Ecological Genetics and Genomics 25 (2022) 100143

Contents lists available at ScienceDirect

Ecological Genetics and Genomics

journal homepage: www.elsevier.com/locate/egg

Genetic diversity and structure of an active germplasm collection of *Annona squamosa* L.

Gisele H. Sá [a], Fatianne C.D. Lima [b], João P.G. Viana [c], Ângela C.A. Lopes [d], Leonardo C.B. Carvalho [e], Sérgio E.S. Valente [b], Paulo S.C. Lima [f,*]

[a] Programa de Pós-Graduação Em Agronomia, Centro de Ciências Agrárias, Universidade Federal do Piauí, Teresina, PI, Brazil
[b] Departamento de Biologia, Universidade Federal do Piauí, Teresina, PI, Brazil
[c] Crop Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[d] Departamento de Fitotecnia, Centro de Ciências Agrárias, Universidade Federal do Piauí, Teresina, PI, Brazil
[e] Programa Nacional de Pós- Doutoramento, Centro de Ciências Agrárias, Universidade Federal do Piauí, Teresina, PI, Brazil
[f] Embrapa Meio-Norte, Teresina, PI, Brazil

ARTICLE INFO

Keywords:
Tropical fruits
ISSR markers
Pinha
Conservation genetics
Genetic polymorphism

ABSTRACT

The fruits of *Annona squamosa* L. are highly appreciated for their sweet taste and their nutritional and medicinal properties. Since the species is threatened by habitat loss, germplasm conservation is of the utmost importance. The present study aimed to characterize genetic variability among 19 accessions of *A. squamosa* maintained in the Active Germplasm Bank (AGB) of Embrapa Meio-Norte, Teresina, Piauí, Brazil, and to determine whether the geographical distribution of the genotypes influenced population structure. The accessions were derived from populations originally located in orchards situated some 200 km apart in three municipalities in the northeast of Brazil, namely Timon, Bom Jesus and Canto do Buriti. Leaf DNA was extracted and analyzed by inter simple sequence repeat (ISSR)-polymerase chain reaction. Nine ISSR primers generated 127 scorable bands of which 101 (79.52%) were polymorphic. The within-population diversity was high (Canto do Buriti), intermediary (Timon) and low (Bom Jesus). Unweighted pair-group method with arithmetic averages and Bayesian analysis of polymorphic bands separated the genotypes into two main clusters according to genetic distance and ancestry (vegetative- and seed-propagated groups), indicating that the geographical location of the populations did not influence the genetic diversity of the accessions. While ISSR marker-based DNA fingerprinting of the 19 studied accessions revealed a degree of genetic diversity between populations, the variability detected within each population was somewhat higher. This information will contribute to a better understanding of genetic variation and population structure of the accessions of *A. squamosa* in the AGB of Embrapa Meio-Norte, and provide crucial information for future conservation and genetic improvement programs of the species.

1. Introduction

Brazil is one of the largest global suppliers of fresh fruit with around 1.24 million tons being exported in 2021, representing a value of more than 1.21 billion US$ [1]. The diversity of climate and soil type within Brazil affords favorable conditions for cultivating tropical, subtropical, and temperate fruits, including both common and exotic varieties [2]. One such exotic species is *Annona squamosa* L., is a diploid species (2n = 14 chromosomes), a small tropical tree that produces compound fruits with large and numerous seeds embedded in a white fleshy pulp. The fruits, which are known locally in Brazil as "pinha" and more widely as sugar apples or sweetsops, are highly appreciated for their sweet taste as well as their nutritive value [3–5]. In addition, the seeds are a potential source of natural oils for the production of biodiesel, while the leaves, roots, bark, and fruits contain numerous bioactive principles [6–8]. In 2017, Brazil produced 8753 tons of pinha, with the state of Bahia being the largest producer [9].

However, species of *Annona* are threatened by genetic erosion resulting from modernization of agriculture, changes in land use, and habitat fragmentation [10]. The process of genetic erosion can be circumvented, or at least minimized, by conserving the variability of the species in active germplasm banks (AGBs) which, in the case of *Annona*

* Corresponding author. Avenida Duque de Caxias, 5650 C.P 001, Buenos Aires, Teresina, PI, 64008-780, Brazil.
E-mail addresses: giselehollanda2@gmail.com (G.H. Sá), fatiannelima@hotmail.com (F.C.D. Lima), joaopaulogomesviana@gmail.com (J.P.G. Viana), acalopes@ufpi.edu.br (Â.C.A. Lopes), cbcleonardo@gmail.com (L.C.B. Carvalho), svalente@ufpi.edu.br (S.E.S. Valente), paulo.costa-lima@embrapa.br (P.S.C. Lima).

https://doi.org/10.1016/j.egg.2022.100143
Received 23 February 2022; Received in revised form 5 September 2022; Accepted 5 September 2022
Available online 10 September 2022
2405-9854/© 2022 Elsevier Inc. All rights reserved.

in general and *A. squamosa* in particular, are somewhat scarce worldwide. Such scant supply of genetic material is a limiting factor in studies on the management, preservation, and genetic improvement of the species [11].

The conservation and maintenance of germplasm require detailed knowledge of existing genetic diversity. In this context, DNA markers have proven extremely useful in germplasm characterization [12–14] and genetic improvement [15–19], for systematic and evolutionary studies [8,20–22], and in the detection of traits of economic importance [23–25]. Methods involving polymerase chain reaction (PCR)-based markers, such as inter-simple sequence repeats (ISSR), can detect polymorphism without sequencing the DNA of the target organisms, thereby affording simple, inexpensive and powerful methods for the assessment of genetic diversity.

The few reports that are available relating to the genetic diversity of *Annona* species, as determined by molecular markers, have revealed variable levels of polymorphism. Thus, Ahmad et al. [26] found that ISSR and random amplified polymorphic DNA (RAPD) primers detected, respectively, 58 and 52% polymorphism among *Annona* species grown in India, while Anuragi et al. [10] reported polymorphism values of 93 and 83% for the corresponding methods. Regarding the intra-specific diversity of *A. squamosa,* Bharad et al. [27] recorded 73% polymorphism when RAPD markers were used to analyze 11 genotypes from different Indian locations, but Guimarães et al. [28] found only 28.7% polymorphism among 64 accessions from southeastern Brazil using the same type of markers.

In view of the limited information available concerning the application of ISSR markers to the analysis of intra-specific diversity in *A. squamosa,* the objectives of the present study were to characterize the genetic variability among accessions maintained at the AGB of Embrapa Meio-Norte (Teresina, PI, Brazil) and to determine whether the population structure was influenced by the geographical location of the genotypes.

