

Genetic structure of *Hancornia speciosa* Gomes (Apocynaceae) in the cerrado and atlantic forest domains reveal two botanical varieties

Jackeline Terto

Universidade Federal Rural de Pernambuco

Ana Kelly Maia

Universidade Federal Rural de Pernambuco

Lázaro José Chaves

Universidade Federal de Goiás

Josué Francisco Silva-Júnior

Embrapa Tabuleiros Costeiros

Elizabeth Ann Veasey

Universidade de São Paulo, ESALQ

Edson Ferreira Silva

Universidade Federal Rural de Pernambuco

Cícero Almeida (✉ cicero@arapiraca.ufal.br)

Universidade Federal de Alagoas

Research Article

Keywords: mangabeira, phylogeography, phylogeny, taxonomy, genetic structure, nSSR, cpDNA, genetic conservation

Posted Date: June 10th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1719926/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Hancornia speciosa Gomes is an important species from Brazil, distributed in the *Cerrado*, *Caatinga* and Atlantic Rain Forest domains. Six botanical varieties are described for *H. speciosa*, of which are distributed in different regions, however, currently, only two varieties are recognized (*pubescens* and *speciosa*). Additionally the phylogeographic and phylogenetic studies have not been reported for *H. speciosa*, which only phylogeographic study has been related for populations from the *Cerrado* domain. Here we aimed to assesses the phylogeographical structure of *H. speciosa* and build the phylogenetic tree using *trnH-psbA* sequences from chloroplast genome and six microsatellites loci of nuclear genome. The results for chloroplast sequence showed 11 haplotypes and phylogenetic tree revealed two main clades, which the first clade containing the botanical varieties *H. speciosa* var *gardneri*, *H. speciosa* var *pubescens* and *H. speciosa* var *cuyabensis* and second clade containing the botanical variety *H. speciosa* var *speciosa*. The microsatellite analysis revealed on group containing botanical variety *H. speciosa* var *speciosa*, while another group consisted of other botanical varieties. The phylogeographic structure revealed five groups: two corresponded to variety *H. speciosa* var *speciosa* (individuals from Caatinga and Atlantic Rain Forest domains) and other three groups contain the botanical varieties *H. speciosa* var *gardneri*, *H. speciosa* var *pubescens* and *H. speciosa* var *cuyabensis* (individuals from Cerrado domain). The genetic structure and phylogenetic data support the two recognized botanical varieties of *H. speciosa*.

Introduction

The climatic oscillation has the main driver of the evolution and diversity, which contributes for biological dispersion, genetic diversity and speciation (Hewitt 2004; Chen et al. 2008; Carnaval et al. 2009; Thomé et al. 2010). In the Brazilian ecosystems, six biomes have been described, including the Atlantic Rain Forest domain, *Caatinga* domain (seasonally dry tropical forest biome) and *Cerrado* domain (Neotropical savanna). These biomes has differences, which the Atlantic Rain Forest domain is humid, with high levels of precipitation and great biodiversity and endemism (Morellato and Haddad, 2000), and there are ecoregions such as the Atlantic Coast *Restingas* (Assis et al. 2011), which the vegetation is composed of elements from the Atlantic Rain Forest and the *Cerrado*, *Caatinga* and Amazonia (Sacramento et al. 2007; Castro et al. 2012). The *Cerrado* is a savanna vegetation, found in dystrophic soils (latosols) and a climate characterized by dry season, the vegetation is composed by high number of endemic plants, which is an important biodiversity hotspot (Myers et al. 2000). The Brazilian *Caatinga* domain is characterized by diverse types of vegetation and found geomorphologically different areas, of which can act as barriers to gene flow (Moro et al. 2015).

The diversity of biomes has resulted in great diversity of species and high intraspecific diversity. Remarkable, some species occurs in the Atlantic Rain forest, Cerrado and Caatinga, of which *Hancornia speciosa* Gomes is one of the most notable. *Hancornia speciosa* is known as *mangaba* or *mangabeira* in the vernacular language, belongs to the family Apocynaceae and is an economically important fruit species, which the fruit is consumed in nature, sweet, ice-cream and juice (Silva Júnior and Léo, 2006).

Six botanical varieties are described: (1) *H. speciosa* var. *speciosa* Gomes, (2) *H. speciosa* var. *maximiliani* A. DC., (3) *H. speciosa* var. *cuyabensis* Malme, (4) *H. speciosa* var. *lundii* A. DC., (5) *H. speciosa* var. *gardneri* (A. DC.) Muell. Arg. and (6) *H. speciosa* var. *pubescens* (Nees. and Martius) Muell. Arg. (Monachino, 1945). The identification of these botanical varieties is based in the morphological characters; however, the origin of the genetic relationships and the putative hybrids of the botanical varieties is not clear (Collevatti et al. 2016; Flores et al. 2017). The botanical varieties are dispersed in the Atlantic Rain forest, *Cerrado* and *Caatinga* domains, of which the differences among biomes is suitable for phylogeographic and phylogenetic studies, which may be useful to understand of the genetic diversification of *H. speciosa*. However, currently, only two varieties are recognized: *pubescens* and *speciosa*, according to Koch et al. (2015).

