (Fig. 4).

In conclusion, the present study has shown that ISSR markers represent a rapid, reliable and cost effective tool for the identification of genotypes and the genetic structuration of populations of *P. insignis*. In the case of the germplasm collection of Embrapa Meio-Norte, it is recommended that

the variability of the genotype pool be augmented by increasing the number of accessions in general, and especially from Barras–PI and Matões–MA since these presented the largest and the smallest variabilities, respectively. The use of ISSR markers will contribute to the conservation and breeding programmes of a potentially valuable resource that has yet to be harnessed effectively.

References

1. Simão S (1998) Tratado de fruticultura. Fundação de Estudos Agrários Luiz de Queiroz, Piracicaba

2. Souza VAB, Araújo ECE, Vasconcelos LFL, Lima PSC (2001) Variabilidade de características físicas e químicas de frutos de germoplasma de bacuri da Região Meio-Norte do Brasil. Rev Bras Frutic 23:677–683

3. Carvalho JEU, Muller CH (2007) Propagação do bacurizeiro. In: Lima MC (ed) Bacuri: agrobiodiversidade. Instituto Interamericano de Cooperação para a Agricultura, São Luís, pp 29–46

4. Ferreira FR, Ferreira SAN, Carvalho JDU (1987) Espécies frutíferas pouco exploradas, com potencial econômico e social para o Brasil. Rev Bras Frutic 9:11–22

5. Nascimento WMO, Carvalho JEU, Muller CH (2007) Ocorrência e distribuição geográfica do bacurizeiro. Rev Bras Frutic 29:657–660

6. Andrade RA, Lemos EGM, Martins ABG, Paula RC (2009) Caracterização morfológica de plantas de rambutan. Acta Sci Agron 31:613–619

7. Rakonjac V, Aksic MF, Nikolic D, Milatovic D, Colic S (2010) Morphological characterization of 'Oblacinska' sour cherry by multivariate analysis. Sci Hortic 125:679–684

8. Leão PCS, Cruz CD, Motoike SY (2011) Genetic diversity of table grape based on morphoagronomic traits. Sci Agric 68:42–49

9. Ikegami H, Nogata H, Hirashima K, Awamura M, Nakahara T (2009) Analysis of genetic diversity among European and Asian fig varieties (*Ficus carica* L.) using ISSR, RAPD, and SSR markers. Genet Resour Crop Evol 56:201–209

10. Dje Y, Tahi CG, Bi ALZ, Baudoin JP, Bertin P (2010) Use of ISSR markers to assess genetic diversity of African edible seeded *Citrullus lanatus* landraces. Sci Hortic 124:159–164

11. Tantasawat P, Trongchuen J, Prajongjai T, Seehalak W, Jittayasothorn Y (2010) Variety identification and comparative analysis of genetic diversity in yardlong bean (*Vigna unguiculata* spp. sesquipedalis) using morphological characters. SSR and ISSR analysis. Sci Hortic 124:204–216

12. Chai X, Chen SL, Xu W (2010) Using inter-simple sequence repeat markers to analyze the genetic structure of natural *Pteroceltis tatarinowii* populations and implications for species conservation. Plant Syst Evol 285:65–73

13. Yang Y, Pan Y, Gong X, Fan M (2010) Genetic variation in the endangered Rutaceae species *Citrus hongheensis* based on ISSR fingerprinting. Genet Resour Crop Evol 57:1239–1248

14. Lahoz I, Fernández JA, Migliaro D, Macua JI, Egea-Gilabert C (2011) Using molecular markers, nutritional traits and field performance data to characterize cultivated cardoon germplasm resources. Sci Hortic 127:188–197

15. Meloni M, Perini D, Filigheddu R, Binelli G (2006) Genetic variation in five Mediterranean populations of *Juniperus phoenicea* as revealed by inter-simple sequence repeat (ISSR) markers. Ann Bot 97:299–304