2. Materials and methods

2.1. Sampling leaf material from accessions of A. squamosa

Nineteen accessions of *A. squamosa* currently maintained in the AGB of Embrapa Meio-Norte and located in the experimental field of native fruit trees (5° 2.255′ S; 42° 48.180’ W) were selected for analysis. These accessions originated from three municipalities in the states of Maranhão and Piauí (Table 1; Fig. 1) and do not represent materials native to the region of the AGB. The accessions in the germplasm collection were propagated at Embrapa Meio-Norte by cuttings (genotypes) or by seed (progenies). Accessions M3F1, M3F2 and M3F3 are progenies originated from seeds of genotypes G2 and G5, which were obtained by vegetative propagation of cuttings collected in the municipality of Timon, in the state of Maranhão. The M1F1 and M1F2 progenies originated from seeds of the G6 and G11 genotypes and the M4F1, M4F3, M4F4, M4F6 and M4F7 progenies were obtained from seeds collected in the G8 genotype, which, like the G6 and G11 genotypes, were cloned from plants located in the municipality of Canto do Buriti in the state of Piauí.

Fresh leaves of each of these accessions were collected, wrapped in paper towels, placed in plastic bags for transportation to the Laboratory of Molecular Biology at Embrapa Meio-Norte, and stored at −20 °C until required for DNA extraction.

Table 1
Origin of the three populations of *Annona squamosa* from northeastern Brazil.

Accessions[a]	Original location - Municipality, State	Coordinates
G2 and G5 (M3F1, M3F2, and M3F3)	Timon, Maranhão	5° 6′ 16.463″ S 42° 49′ 44.108″ W
G4, G10, and G16	Bom Jesus, Piauí	9° 4′ 18.491″ S 44° 21′ 34.826″ W
G6 and G11 (M1F1, M1F2), G7, and G8 (M4F1, M4F3, M4F4, M4F6, M4F7)	Canto do Buriti, Piauí	8° 6′ 41.436″ S 42° 57′ 7.578″ W

[a] G - genotypes propagated through cuttings (vegetative propagation); MF - progenies originated from G by seed propagation.

2.2. DNA extraction and analysis

Genomic DNA was extracted using Invitrogen (Thermo Fisher Scientific, Waltham, MA, USA) PureLink™ genomic DNA mini kits following the manufacturer instructions but with two minor modifications in that the amount of plant tissue employed was increased from 25 to 50 mg and maceration was carried out in a Precellys®24 tissue homogenizer (Bertin Instruments, Montigny-le-Bretonneux, France). In order to evaluate the integrity and purity of extracted DNA, samples and a lambda DNA marker were subjected to 0.8% agarose gel electrophoresis in 0.5 X TBE running buffer (44 mM Tris; 44 mM boric acid and 1 mM ethylenediaminetetraacetic acid, pH 7.5) at 80 V for 1 h, followed by staining with GelRed® Nucleic Acid Stain (Biotium, Fremont, CA, USA). DNA was quantified spectrophotometrically and fluorimetrically using a Qubit fluorometer (Thermo Fisher Scientific) by comparison with lambda DNA of known concentration (0.1 μg/μL). DNA samples were stored at −20 °C until required for PCR amplification.

2.3. ISSR-PCR analysis

Nine 18-mer ISSR primers, acquired from the University of British Columbia (Vancouver, BC, Canada), were selected for use in the amplification reactions (Table 2). The PCR mixture contained 6.5 μL water, 1 μL amplification buffer (100 mM Tris-HCl pH 8.4 and 500 mM KCl pH 8.5), 0.1 μL Quatro G Taq DNA polymerase (1U/μL; Pesquisa e Desenvolvimento, Porto Alegre, RS, Brazil), 1 μL 0.8 mM dNTPs, 0.8 μL 0.4 μM ISSR primer and 0.6 μL of genomic DNA (7 ng/μL) in a total volume of 10 μL. Amplifications were carried out in a VeritiTM 96 Well Thermal Cycler (Thermo Fisher Scientific) using the following parameters: initial denaturation at 94 °C for 1 min; 40 cycles of denaturation at 94 °C for 40 s, annealing at 50–57 °C (primer dependent; Table 2) for 45 s, extension at 72 °C for 2 min; and final extension at 72 °C for 7 min. Amplicons from all accessions, together with 1 kb lambda DNA ladders (Thermo Fisher Scientific or Bioron, Römerberg, Germany), were submitted to electrophoresis on 1.5% agarose gel pre-stained with 0.8 μL GelRed™ in 0.5 X TBE running buffer at 110 V for 4 h. The gels were visualized on a transilluminator and images were captured and digitized using the MiniBis Pro gel documentation system (DNR Bio Imaging System, Neve Yamin, Israel).

2.4. Data analysis

Well-resolved amplicons produced by the same ISSR primer and present in the same location on gels of different accessions were considered to belong to the same gene locus and were scored visually as absent (0) or present (1) for each of the 19 genotypes. A binary similarity matrix was constructed from the ISSR marker scores and a distance matrix obtained using the expression 1-J where J is the Jaccard coefficient. Cluster analysis was carried out by applying the unweighted pair-group method with arithmetic averages (UPGMA). Bootstrap analysis with 1000 replicates was performed to test the reliability of individual branching points of the phylogenetic tree using the approximately unbiased (AU) test. All calculations were performed with the help of R software [29] using the algorithm *pvclust* to calculate the *P* values for the AU test [30].

The parameters of genetic diversity within and between populations were estimated using GPOPGENE software version 1.32 [31], and these