Genetic studies have been described for *H. speciosa* using morphological (Ganga et al. 2010; Freitas et al. 2012; Nascimento et al. 2014; isoenzymatic (Martins et al. 2012) and molecular markers (Moura et al. 2011; Silva et al. 2011; Costa et al. 2011; Silva et al. 2012; Amorim et al. 2015; Collevatti et al. 2016; Costa et al. 2017). Recent genetic population studies were performed in 28 populations from the *Cerrado* domain, using seven microsatellite markers and showed isolation by distance and partially supported the classification of the botanical varieties from Monachino (Collevatti et al. 2018). However, the complete phylogenetic and phylogeographic studies is need in order to obtain robust inferences and understanding of the relationships among putative botanical varieties as well the spatial distributions of *H. speciosa*. Here we investigated the questions: (1) Are there phylogenetic supports for the botanical varieties of *H. speciosa*? (2) Do genetic diversity show phylogeographic patterns associate with biomes? To answer these questions, the present study analyzed the phylogeographic structure of *H. speciosa* using sequences from the chloroplast genome and six nuclear microsatellite loci.

Material And Methods

Sampling and DNA extraction

Leaves were sampled in eight localities across to the three biomes. Three localities in the *Cerrado* (Goiás, Minas Gerais and Maranhão), two in the *Caatinga* (Bahia and Ceará) and three from Atlantic forest composed by localities Sergipe, Pernambuco and Paraíba. The samples correspond to a large area of distribution of the species, allowing to obtain robust inferences about the evolution of the species (Figure S1, in blue represent the distribution of specie and red represent the individuals sampled) Additionally, nine individuals identified as the botanical varieties *gardneri*, *pubescens* and *cuyabensis* (three individuals for each botanical variety) were sampled of the Active Germplasm Bank of *Mangaba* of the Universidade Federal de Goiás.. DNA extraction was performed according to the protocol in Doyle and Doyle (1990), and the quantity and quality of DNA were analyzed using 1% agarose gel electrophoresis.

cpDNA: amplification and sequencing

The amplification was performed for three regions of cpDNA (*trnH-psbA* intergenic spacer and *matK* and *rbcL* genes) using primers described by Scarcelli et al. (2011). The reactions were performed in a total of

50 µL, containing 5 µL reaction buffer, 1.5 mM MgCl₂, 0.2 mM dNTP, 1.25 U *Taq* DNA polymerase, primers (0.5 µM each) and 200 ng DNA. Amplifications were performed with an initial denaturation at 94 °C for 3 min, followed by 40 cycles at 94 °C for 30 s and annealing at 48 °C - 60 °C for 30 s, extension at 72 °C for 1 min and a final extension at 72 °C for 10 min. PCR reactions were performed in a thermal cycler (BioCycler) and PCR products were subjected to electrophoresis in 1% agarose gel to confirm amplification. PCR products were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems®) and analyzed by electrophoresis in a 3500 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA).

cpDNA: haplotype analysis

Sequences were edited using the Geneious R9 software and aligned with MAFFT (Kato et al. 2002) implemented in Geneious. The alignment was performed using the default settings and optimized manually when necessary. Haplotype diversity (*h*), nucleotide diversity (π) (Rogers and Harpending, 1992) and mismatch distribution (Rogers and Harpending, 1992) were calculated using the DnaSP 5.10.01 software (Librado and Rozas, 2009). The mismatch distribution unimodal pattern is expected for samples under recent demographic expansion and a multimodal pattern for populations at demographic equilibrium.

The best-fit substitution model was obtained using the MEGA7 (Kumar et al. 2016) to help guide the selection of a model for the molecular evolution of haplotypes, and the GTR+I+G model was used for subsequent phylogenetic inferences using Bayesian analysis. This analysis was performed using Beast v1.8.0 (Drummond and Rambaut, 2007), and the subsequent distribution was approximated using the Markov Chain Monte Carlo method for 10 million generations. Convergence of parameters was assessed using the Tracer 1.5 software (Rambaut and Drummond, 2009). Genealogic relationships between haplotypes were estimated using the maximum parsimony statistical method (Templeton et al. 1992), implemented in the R package 'haplotypes'. The genetic relationships among species were also investigated through a principal component analysis (PCA) using the function *glPca* in R package *ade4* (Jombart and Ahmed, 2011) and graphical obtained using the function *plot3D* e "*rgl*."

cpDNA: population structure

Spatial analysis of molecular variance was performed using the Spatial Analysis of Molecular Variation (SAMOVA) software (Dupanloup et al. 2002). This approach uses the simulations to determine the groups of populations (*k*) that are geographically homogeneous and that maximize the differences between groups. The SAMOVA analyses were performed using 1000 permutations for *k* = {2, ..., 8} groups.

Nuclear SSR (nSSR) amplification and genotyping

A set of six SSR markers developed by Rodrigues et al. (2015) was used for polymorphism analysis (Table S2). The reactions were performed in a total volume of 50 µL containing 5 µL reaction buffer, 1.5 mM MgCl₂, 0.2 mM dNTP, 1.25 U *Taq* DNA polymerase, primers (0.5 µM each) labeled with 5¢-FAM and

200 ng DNA. Amplifications were performed with an initial denaturation at 95 °C for 10 min, followed by 15 cycles at 94 °C for 30 s and annealing at 55 °C–60 °C for 1 min; 20 additional at 89 °C for 1 min and annealing at 55 °C–60 °C for 1 min, extension at 72 °C for 1 min and a final extension at 72 °C for 1 h. PCR reactions were performed in a thermal cycler (BioCycler), and PCR products were subjected to capillary electrophoresis in a 3500 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA), using the GeneScan™ 600 Liz® (Applied Biosystems) as size standard.

nSSR analysis

The allele number (N_a), expected heterozygosity (H_E), observed heterozygosity (H_O) and deviation from Hardy–Weinberg equilibrium (HWE) were obtained using the R package Pegas (Paradis et al. 2017) and the R package PopGenReport 3.0 (Gruber and Adamack, 2017). The minimum-spanning networks were produced in the R package poppr 2.6.1 (Kamvar et al. 2014), using the genetic distance measure of Bruvo et al. (2004) to examine the genetic relationships among multilocus microsatellite genotypes, and PCA was produced in the R package adegenet (Jombart and Ahmed, 2011). The population structure was estimated using the R package Geneland 2.0.9 (Estoup et al. 2007).