16. Weising K, Nybom H, Wolff K, Meyer W (1995) DNA fingerprints in plants and fungi. CRC Press, Boca Raton

17. Qian W, Ge S, Hong DY (2001) Genetic variation within and among populations of a wild rice *Oryza granulata* from China detected by RAPD and ISSR markers. Theor Appl Genet 102:440–449

18. Almeida HJS, Costa JTA, Benbadis AK, Innvecco R, Aloufa MAI, Carvalho ACPP (2007) Aplicação de marcador molecular (RAPD) para estudos da diversidade genética em bacurizeiro. In: Lima MC (ed) Bacuri: agrobiodiversidade. Instituto Interamericano de Cooperação para a Agricultura, São Luís, pp 157–170

19. Williams JGK, Rafalski JA, Kubelik AR, Livak KJ, Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

20. Sokal RR, Rohlf FJ (1962) The comparison of dendrograms by objective methods. Taxon 11:30–40

21. Hammer O, Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. Paleontol Electron 4(art. 4):9. http://palaeo-electronica.org/2001_1/past/past.pdf. Accessed 24 Dec 2012

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

23. 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

24. Lewontin RC (1972) The apportionment of human diversity. Evol Biol 6:381–398

25. Kimura M, Crow JF (1964) Number of alleles that can be maintained in finite population. Genetics 49:725–738

26. Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

27. Lowe A, Harris S, Ashton P (2004) Ecological genetics: design analysis and application. Blackwell, Oxford

28. Yeh FC, Boyle TYZ, Xiyan JM (1999) POPGENE version 1.31: Microsoft Windows based freeware for populations genetic analysis. University of Alberta and Center for International Forestry Research, Edmonton

29. Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, pp 187–192

30. Wright S (1978) Evolution and the genetics of population. University of Chicago Press, Chicago

31. Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395–420

32. Excoffier L, Laval G, Schneider S (2006) Arlequin (version 3.1): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50 http://www.la-press.com/arlequin-version-30-an-integrated-software-package-for-population-gene-article-a188. Accessed 24 Dec 2012

33. Cavalli-Sforza LL, Bodmer WF (1971) The genetics of human populations. WH Freeman, San Francisco

34. Waugh R, Powell W (1992) Using RAPD markers for crop improvement. Trends Biotechnol 10:186–191

35. Wu YG, Guo QS, He JC, Lin YF, Luo LJ, Liu GD (2010) Genetic diversity analysis among and within populations of *Pogostemon cablin* from China with ISSR and SRAP markers. Biochem Syst Ecol 38:63–72

36. Shannon CWW (1949) The mathematical theory of communication. University of Illinois Press, Urbana

37. Becelaere GV, Edward LL, Paterson AH, Chee PW, Pedigree VS (2005) DNA marker-based genetic similarity estimates in cotton. Crop Sci 45:2281–2287

38. Ali ML, Rajewski JF, Baenziger PS, Gill KS, Eskridge KM (2008) Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Mol Breed 21:497–509

39. Takrouni MM, Boussaid M (2010) Genetic diversity and population's structure in Tunisian strawberry tree (*Arbutus unedo* L.). Sci Hortic 126:330–337

40. Hamrick JL, Godt MJW, Murawski DA, Loveless MD (1991) Correlations between species traits and allozyme diversity: implications for conservation biology. In: Falk DAS, Holsinger KE (eds) Genetics and conservation of rare plants. Oxford University Press, Oxford, pp 75–86

41. Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. Ann Rev Ecol Syst 15: 65–95

42. Nas MN, Bolek Y, Bardak A (2011) Genetic diversity and phylogenetic relationships of *Prunus microcarpa* C.A. Mey. *subsp. tortusa* analyzed by simple sequence repeats (SSRs). Sci Hortic 127:220–227

43. Oliveira MSP, Silva KJD (2008) Diferenciação genética entre procedências de açaizeiro por marcadores RAPD e SSR. Rev Bras Frutic 30:438–443