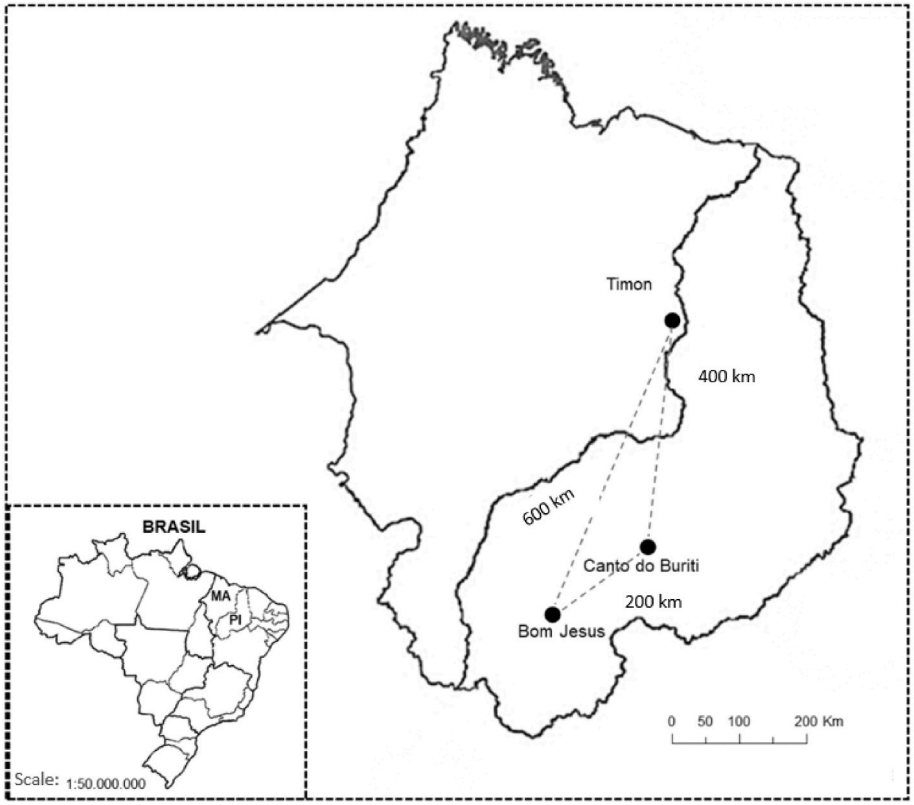

Fig. 1. Original locations of the populations of *Annona squamosa* in the Brazilian states of Maranhão (MA) and Piauí (PI) showing "straight-line" distances between the municipalities of Timon, Canto do Buriti and Bom Jesus. The maps were prepared from the cartographic base of the Instituto Brasileiro de Geografia e Estatística.

Table 2
ISSR primers used in the analysis of the genetic diversity of *Annona squamosa* accessions.

Primer	Tm[a] (°C)	Ta[b] (°C)	Sequence 5′-3′	GC (%)	Number of loci Total	Polymorphic
UBC 812	45.7	51.0	GAG AGA GAG AGA GAG AA	47.06	18	18
UBC 826	57.2	57.0	ACA CAC ACA CAC ACA CC	52.94	13	10
UBC 827	53.0	54.0	ACA CAC ACA CAC ACA CG	52.94	17	13
UBC 834	49.2	54.0	AGA GAG AGA GAG AGA GYT	50.00	14	13
UBC 841	48.5	50.0	GAG AGA GAG AGA GAG AYC	55.55	14	10
UBC 842	48.8	52.0	GAG AGA GAG AGA GAG AYG	55.55	15	10
UBC 856	52.8	54.0	ACA CAC ACA CAC ACA CYA	50.00	13	12
UBC 864	43.6	50.0	ATG ATG ATG ATG ATG ATG	33.33	13	10
UBC 890	52.2	55.0	VHV GTG TGT GTG TGT GT	58.82	10	5
Total	–	–	–	–	127	101

[a] Melting temperature.
[b] Annealing temperature.

included the number and percentage of polymorphic loci, Nei's genetic diversity index (*H*) [32], Shannon's information index (*I*) [33], the genetic differentiation coefficient (G_{ST}) and the number of migrants per generation (N_m), the latter being calculated from the G_{ST} value. In order to infer population structure, nonhierarchical Bayesian clustering was performed using the program STRUCTURE 2.3.4 [34] adopting the admixture model with correlated allelic frequencies. Ten independent runs were performed for each value of *K*, which ranged from 1 to 10, with 1,000,000 Markov Chain Monte Carlo simulations and 500,000 burn-in generations. The most probable *K* was determined by calculating delta-*K* according to the method described by Evanno et al. [35] with the aid of STRUCTURE HARVESTER software version 0.6.9 [36].

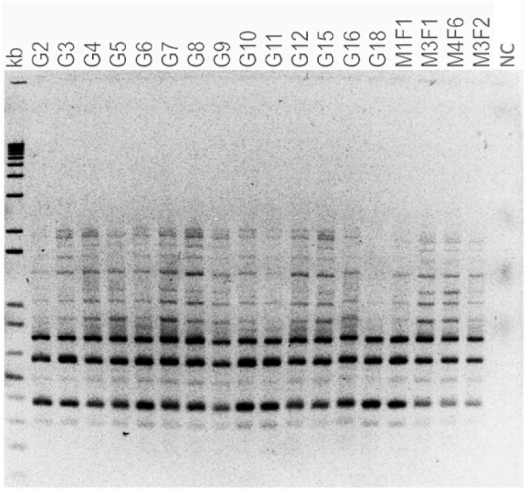

Fig. 2. Agarose gel electrophoresis of ISSR fragments of 18 *Annona squamosa* accessions. Kb, molecular weight marker, NC, negative control.

3. Results

3.1. Assessment of ISSR bands

The nine selected ISSR primers were very efficient in amplifying DNA fragments across the 19 *A. squamosa* genotypes (Fig. 2), with an average of 14.11 bands per primer and band sizes varying between 250 and 4000 bp. Of the 127 amplified bands that could be scored, 101 (79.52%) were polymorphic. The highest number of polymorphic bands ($n = 18$) was obtained with primer UBC812 and the lowest ($n = 5$) with primer UBC890.

3.2. Genetic diversity within- and between-populations

As shown in Table 3, the highest percentage of polymorphic loci was observed within the Canto do Buriti population (73.23%), whereas the lowest was within the population from Bom Jesus (39.37%). The overall proportion of polymorphic loci in the combined populations was 79.52%.

Nei's index of within-population genetic diversity ranged from 0.1750 to 0.2574 with an average of 0.2223, while Shannon's information index varied from 0.2506 to 0.3861 with an average of 0.3266. The between-population variability was 15.85% ($G_{ST} = 0.1585$) and the between-population gene flow was moderate with the occurrence of around three migrants per generation ($N_m = 2.6539$).

3.3. Phylogenetic analysis

Pairwise comparisons based on the binary matrix revealed that the Jaccard coefficients varied from 0.17 to 0.54 with an average value of 0.38. The accessions presenting the least genetic distant were G4 from Bom Jesus and G8 from Canto do Buriti, while the most genetically distant were M1F1 and G11 both from Canto do Buriti. The UPGMA clustering method generated a rooted tree that separated the accessions into three groups and the reliability of the grouping was verified by the high AU percentages (Fig. 3). Group 1 comprised exclusively accession M1F1 from Canto do Buriti, while groups 2 and 3 contained accessions from all three populations. Group 2 was formed only of accessions derived by vegetative propagation, whereas group 3 included accessions that, with the single exception of G2, were derived entirely from seed.