Results

Haplotype networks and phylogenetic analyses

Of the three chloroplast regions sequenced, only the *trnH-psbA* showed polymorphism in the 151 samples analyzed, resulting in a total of 427 bp length. The analysis of the sequences revealed 11 haplotypes (Table S1), haplotype diversity (h) of 0.683 ± 0.022 and the mean nucleotides diversity (π) of 0.00458 ± 0.004 . The haplotypes were named using H letter and numbers from 1 to 11, of which H4, H5, H6, H7, H10 and H11 were private to variety *H. speciosa* var *speciosa*, H1 to variety *H. speciosa* var *pubescens*, H2, H3 and H9 to variety *H. speciosa* var *gardneri* and H8 to variety *H. speciosa* var *cuyabensis* (Figures 1 and 2). Among the haplotypes private to variety *H. speciosa* var *speciosa*, the H4, H5, H6, and H7 were exclusive from the *Restinga*, while haplotypes H10 and H11 were from the Cerrado domain. Markedly, haplotype H3 (*H. speciosa* var *cuyabensis*) was found dispersed in the Caatinga domain (Figure 2A).

The relationship among haplotypes using PCA shows clear support for botanical varieties delimitation in *H. speciosa* (Figure 1A), where variety *H. speciosa* var *speciosa* was displayed in one group and the varieties *H. speciosa* var *gardneri*, *H. speciosa* var *cuyabensis* and *H. speciosa* var *pubescens* in another group. Among the varieties of the second group, variety *H. speciosa* var *pubescens* was separated from varieties *H. speciosa* var *gardneri* and *H. speciosa* var *cuyabensis* (Figure 1A). The phylogenetic analyses using Bayesian (Figure 1B), showed two main clades, of which one clade was constituted of varieties *H. speciosa* var *gardneri*, *H. speciosa* var *cuyabensis* and *H. speciosa* var *pubescens* and another clade was formed by variety *speciosa* (Figure 1B). Remarkedly, the variety *H. speciosa* var *gardneri* was found in the first clade close to varieties *H. speciosa* var *cuyabensis* and *H. speciosa* var *pubescens* (Figure 1B). According to the NETWORK results, haplotypes from variety *H. speciosa* var *speciosa* contained four

nucleotide substitutions, while the haplotypes from *H. speciosa* var *gardneri*, *H. speciosa* var *cuyabensis* and *H. speciosa* var *pubescens* showed one to three substitutions (Figure 2B).

nSSR analysis

A total of 176 samples were genotyped with six nSSR loci and the results displayed 138 alleles with 23 alleles on average by *locus*, ranging from 14 to 41 alleles/locus (Table S2). The observed heterozygosity (H_O) ranged from 0.34 to 0.79 and the expected heterozygosity (H_E) ranged from 0.66 to 0.91. The genetic relationships among genotypes using a minimum-spanning network showed that most of the genotypes from *H. speciosa* var *speciosa* (individuals from Sergipe, Pernambuco, Paraíba and Maranhão) were observed in a large cluster, while the genotypes from *H. speciosa* var *gardneri*, *H. speciosa* var *cuyabensis* and *H. speciosa* var *pubescens* (individuals from Bahia, Ceará, Goiás, and Minas Gerais) were observed in another large cluster (Figure 3A).

Within the *H. speciosa* var *speciosa* cluster, there is a separated sub cluster with abundance individuals from Maranhão locality (*Cerrado* domain) and another sub cluster with individuals from Pernambuco, Paraíba and Sergipe (Atlantic Coast Restingas). Another larger cluster showed a mixture of the other varieties; however, a predominance of individuals from Bahia and Ceará (*H. speciosa* var *gardneri*) was found in a sub cluster, while individuals from Goiás and Minas Gerais (*H. speciosa* var *gardneri*, *H. speciosa* var *cuyabensis* and *H. speciosa* var *pubescens*) were observed in another sub cluster (Figure 3A). The relationship among genotypes using the PCA showed clear groups constituted by individuals of *H. speciosa* var *speciosa* (individuals from Sergipe, Maranhão, Pernambuco and Paraíba) and another group formed by individuals from *H. speciosa* var *gardneri*, *H. speciosa* var *cuyabensis* and *H. speciosa* var *pubescens* (individuals from Bahia, Ceará, Goiás, and Minas Gerais) (Figure 3B).

Population genetic structure for cpDNA and nSSR

For cpDNA, the best description of population structure in the SAMOVA analysis was obtained when $k = 5$, thereby enabling the identification of five population groups with $F_{ST} = 0.94$ ($P < 0.001$), comprising 93.9% of the variation between groups (Table 1). Group A contains the Paraíba, Pernambuco and Sergipe localities and group B contains individuals from Maranhão locality. Both groups A and B correspond to variety *H. speciosa* var *speciosa* from Coastal Tablelands Restingas and Cerrado domains, respectively. Group C is characterized by individuals of *H. speciosa* var *gardneri* from Bahia and Ceará and groups D and E contain individuals of *H. speciosa* var *gardneri*, *H. speciosa* var *cuyabensis* and *H. speciosa* var *pubescens* from Goiás and Minas Gerais (Table 1, Figure 1A and Figure S1). The SAMOVA analysis for nSSR among groups determined by cpDNA showed $F_{ST} = 0.47$ ($P < 0.001$), comprising 43.02% of the variation between groups and 52.69% within of groups (Table 1). The mismatch distribution showed populations at demographic equilibrium, of which the pairwise distribution displayed a typical multimodal distribution to a model of sudden expansion (see Figure S2).