The overall average genetic distance between all pairs of accessions was 0.23, while the average distances between pairs within each group were 0.18 (group 2) and 0.26 (group 3). The most distant pairs of accessions were M1F2/M4F4, M4F6/M4F3 and M3F2/M4F7 (0.27 each pair), whereas the pair most closely related was G4/G8 (0.17). The similarity relationships within the groups are illustrated in the heat map presented in Fig. 4.

3.4. Population structure

Bayesian clustering analysis indicated that the optimal number of clusters of *A. squamosa* accessions was two ($K = 2$; Fig. 5) and specified the proportion of the genome of an individual that likely originated from each inferred population (Fig. 6). According to this analysis, there was no association between the two distinct genetic groups and their geographical origins. Of the 19 accessions of A. squamosa, 14 (73.7%) shared membership coefficients in the two clusters with different degrees of admixture. The clone G2, for example had approximately 0.40 membership of cluster 1 (shown red in Figs. 6) and 0.60 of cluster II (shown in green). Only 5 (26.3%) of the accessions, namely G5, G6, G7, M3F2 and M1F2, showed uniformity in their genetic structure and belonged essentially to one cluster or the other. Interestingly, cluster I (predominantly red in Fig. 6) encompassed all accessions (with the exception of G2) that had been propagated by cuttings, namely G5, G4, G10, G16, G6, G11, G7 and G8. On the other hand, cluster II (predominantly green) incorporated the 10 accessions propagated by seed (M3F2, M3F3, M3F1, M1F1, M1F2, M4F1, M4F3, M4F4, M4F7 and M4F6) with about 90% compatibility within the group.

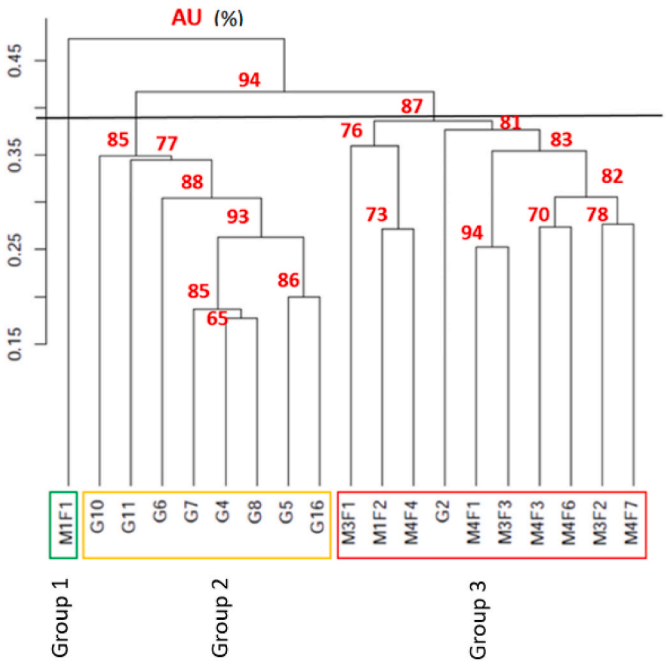

Fig. 3. Dendrogram constructed using the unweighted pair-group method with arithmetic averages (UPGMA) based on Jaccard distances calculated for the 19 *Annona squamosa* accessions studied. The values in red represent the approximately unbiased (AU) percentages of reliability of the clustering method (AU ≥ 90% for ideal reliability). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

In the triangle plot shown in Fig. 7, each accession is represented by a point, the color of which refers to one of the three populations (Timon, Canto do Buriti or Bom Jesus), and the vertices correspond to clusters. The nearer an accession is to a vertex, the greater the genetic proximity to that particular cluster. However, when points of one population are found close to the points of another population, the accessions are considered migrants. In the present case, the overlapping of populations in clusters I and II confirm the presence of migrants or their descendants within the populations.

4. Discussion

The high overall polymorphism reported herein (79.52%) demonstrates the resolving power and efficiency of ISSR primers in detecting differences between *A. squamosa* genotypes and is comparable with levels previously reported for members of the genus *Annona*. For example, Nagori et al. [37] reported that the combined ISSR and RAPD

Table 3
Genetic diversity within- and between-populations of *Annona squamosa* grown in northeastern Brazil as determined by analysis of 127 amplified ISSR loci.

Population	NPL[a]	PPL[b]	H^c	I^d	Nm^e	G_{ST}^f
Timon	74	58.27%	0.2343	0.3432	–	–
Bom Jesus	50	39.37%	0.1750	0.2506	–	–
Canto do Buriti	93	73.23%	0.2574	0.3861	–	–
Total	101	79.52%	0.2223	0.3266	2.6539	0.1585

[a] Number of polymorphic loci.
[b] Percentage of polymorphic loci.
[c] Nei's genetic diversity index.
[d] Shannon's information index.
[e] Number of migrants per generation.
[f] Genetic differentiation coefficient.

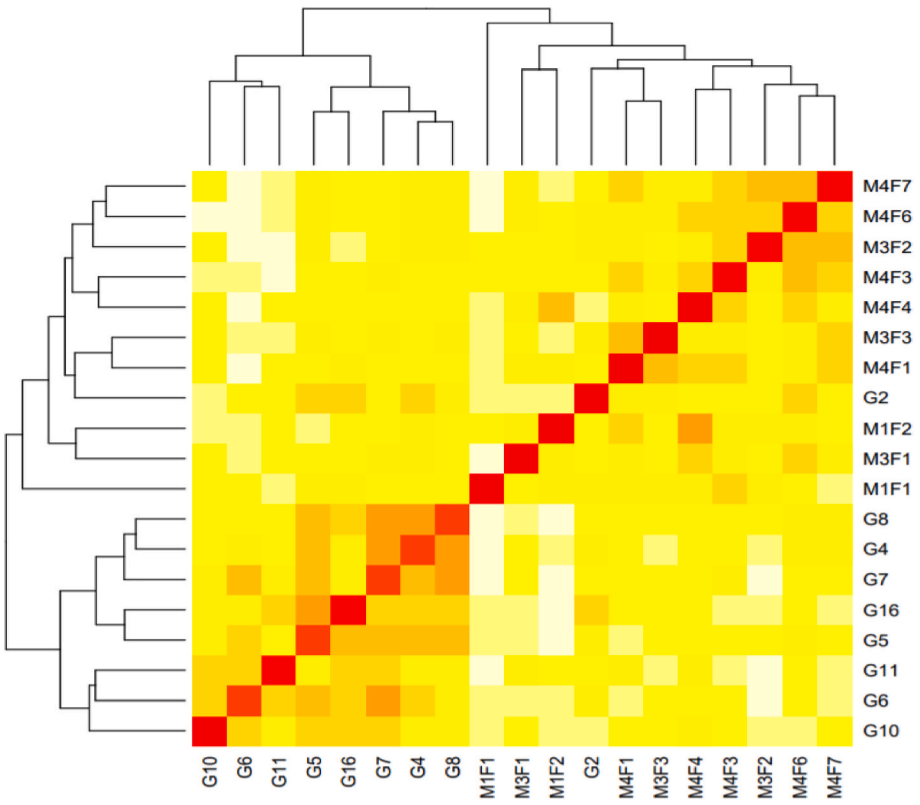

Fig. 4. Heat map representing the pairwise genetic distances between the 19 *Annona squamosa* accessions studied. Values were calculated with 1000 permutations. Genetic similarity is low to moderate (white to yellow) or high (orange to red). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

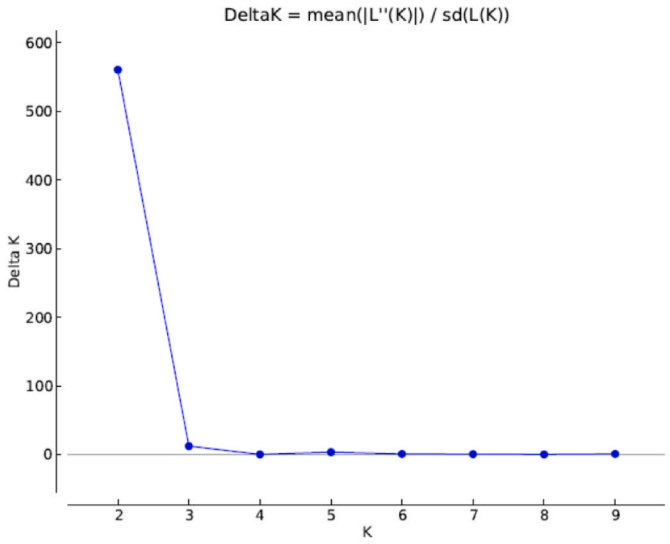

Fig. 5. Estimation of the most probable number of groups (*K*) represented by the highest delta-*K* value determined with the aid of STRUCTURE software.

data from 21 *A. squamosa* populations from India revealed an overall polymorphism of 73.91%, while Gwinner et al. [38] employed ISSR primers to detect an overall polymorphism of 73.72% among 55 accessions of *A. crassiflora* Mart.

ISSR-based analysis of genetic diversity revealed moderate genetic variation within and between populations of *A. squamosa*. The values of the Nei diversity index *H*, which range from 0 to 1 with higher values indicating greater genetic diversity, reveal that the population from Bom Jesus ($H = 0.1750$) exhibited less diversity in comparison with those from Canto do Buriti ($H = 0.2574$) and Timon ($H = 0.2343$). Values of *H* in the range 0.112–0.300 have been reported for populations of *A. crassiflora* analyzed using amplified fragment length polymorphism (AFLP) primers [39], while Cota et al. [40] obtained values between 0.210 and 0.300 for populations of the same species analyzed by RAPD.

The high diversity within the population of *A. squamosa* from Canto do Buriti observed in the present study can be explained by the ten accessions that were propagated via seed, which allowed the random exchange of alleles. The population from Bom Jesus showed the least diversity probably because it consists entirely of clones G4, G10, and G16 that may share common ancestry, while the intermediate genetic variability within the Timon population may be explained by the presence of some half siblings, i.e. the sexually propagated accessions (M3F1, M3F2, and M3F3) originating from G5 that was propagated via cuttings.

According to Shannon's diversity index *I*, which ranges from 0 to 1 with higher values indicating greater diversity [40], the *A. squamosa* population from Bom Jesus presented low diversity ($I = 0.2506$), while those from Timon and Canto do Buriti exhibited moderate diversities ($I = 0.3432$ and 0.3861, respectively). Values reported previously for *A. crassiflora* populations varied from 0.119 to 0.323 [39] and from 0.310 to 0.440 [40].

The values of GST also range from 0 to 1, with 0 representing no differences in allele frequencies between two populations and 1 indicating that the populations interchange alleles. The G_{ST} value of 0.1585 obtained for the three populations of *A. squamosa* is within the range 0.05–0.15 that is considered as high genetic differentiation [41]. GST is related directly to gene flow and mutation rate but, in many cases, gene flow is assumed to be high in relation to mutation rate such that the coefficient may be used to assess migration rates [42]. The occurrence of gene flow between-populations was verified by the N_m value (2.6539), which is an estimate of the number of migrants successfully entering a

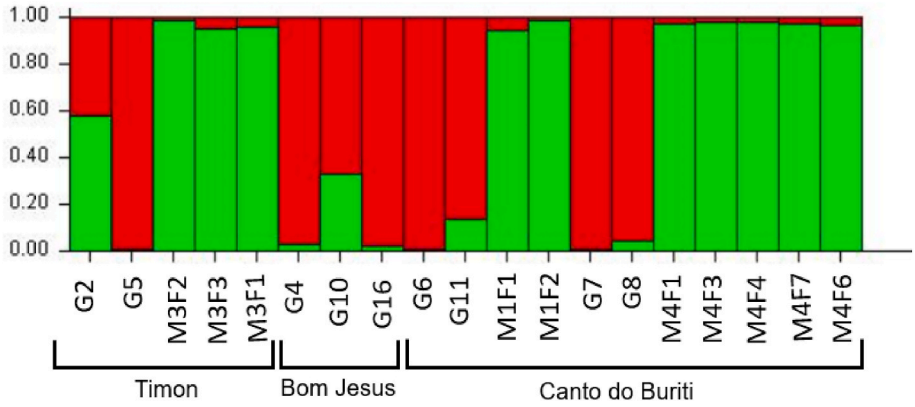

Fig. 6. Bayesian analysis, performed with the aid of STRUCTURE software, of 19 *Annona squamosa* accessions from Timon, Bom Jesus and Canto do Buriti showing separation into two clusters. Each color represents one group and the length of the colored segment represents the estimated membership proportion of each accession to the designated group. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