Discussion

***H. speciosa* varieties relationships**

Hancornia speciosa is spread in the *Cerrado* (tropical savanna) and Atlantic coast and here we show the first complete phylogenetic and phylogeographic study using sequences from chloroplast DNA and SSR data, including four botanical varieties and individuals from all regions where the species occurs. Our findings show clear genetic differentiation among varieties of *H. speciosa*, of which *H. speciosa* var. *speciosa* shows high differentiation and distinct clade when compared with putative botanical varieties *H. speciosa* var. *cuyabensis*, *H. speciosa* var. *gardneri* and *H. speciosa* var. *pubescens*. This clear differentiation was observed using both chloroplast DNA sequence and SSR data. Among the varieties from the south-west *Cerrado* domain, there was a lower distinction, in which the *H. speciosa* var. *cuyabensis* was displayed in one clade and *H. speciosa* var. *pubescens* in another clade; however, *H. speciosa* var. *gardneri* showed haplotypes close to *H. speciosa* var. *cuyabensis* or *H. speciosa* var. *pubescens*, suggesting that *H. speciosa* var. *gardneri* may be originated by cross-pollination. The hybridization among *H. speciosa* varieties was demonstrated by Collevatti et al. (2016), showing that there is gene flow among botanical varieties.

The high similarity among individuals from the *Cerrado* domain was observed by Collevatti et al. (2018) using SSR data, in which the varieties *H. speciosa* var. *cuyabensis*, *H. speciosa* var. *gardneri* and *H. speciosa* var. *pubescens* generally clustered together. Other studies using chemometric analyses showed that *H. speciosa* var. *speciosa* differs from other varieties (Flores et al. 2017). Our phylogenetic analysis shows correlation with morphological descriptions suggested by Monachino (1945), in which the varieties *H. speciosa* var. *gardneri* and *H. speciosa* var. *pubescens* differ due to the presence of pubescence in the leaves of *H. speciosa* var. *pubescens* (Monachino, 1945; Ganga et al. 2010). Our data support two botanical varieties identified as *H. speciosa* var. *speciosa* from Atlantic coast and *H. speciosa* var. *pubescens* from central Brazil. Two botanical varieties have been recognized by Koch et al. (2015), which belong the *H. speciosa* var. *speciosa* and *H. speciosa* var. *pubescens*. Studies using SSR data (Collevatti et al. (2018) suggested partially support Monachino's classification at least for the four sampled varieties, however, our study we used samples from *Caatinga* and Atlantic forest domains and two differences genetic data (sequence data from chloroplast and SSR data from nuclear genome).

Phylogeographic structure

Phylogeographic structure in this work corroborates with the main phylogeographical patterns described for the *Cerrado* domain, as reported for *Dalbergia miscolobium* Benth (Fabaceae) (Novaes et al. 2013) and for *Qualea grandiflora* Mart (Vochysiaceae) (Buzatti et al. 2018). Remarkably, the phylogeographical patterns reveal a clear differentiation between the *Cerrado* domains and Atlantic Coast *Restinga* and show a clear similarity to the *Cerrado* domain's phylogeographic structure proposed by Collevatti et al. (2018) using SSR markers for *H. speciosa*. However, our data contain individuals from *Cerrado*, from the transition between *Cerrado* and *Caatinga* domains and from Atlantic Coast *Restinga*, including four botanical varieties and complete spatial geographical distribution analyzed by SSR markers and sequences from chloroplast DNA.

The phylogeographic structure showed that *H. speciosa* var *speciosa* occurs in the Coastal Tablelands and Restinga and in transitions areas between the Cerrado and Caatinga domains, while other botanical varieties were exclusive in the Cerrado domain, except for *H. speciosa* var *gardneri* which was found in the Caatinga domain from Bahia and Ceará. No samples of *H. speciosa* var *speciosa* from Central Brazil were found in this study, possibly the *H. speciosa* var *speciosa* from Central Brazil is different to found in the Coastal Tablelands and Restinga. The structure from central Brazil showed that populations clustered together, independent of botanical varieties, and this pattern was also found by Collevatti et al. (2018) for populations from western Cerrado domains. However, we observed a separation between individuals from Eastern and Western Central Cerrado domain. This separation in the Cerrado domain may be explained by the different hypothesis involving climate, soil fertility, past climatic changes and contact with other vegetation from the Atlantic Rain Forest, but our results suggest that some forces have persisted and modeled the diversification of *H. speciosa*.

Genetic conservation

The genetic diversity of *H. speciosa* did not show random distribution in the geographical space and was not separated by botanical varieties. There are four regions characterized by different lineages distributed in the three botanical varieties. The first is located in the Atlantic Coast Restinga consisting of *H. speciosa* var *speciosa*, whereas the second is located between Cerrado and Caatinga domains also consisting of individuals of *H. speciosa* var *speciosa*. The third and fourth gene pools are located in central Brazil, in the Eastern and Western Central Cerrado domains and are characterized by individuals of *H. speciosa* var *pubescens* e *H. speciosa* var *cuyabensis*. The results have implications for conservation, and it would be important to design strategies to conservation of the genetic diversity, mainly in the Atlantic Coast Restinga, where the expansions of urban areas and agriculture have been damaging the natural populations. Whereas two of six varieties reported by Monachino (1945); *gardneri* and *maximiliani* were not found, although, the reported areas of occurrence have been examined.