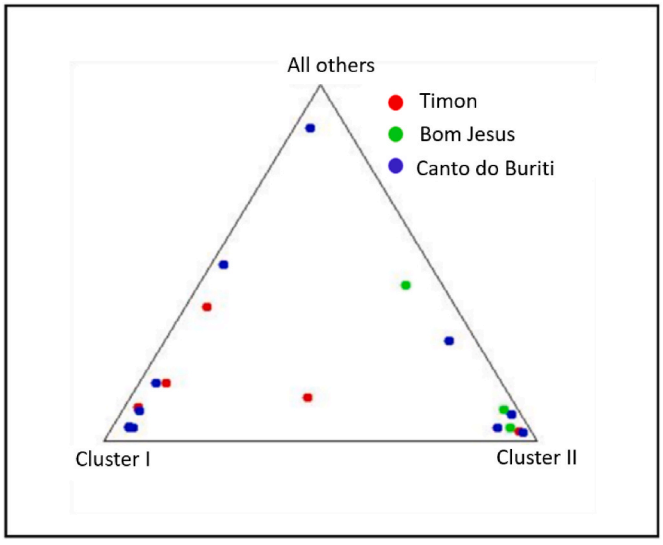

Fig. 7. Triangle plot constructed with the aid of STRUCTURE software using the functions USEPOPINFO, MIGRPRIOR (0.05) and GENSBACK (3). Each accession is represented as a colored dot and each color corresponds to one of the three populations of *Annona squamosa*. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

population at each generation. Considering the long distances between the study populations, gene flow among the accessions probably resulted from anthropic activities rather than from natural dispersion by Coleopterans [43].

All the indices presented herein indicated that genetic variabilities within-populations of *A. squamosa* were generally greater than those between-populations, thereby confirming findings reported previously for this species [44], for *A. crassiflora* [39] and for *A. senegalensis* Pers [45]. This finding suggests that efforts to augment the gene pool of *A. squamosa* in the germplasm bank at Embrapa Meio-Norte should be concentrated on each separate population to prevent genetic redundancy within the collection [46].

The UPGMA dendrogram separated the genotypes into three groups (Fig. 3) according to genetic relatedness, whereby group 2 consisted exclusively of accessions propagated vegetatively while group 3 comprised primarily the seed-propagated accessions along with clone G2. The unique member of group1 was accession M1F1, which was genetically diversified in relation to all other accessions. The UPGMA groups were comparable with the clusters separated according to ancestry using the Bayesian method (Fig. 6), while the heat map (Fig. 4) permitted similar inferences regarding the genetic diversity of the studied genotypes. These results demonstrate that the half-sibling accessions G5 (M3F1, M3F2 and M3F3), G6/G11 (M1F1, M1F2) and G7/G8 (M4F1, M4F3, M4F4, M4F6 and M4F7) present distinct genetic structures.

The separation of *A. squamosa* genotypes into genetically distinct clusters was associated with the type of reproduction while geographical provenance seemed not to influence the distribution. Vegetative reproduction produces individuals that are genetically identical, or nearly identical, to the parent plant, while progenies generated via sexual reproduction are genetically different from their parent plants because of genetic recombination. In this context, *A. squamosa* is predominantly allogamous [47]. This phenomenon is due to a condition known as protogynous dichogamy in which the sexual organs mature at different times, thereby preventing self-pollination [48,49]. Gene flow by cross-pollination produces seeds with high genetic variability as exemplified by the seed-propagated accessions of *Annona* species [10,44].

The low genetic diversity between-populations of *A. squamosa* could have been caused by human-assisted gene flow and by parental relatedness between the accessions. It is most likely that the three populations did not originate at their sampling sites but at one single location. The geographical dissociation of the three populations likely resulted from the interaction between indigenous communities in the area, similar to the situation described by Salazar et al. [44] for Mayan domestic gardens in the Yucatan Peninsula.

5. Conclusions

While ISSR marker-based DNA fingerprinting of the 19 studied accessions revealed a degree of genetic diversity between populations, the variability detected within each population was somewhat higher. The low variability between the populations can be explained by their historical connections and shared genetic material in combination with the allogamous reproduction of the species. Bayesian and UPGMA analysis separated the accessions on the basis of their form of propagation.

The potential loss of species and continuing habitat fragmentation lend urgency to the need to conserve germplasm of *A. squamosa.* One strategy by which this might be accomplished is through the establishment of AGBs. The results recorded herein will contribute to a better understanding of genetic variation and population structure of the accessions of *A. squamosa* in the AGB of Embrapa Meio-Norte AGB and provide crucial information for future conservation and genetic improvement programs of the species

CRediT authorship contribution statement

Gisele H. Sá: conception and design, acquisition of data, Formal analysis, drafting the article, critical review of important intellectual content. **Fatianne C.D. Lima:** acquisition of data, Formal analysis, critical review of important intellectual content. **João P.G. Viana:** critical review of important intellectual content. **Ângela C.A. Lopes:** critical review of important intellectual content. **Leonardo C.B. Carvalho:** conception and design, acquisition of data, Formal analysis, drafting the article, critical review of important intellectual content. **Sérgio E.S. Valente:** Formal analysis, drafting the article, critical review of important intellectual content. **Paulo S.C. Lima:** conception and design, Formal analysis, drafting the article, critical review of important intellectual content. All authors approved the final version of the manuscript.

Declaration of competing interest

The authors declare no competing interest.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by Embrapa (grant number 11.15.02.002.06.06.013) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