Declarations

Data archiving statement

Haplotype sequence data for intergenic region have been submitted to GenBank: accession numbers MN249239 - MN249393

Acknowledgments

This work was supported by the project CNPq (process number 484847/2013-9) that the authors gratefully acknowledge.

Funding

This study was financially supported by Brazilian Science Agency -CNPq (process number 484847/2013-9)

Conflict of interest

The authors declare that they have no conflict of interest.

Authors' contributions

J.T. and A.K.M. conducted the experiment. L.J.C, J.F.S.J and E.F.S conceived the project. C.A. analyzed the results and drafted the manuscript. All authors contributed to the final manuscript draft.

Availability of data

Sequences have been submitted to GenBank (MN249239 to MN249393)

References

- Amorim JAE, Mata L R, Lédo AS, Azevedo VCR, Silva AVC (2015) Diversity and genetic structure of mangaba remnants in states of northeastern Brazil. *Genetics and Molecular Research* 14:823-833
- Assis MA, Prata EMB, Pedroni F, Sanchez M, Eisenlohr PV, Martins FR, Joly CA (2011) Restinga and Lowland forests in coastal plain of southeastern Brazil: vegetation and environmental heterogeneity. *Biota Neotropica* 11:83-101 <http://www.biotaneotropica.org.br/v11n2/en/abstract?article+bn02111022011>
- Bruvo R, Michiels NK, D'Souza TG, Schulenburg H (2004) A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* 13(7):2101-2106
- Buzatti RSO, Pfeilsticker TR, Magalhães RF, Bueno ML, Lemos-Filho JP, Lovato MB (2018) Genetic and historical colonization analyses of an endemic Savanna tree, *Qualea grandiflora*, reveal ancient connections between Amazonian savannas and Cerrado core. *Frontiers in Plant Science* 9:981. <https://doi.org/10.3389/fpls.2018.00981>
- Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability Predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science* 323(5915):785-789. <https://doi.org/10.1126/science.1166955>
- Castro ASF, Moro MF Menezes MOT (2012) O complexo vegetacional da zona litorânea no Ceará: Pecém, São Gonçalo do Amarante. *Acta Botanica Brasilica* 26:108-124
- Chen F, Yu Z, Yang M, Ito E, Wang S, Madsen DB, Wünnemann B (2008) Holocene moisture evolution in arid central Asia and its out-of-phase relationship with Asian monsoon history. *Quaternary Science Reviews* 27:351-364

- Collevatti RG, Olivatii AM, Telles MPC, Chaves LJ (2016) Gene flow among *Hancornia speciosa* (Apocynaceae) varieties and hybrid fitness. *Tree Genetics and Genomics* 12:1-12.
- Collevatti RG, Rodrigues EE, Vitorino LC, Lima-Ribeiro MS, Chaves LJ, Telles MPC (2018) Unravelling the genetic differentiation among varieties of the Neotropical savanna tree *Hancornia speciosa* Gomes. *Annals of Botany* 122:973-984. <https://doi.org/10.1093/aob/mcy060>
- Costa CF, Collevatti RG, Chaves LJ, Lima JS, Soares TN, Telles MPC (2017) Genetic diversity and fine-scale genetic structure in *Hancornia speciosa* Gomes (Apocynaceae). *Biochemical Systematics and Ecology* 1:1-5
- Costa TS, Silva AVC, Lêdo AS, Santos ARF, Silva Júnior JF (2011) Diversidade genética de acessos do banco de germoplasma de mangaba em Sergipe. *Pesquisa Agropecuária Brasileira* 46:499-508.
- Doyle JJ, Doyle JL (1990) A rapid total DNA preparation procedure for fresh plant tissue. *Focus* 12:13-15
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* 11:2571-2581
- Estoup A, Guillot G, Santos F (2007) Geneland: simulation and MCMC inference in landscape genetics R package version 2.0.5 URL: <http://cran.r-project.org/>
- Flores IS, Silva SK, Furquin LC, Castro CFS, Chaves LJ, Collevatti RG, Lião LM (2017) HR-MAS NMR allied to chemometric on *Hancornia speciosa* varieties differentiation. *Journal of the Brazilian Chemical Society* 1:1-7
- Freitas MKC, Coimbra RR, Aguiar GB, Aguiar CBN, Chagas DBC, Ferreira WM, Oliveira RJ (2012) Variabilidade fenotípica e caracterização morfológica de uma população natural de *Hancornia speciosa* Gomes. *Bioscience Journal* 28:833-841
- Ganga RMD, Ferreira GA, Chaves LJ, Naves RV, Nascimento JL (2010) Caracterização de frutos e árvores de populações naturais de *Hancornia speciosa* Gomes do cerrado. *Revista Brasileira de Fruticultura* 32:101-113
- Gruber B, Adamack A (2017) PopGenReport v3.0.0 <https://github.com/green-striped-gecko/PopGenReport> A simple framework to analyse population and landscape genetic data
- Hewitt GM (2004) Genetic consequence of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci* 359:183-195
- Jombart T, Ahmed I (2011) Adegnet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27 (21):3070-3071.

- Katoh K., Misawa K., Kuma K., & Miyata T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 15, 30(14), 3059-3066.
- Koch I, Rapini A, Simões AO, Kinoshita LS, Spina AP and Castello ACD (2015) Apocynaceae in lista de espécies da flora do Brasil. Jardim Botânico do Rio de Janeiro. Available at <>. Accessed on December 20, 2021.
- Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33:1870-1874
- Librato P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452
- Martins GV, Martins LSM, Veasey EA, Lederman IE, Silva EF (2012) Diversity and genetic structure in natural populations of *Hancornia speciosa* var *speciosa* Gomes in northeastern Brazil. *Revista Brasileira de Fruticultura* 34:1143-1153
- Monachino JÁ (1945) revision of *Hancornia* (Apocynaceae). *Lilloa*, 11:19-48
- Morellato LPC, Haddad CFB (2000) Introduction: the Brazilian Atlantic Forest. *Biotropica* 32:786-792
- Moro MF, Silva IA, Araújo FS, Lughadha EN, Meagher TC, Martins FR (2015) The role of edaphic environment and climate in structuring phylogenetic pattern in seasonally dry tropical plant communities. *Journal Plos one* 13:083–970
- Moura NF, Chaves LJ, Venkovsky R, Naves RVN, Aguiar AV, Moura MF (2011) Genetic structure of mangaba (*Hancornia speciosa* Gomes) populations in the cerrado region of central Brazil. *Bioscience Journal* 27:473-481
- Myers, N., Mittermeier, R., Mittermeier, C. *et al.* Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858 (2000). <https://doi.org/10.1038/35002501>
- Nascimento RSM, Cardoso JA, Coccozza FDM (2014) Caracterização física e físico-química de frutos de mangabeira (*Hancornia speciosa* Gomes) no oeste da Bahia. *Revista Brasileira de Engenharia Agrícola e Ambiental* 18:856-860
- Novaes RM, Ribeiro RA, Lemos-Filho JP, Lovato MB (2013) Concordance between phylogeographical and biogeographical patterns in the Brazilian Cerrado: diversification of the endemic tree *Dalbergia miscolobium* (Fabaceae). *PLoS One* 8:e82198
- Paradis E, Gosselin T, Goudet J, Jombart T, Schliep K (2017) Linking genomics and population genetics with R. *Molecular Ecology Resources* 17:54-66
- Rambaut A, Drummond AJ (2009) Tracer version 1.5 [computer program] <http://beast.bio.ed.ac.uk>

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9(3):552-69

Sacramento ACS, Almeida JrEB, Zickel CS (2007) Aspectos florísticos da vegetação de restinga no litoral de Pernambuco. *Revista Árvore* 31:1121-1130

Scarcelli N, Barnaud A, Eiserhardt W, Treier UA, Seveno M, D’Anfray A, Vigouroux Y, Pintaud JC (2011) Set of 100 chloroplast DNA primer pairs to study population genetics and phylogeny in monocotyledons. *Plos One* 6:e19954.

Silva AVC, Rabbani ARC, Sena Filho JG, Almeida CS, Feitosa RB (2012) Genetic diversity analysis of mangaba (*Hancornia speciosa* Gomes), an exotic brazilian tropical species. *Tropical and Subtropical Agroecosystems* 15:217-225

Silva AVC, Santos ARF, Wickert E, Silva Júnior JF, Costa TS (2011) Divergência genética entre acessos de mangabeira (*Hancornia speciosa* Gomes). *Revista Brasileira de Ciências Agrárias* 6:572-578

Silva Júnior JF, Lêdo AS (2006) A cultura da mangaba Embrapa Tabuleiros Costeiros, Aracaju, 253p.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data III Cladogram estimation. *Genetics* 132(2):619-33

Thomé MTC, Zamudio KR, Giovanelli JGR, Haddad CFB, Baldissera FA, Alexandrino J (2010) Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Molecular Phylogenetics and Evolution* 55:1018-1031

Tables

Table 1. Spatial Analysis of Molecular Variance (SAMOVA) between groups

Source of variation	cpDNA		nSSR	
	Variance Components	Percentage of variation	Variance Components	Percentage of variation
Among groups*	1.38	93.90	10.024	43.02
Among population within groups	-0.0036	-0.025	1.0	4.29
Within populations	0.09353	6.34	12.28	52.69
Total	1.47	100	23.3	100
F _{ST}	0.94 (p<0.001)	-	0.47 (p<0.001)	

* Group A (Paraíba, Pernambuco, Sergipe); Group B (Maranhão); Group C (Bahia, Ceará); Group D (Goiás); Group E (Minas Gerais)