[1] Companhia Nacional de Abastecimento – CONAB, Boletim Hortigranjeiro Janeiro 2022, 2022. https://www.conab.gov.br/info-agro/hortigranjeiros-prohort/boletim-hortigranjeiro. (Accessed 24 March 2022).
[2] J.C. Fachinello, J.C. Nachitigal, E. Kersten, Poda de plantas frutíferas, in: Fruticultura: Fundamentos e Práticas, Embrapa Clima Temperado, 2008, pp. 93–102. Pelotas.
[3] E.H.A. Andrade, M.G.B. Zoghbi, J.G.S. Maia, H. Fabricius, F. Marx, Chemical characterization of the fruit of *Annona squamosa* L. occurring in the Amazon, J. Food Compos. Anal. 14 (2001) 227–232, https://doi.org/10.1006/jfca.2000.0968.
[4] M.M.B. Almeida, P.H.M.D. Sousa, M.L. Fonseca, C.E.C. Magalhães, M.D.F.G. Lopes, T.L.G.D. Lemos, Evaluation of macro and micro-mineral content in tropical fruits cultivated in the northeast of Brazil, Food Sci. Technol. 29 (2009) 581–586, https://doi.org/10.1590/S0101-20612009000300020.
[5] M.A.Y. Abdualrahman, H. Ma, C. Zhou, A.E.A. Yagoub, A.O. Ali, H.E. Tahir, A. Wali, Postharvest physicochemical properties of the pulp and seed oil from *Annona squamosa* L. (Gishta) fruit grown in Darfur region, Sudan, Arab. J. Chem. 12 (2019) 4514–4521, https://doi.org/10.1016/j.arabjc.2016.07.008.
[6] R. Senthil, R. Silambarasan, *Annona*: a new biodiesel for diesel engine: a comparative experimental investigation, J. Energy Inst. 88 (2015) 459–469, https://doi.org/10.1016/j.joei.2014.09.011.
[7] H.A. Salman, S.M. Ramasamy, B.S. Mahmood, Detection of caffeic and chlorogenic acids from methanolic extract of *Annona squamosa* bark by LC-ESI-MS/MS, J. Intercult. Ethnopharmacol. 7 (2018) 76–81, https://doi.org/10.5455/jice.20171011073247.
[8] H.X. Li, S.Y. Yang, Y.S. Kim, H.D. Jang, Y.H. Kim, W. Li, Nitro derivatives and other compounds from sugar apple (*Annona squamosa* L.) leaves exhibit soluble epoxide hydrolase inhibitory activity, Med. Chem. Res. 28 (2019) 1939–1944, https://doi.org/10.1007/s00044-019-02425-2.
[9] I.B.G.E. Instituto Brasileiro de Geografia e Estatística –, Censo Agro, 2017. https://censoagro2017.ibge.gov.br/templates/censo_agro/resultadosagro/agricultura.html?localidade=0&tema=76297. (Accessed 24 March 2022).
[10] H. Anuragi, H.L. Dhaduk, S. Kumar, J.J. Dhruve, M.J. Parekh, A.A. Sakure, Molecular diversity of *Annona* species and proximate fruit composition of selected genotypes, 3 Biotech 6 (2016) 204, https://doi.org/10.1007/s13205-016-0520-9.
[11] A.C.Q. Pinto, *Annona* Species, International Centre of Under-Utilized Crops, University of Southampton Press, Southampton, 2005, pp. 71–126.
[12] A.N.C. Borges, A.C.A. Lopes, F.B. Brito, L.F.L. Vasconcelos, P.S.C. Lima, Genetic diversity in a cajuí (*Anacardium* spp.) germplasm bank as determined by ISSR markers, Genet. Mol. Res. 17 (2018) 1–14, https://doi.org/10.4238/gmr18212.
[13] R.L.F. Gomes, M.F. Costa, A. Alves-Pereira, M.M. Bajay, J.P.G. Viana, S.E.D. S. Valente, J.B. Pinheiro, A lima bean core collection based on molecular markers, Sci. Agric. 77 (2020) 1–8, https://doi.org/10.1590/1678-992x-2018-0140.
[14] R.R.A. Giachino, Investigation of the genetic variation of anise (*Pimpinella anisum* L.) using RAPD and ISSR markers, Genet. Resour. Crop Evol. 67 (2020) 763–780, https://doi.org/10.1007/s10722-019-00861-y.
[15] D.W. Nogueira, D.G. Nogueira, W.R. Maluf, G.M. Maciel, A.D.R. Figueira, C.B. D. Menezes, Seleção assistida com uso de marcador molecular para resistência a potyvírus em pimentão, Pesqui. Agropecu. Bras. 47 (2012) 955–963, https://doi.org/10.1590/S0100-204X2012000700012.
[16] J.S. Penha, A.C.A. Lopes, R.L.F. Gomes, J.B. Pinheiro, J.R. Assunção Filho, E. A. Silvestre, J.P.G. Viana, J. Martinez-Castilho, Estimation of natural outcrossing rate and genetic diversity in Lima bean (*Phaseolus lunatus* L. var. *lunatus*) from Brazil using SSR markers: implications for conservation and breeding, Genet. Resour. Crop Evol. 64 (2017) 1355–1364, https://doi.org/10.1007/s10722-016-0441-9.
[17] S. Adak, S. Datta, S. Bhattacharya, T.K. Ghose, A.L. Majumder, Diversity analysis of selected rice landraces from West Bengal and their linked molecular markers for salinity tolerance, Physiol. Mol. Biol. Plants 26 (2020) 669–682, https://doi.org/10.1007/s12298-020-00772-8.
[18] E. Julio, A. Malpica, J. Cotucheau, S. Bachet, R. Volpatti, C. Decorps, F.D. Borne, RNA-Seq analysis of *Orobanche* resistance in *Nicotiana tabacum*: development of molecular markers for breeding recessive tolerance from 'Wika' tobacco variety, Euphytica 216 (2020) 6, https://doi.org/10.1007/s10681-019-2544-9.
[19] J. Li, L. Demesyeux, M. Brym, A.H. Chambers, Development of species specific molecular markers in *Vanilla* for seedling selection of hybrids, Mol. Biol. Rep. 47 (2020) 1905–1920, https://doi.org/10.1007/s11033-020-05287-9.
[20] N.R. Sousa, D.P. Rodrigues, C.R. Clement, E.O. Nagao, S. Astolfi-Filho, Discriminação de raças primitivas de pupunha (*Bactris gasipaes*) na Amazônia brasileira por meio de marcadores moleculares (RAPDS), Acta Amazonica 31 (2001) 539, https://doi.org/10.1590/1809-43922001314545, 539.
[21] S. Benor, S. Demissew, K. Hammer, F.R. Blattner, Genetic diversity and relationships in *Corchorus olitorius* (Malvaceae sl) inferred from molecular and morphological data, Genet. Resour. Crop Evol. 59 (2012) 1125–1146, https://doi.org/10.1007/s10722-011-9748-8.
[22] Z.A. Desta, D.J. Koning, R. Ortiz, Molecular mapping and identification of qualitative trait loci for domestication traits in the field cress (Lepitidium campestre L.) genome, Heredity 124 (2020) 579–591, https://doi.org/10.1038/s41437-020-0296-x.
[23] M.S. Uma, N. Hegde, S. Hittalmani, Identification of SSR marker associated with rust resistance in cowpea (*Vigna unguiculata* L.) using bulk segregant analysis, Legume Res. 39 (2016) 39–42, https://doi.org/10.18805/lr.v39i1.8861.
[24] C.G. Fajardo, R.A.R. Silva, K.P.T. Chagas, F.A. Vieira, Genetic and phenotypic association of the carnauba palm tree evaluated by inter-simple sequence repeat and biometric traits, Genet. Mol. Res. 17 (2018) 1–9, https://doi.org/10.4238/gmr18018.
[25] K.T. Ravikiran, S.L. Krishnamurthy, A.S. Warraich, P.C. Sharma, Diversity and haplotypes of rice genotypes for seedling stage salinity tolerance analyzed through morpho-physiological and SSR markers, Field Crop. Res. 220 (2018) 10–18, https://doi.org/10.1016/j.fcr.2017.04.006.
[26] I. Ahmad, S. Bhagat, T.V.R.S. Sharma, K. Kumar, P. Simachalam, R.C. Srivastava, ISSR and RAPD marker based DNA fingerprinting and diversity assessment of *Annona* spp. in South Andamans, Indian J. Hortic. 67 (2010) 147–151.
[27] S.G. Bharad, P.L. Kulwal, S.A. Bagal, Genetic diversity study in *A. squamosa* by morphological, biochemical and RAPD markers, Acta Hortic. 83 (2009) 615–623, https://doi.org/10.17660/FActaHortic.2009.839.84.
[28] J.F.R. Guimarães, S. Nietsche, M.R. Costa, G.B.R. Moreira, M.C.T. Pereira, W. Vendrame, Genetic diversity in sugar apple (*Annona squamosa* L.) by using RAPD markers, Rev. Ceres 60 (2013) 428–431, https://doi.org/10.1590/S0034-737X2013000300017.
[29] R Core Team, R, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, 2018.
[30] R. Suzuki, Y. Terada, H. Shimodaira, Pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling. R Package Version 2.2-0, 2019. https://cran.r-project.org/web/packages/pvclust/pvclust.pdf. (Accessed 20 March 2020).
[31] F.C. Yeh, R.C. Yang, C. Boyle, POPGENE Version 1.31, University of Alberta and Center for International Research, 1999.
[32] M. Nei, Analysis of gene diversity in subdivided populations, Proc. Natl. Acad. Sci. USA 70 (1973) 3321–3323, https://doi.org/10.1073/pnas.70.12.3321.
[33] R.C. Lewontin, The apportionment of human diversity, in: T. Dobzhansky, M. K. Hecht, W.C. Steere (Eds.), Evolutionary Biology, Springer, New York, 1972, pp. 381–398, https://doi.org/10.1007/978-1-4684-9063-3_14.
[34] J.K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using multilocus genotype data, Genetics 155 (2000) 945–959, https://doi.org/10.1093/genetics/155.2.945.
[35] G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, Mol. Ecol. 14 (2005) 2611–2620, https://doi.org/10.1111/j.1365-294X.2005.02553.x.
[36] D.A. Earl, B.M. Von Holdt, Structure HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method, Conserv. Genet. Resour. 4 (2012) 359–361, https://doi.org/10.1007/s12686-011-9548-7.
[37] R. Nagori, P. Sharma, S.D. Purohit, Molecular characterization of diversity in *Annona squamosa* L.: a high value tropical medicinal plant, Int. J. Phytocosmet. Nat. Ingred. 5 (2018) 1–13, https://doi.org/10.15171/ijpni.2018.04.
[38] R. Gwinner, T.A. Setotaw, F.A. Rodrigues, D.V. França, F.A. da Silveira, L.A. Pio, M. Pasqual, Population structure of *Annona crassiflora*: an endemic plant species of the Brazilian Cerrado, Genet. Mol. Res. 15 (2016) 1–13.
[39] A.P.M. Egydio-Brandão, C.M. Furlan, D.Y.A.C. Santos, Genetic diversity and structure of populations of *Annona crassiflora* Mart. of Brazilian savanna and its