Figures

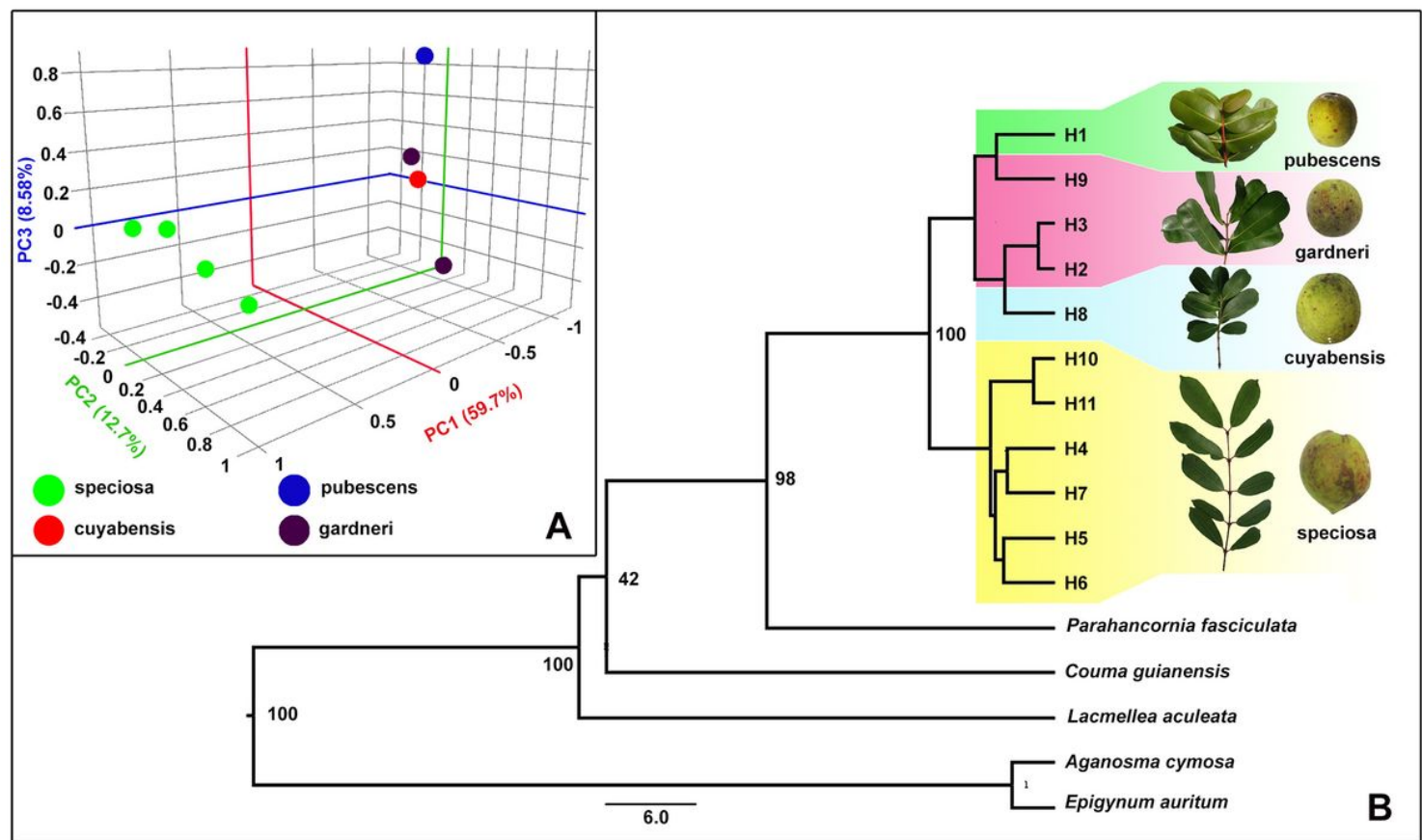


Figure 1

Genetic relationships among haplotypes of the *trnH-psbA* sequence. (A) Principal component analysis (PCA) revealing a three-dimensional distribution of *Hancornia speciosa* genotypes according to the botanical varieties. (B) Phylogenetic analysis using Bayesian analysis, with support inferred by posterior probability (here, represented in percentage).