association with chemical variability, Chem. Biodivers. 13 (2016) 990–997, https://doi.org/10.1002/cbdv.201500320.
[40] L.G. Cota, F.A. Vieira, A.F. Melo Júnior, M.M. Brandão, K.N.O. Santana, M. L. Guedes, D.A. Oliveira, Genetic diversity of *Annona crassiflora* (annonaceae) in northern minas gerais state, Genet. Mol. Res. 10 (2011) 2172–2180, https://doi.org/10.4238/vol10-3gmr1188.
[41] S. Wright, Evolution and the genetics of populations, in: Variability within and Among Natural Populations, vol. 4, University of Chicago Press, Chicago, 1978.
[42] M. Whitlock, G'ST and D do not replace FST, Mol. Ecol. 20 (2011) 1083–1091, https://doi.org/10.1111/j.1365-294x.2010.04996.x.
[43] A. Endres, A.J. Creao-Duarte, M.I.M. Hernandez, Diversidade de Scarabaeidae s. str.(Coleoptera) da Reserva Biológica Guaribas, Mamanguape, Paraíba, Brasil, Brasil: uma comparação entre Mata Atlântica e Tabuleiro Nordestino, Braz. J. Entomol. 51 (2007) 67–71, https://doi.org/10.1590/S0085-56262007000100012.
[44] C. Salazar, C.F. Vargas-Mendoza, J.S. Flores, Estructura y diversidad genética de *Annona squamosa* en huertos familiares mayas de la península de Yucatán, Rev. Mex. Biodivers. 81 (2010) 759–770.
[45] K. Kwapata, W.F. Mwase, J.M. Bokosi, M.B. Kwapata, P. Munyenyembe, Genetic diversity of *Annona senegalensis* Pers. populations as revealed by simple sequence repeats (SSRs), Afr. J. Biotechnol. 6 (2007) 1239–1247.
[46] H.E. Fadoul, M.A. El Siddig, A.A. El Hussein, Assessment of genetic diversity among Sudanese wheat cultivars using RAPD markers, Int. J. Curr. Sci. 6 (2013) 51–57.
[47] L.H.P. Kiill, J.G. Costa, Biologia floral e sistema de reprodução de *Annona squamosa L.* (Annonaceae) na região de Petrolina-PE, Ciência Rural. 33 (2003) 851–856, https://doi.org/10.1590/S0103-84782003000500009.
[48] Z.M.M. Modesto, N.J.B. Siqueira, Curricolo de estudos de biologia. Botânica, Pedagógica e Universitária, São Paulo, 1981.
[49] M.C.T. Pereira, S. Nietsche, M.R. Costa, J.H. Crane, C.D.A. Corsato, E.H. Mizobutsi, Anonáceas: pinha, atemoia e graviola, Inf. Agropecu. (Belo Horiz.) 32 (2011) 26–34.