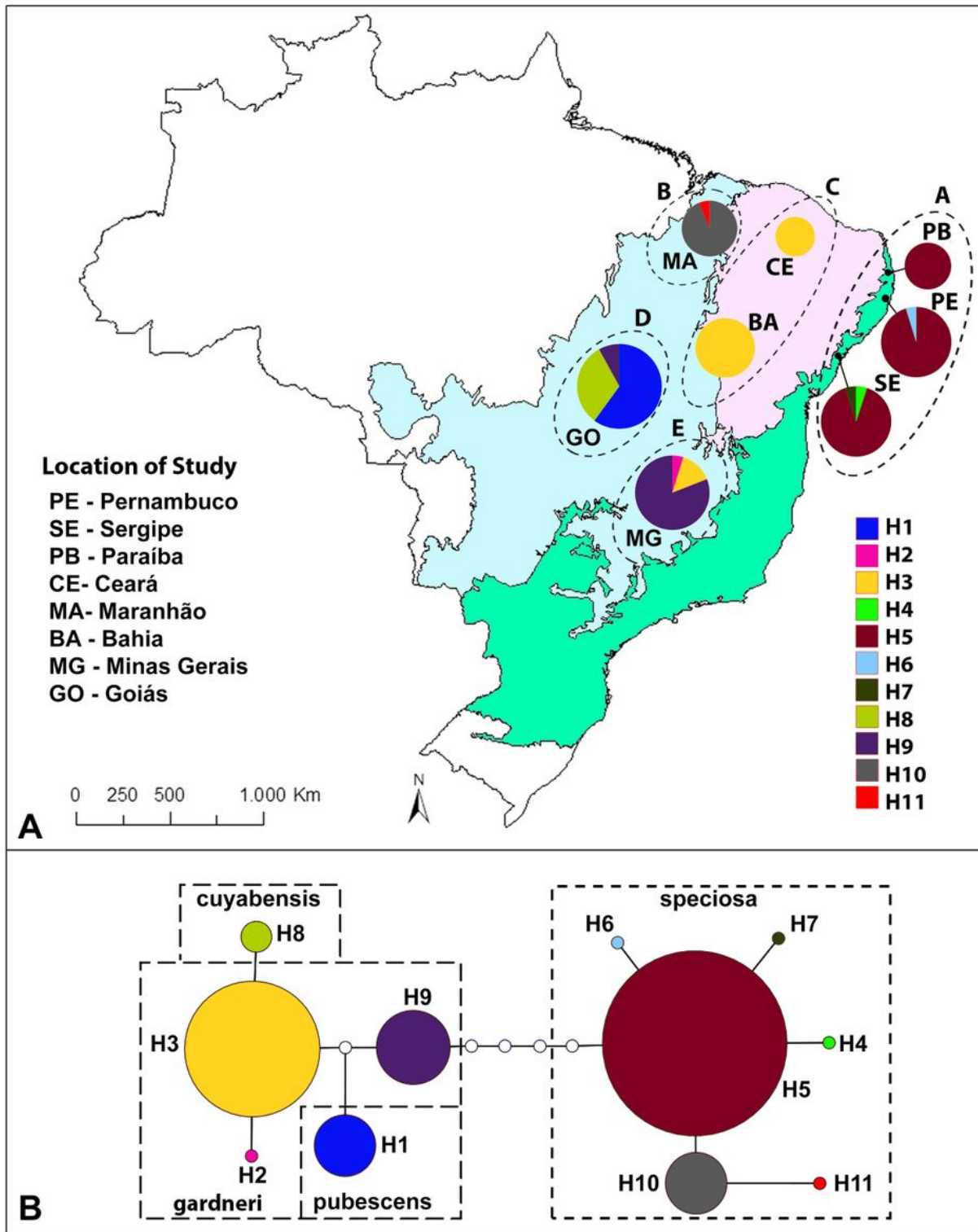


Figure 2

Patterns of distribution of *Hancornia speciosa* haplotypes. (A) Frequency of haplotypes for the *trnH-psbA* intergenic spacer per region site. The size of the graphic is proportional to the size of the sample; the coloured portions represent the *Caatinga* domain (pink), *Cerrado* domain (sky blue) and Atlantic Forest domain (green). The colours on the pie graph correspond to haplotypes following the names in the legend and SAMOVA groups, which are encircled by dotted lines. (B) Haplotypes network obtained by maximum

parsimony statistics using the haplotype software. The sizes of the circles correspond to the proportion of each haplotype and the names correspond to the botanical varieties.

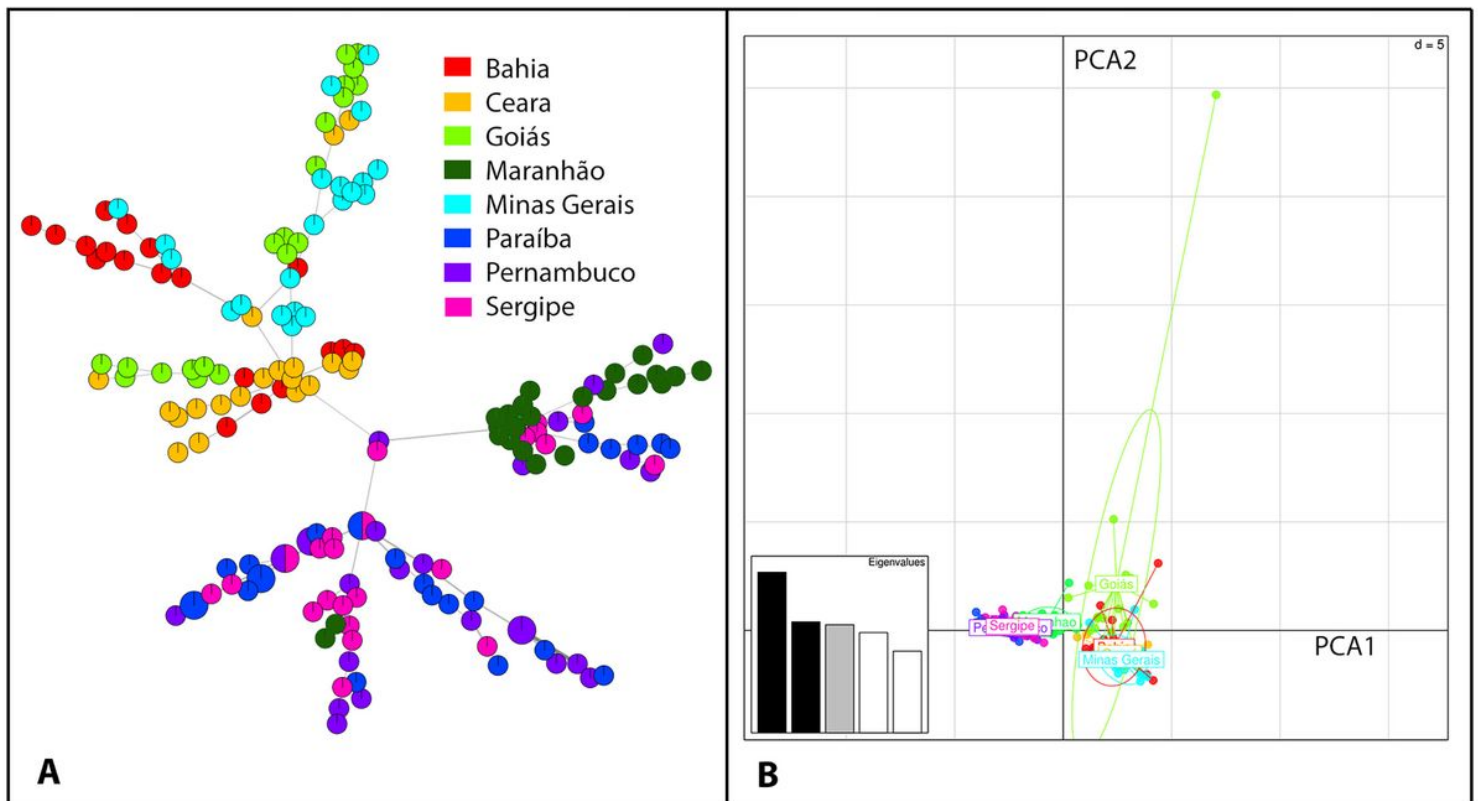


Figure 3

Genetic relationships among genotypes of SSR loci. (A) Minimum-spanning network depicting the distribution of *Hancornia speciosa* multilocus microsatellites. (B) Principal component analysis (PCA) revealing a two-dimensional distribution of *Hancornia speciosa* genotypes

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [TableS2.docx](#)
- [FigureS1.tif](#)
- [FigureS2.tif](#)
- [FigureS3.tif](#)