Genetic diversity of *Lippia sidoides* Cham. and *L. gracilis* Schauer germplasm

C.P. Santos[1], J.B. Pinheiro[2], M.I. Zucchi[3], M.M. Bajay[2], J.B. Campos[3], M.F. Arrigoni-Blank[1], J.A.O. Pinto[1] and A.F. Blank[1]

[1]Laboratório de Recursos Genéticos Vegetais e Óleos Essenciais, Departamento de Engenharia Agronômica, Universidade Federal de Sergipe, São Cristóvão, SE, Brasil
[2]Laboratório de Diversidade Genética e Melhoramento, Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, SP, Brasil
[3]Laboratório de Biologia Molecular, Agência Paulista de Tecnologia dos Agronegócios, Polo Centro Sul, Piracicaba, SP, Brasil

Corresponding author: A.F. Blank
E-mail: afblank@ufs.br

Genet. Mol. Res. 15 (3): gmr.15037781
Received January 4, 2016
Accepted August 4, 2016
Published September 23, 2016
DOI http://dx.doi.org/10.4238/gmr.15037781

ABSTRACT. The conservation of plants in germplasm banks ensures the characterization and availability of these resources for future generations. The present study used DNA markers to obtain genetic information about germplasm collections of *Lippia sidoides* and *L. gracilis*, which are maintained in an Active Germplasm Bank (AGB). Genetic variability of samples in the AGB was assessed using 12 combinations of amplified fragment length polymorphism (AFLP) primers (*Eco*RI/*Mse*I). Twenty simple sequence repeat primers designed for *L. alba* were tested to determine their transferability in

L. sidoides and *L. gracilis*. The AFLP markers generated 789 markers. The assessed loci exhibited a moderate Shannon diversity index (I = 0.42) in both species, suggesting that the conserved accessions possess an intermediate level of genetic diversity. Twelve microsatellite loci amplified satisfactorily, and nine loci were polymorphic in each species. A total of 23, 22, and 36 alleles, with an average of 2.5, 2.4, and 3.27 alleles per locus were identified for *L. sidoides* and *L. gracilis* accessions in the AGB, and *Lippia* sp sampled plants, respectively. Analyses of genetic structure permitted the identification of three different groups using both sets of markers, of which two were representative of *L. sidoides*. The information generated in this study may help to create, expand, and maintain collections of these species and may assist in genetic-breeding programs.

Key words: *Lippia sidoides*; *Lippia gracilis*; Medicinal plants; AFLP markers; Microsatellites; Accessions

INTRODUCTION

Brazil is a leading center of species diversity within the genus *Lippia* because it possesses approximately 70-75% of all catalogued species. Most of these species are endemic to Brazil and are found in risk areas because of anthropogenic action (Viccini et al., 2006).

Lippia sidoides Cham., commonly known as "*alecrim-pimenta*", is an aromatic shrubby species, native to the northeastern Brazilian semi-arid region (Lorenzi and Matos, 2002). Due to the chemical composition of its essential oil, several biological studies have demonstrated the medicinal potential of this species, with its bactericidal and fungicidal activities (Botelho et al., 2007), larvicidal (*Aedes aegypti*) (Carvalho et al., 2003), and acaricidal (*Tetranychus urticae*) (Cavalcanti et al., 2010) activities being the most prominent.

Another species, *Lippia gracilis* Schauer, commonly known as "*alecrim de tabuleiro*", is endemic to northeastern Brazil (Lorenzi and Matos, 2002). Several biological activities have already been reported for the essential oil of this species, which include bactericidal and fungicidal (Pessoa et al., 2005), larvicidal (Silva et al., 2008), and acaricidal (Cruz et al., 2013) activities. These properties are due to the presence of the monoterpenes carvacrol and thymol.

These species are found in the Caatinga biome, which is relatively representative of the northeastern Brazilian semi-arid region because it contains a wide biodiversity of plants and animals. However, this region has been subjected to increasingly severe desertification processes with the introduction of mechanized agriculture and rapid population growth, which are aggravated by prolonged droughts. These processes result in the loss of genetic resources and often involve species that are endemic to the region (Goedert, 2007).

The medicinal and aromatic species, *L. sidoides* and *L. gracilis*, are important plant genetic resources that need to be conserved and characterized, given their medicinal importance and sustainable economic potential.

Various Brazilian research centers maintain field germplasm collections of these species to overcome excessive germplasm loss. Universidade Federal de Sergipe established an Active Germplasm Bank (AGB) of *L. sidoides* and *L. gracilis*, and their accessions are maintained in the field. However, molecular studies are essential for characterizing germplasm

banks because knowledge of the genetic diversity in them is required for the proper management and implementation of efficient conservation strategies. Furthermore, these studies permit the most representative genotypes or clones to be identified, therefore reducing costs associated with the maintenance of genetically identical accessions.

The collection of germplasms derived from native and endemic species continues to demand increased human and financial resources. Nevertheless, efforts have been made, and a collection of *Lippia* sp germplasm was collected to expand the AGB by introducing the most representative genotypes.

In this study, to assess genetic diversity of germplasm collections, both amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) markers were used, which permits increased genome coverage and generates more robust results. Microsatellite or SSR markers are commonly used in the molecular characterization of germplasm banks because of their high polymorphism and reproducibility, multi-allelic nature, codominance, and wide distribution throughout the genome (Guimarães et al., 2009). Furthermore, SSR markers may be transferred between evolutionarily related species because the flanking regions are mostly conserved, which reduces the cost and time required to develop site-specific markers (Kuleung et al., 2004). AFLP markers were used in addition to microsatellite markers. Despite its dominant nature, the AFLP method has high reproducibility and a multi-loci nature, enabling wide genome coverage without any prior knowledge. Therefore, this method is highly informative, robust, and reliable (Kumar et al., 2009).

The use of molecular markers is essential because it enables more effective access to the variability that exists within a germplasm bank. Data may then be correlated with the chemical analyses, and these two tools are complementary when uncorrelated (Morone-Fortunato et al., 2010), enabling improved differentiation between the accessions studied; thus, selection of the most representative accessions based on the trait of interest (i.e., the essential oil) can be determined.

Therefore, the objective of the present study was to characterize the structure and genetic variability of the accessions and collected plant samples of *L. sidoides* and *L. gracilis*, using AFLP and SSR molecular markers to obtain data for the proper management of the AGB and for the introduction of genotypes to expand the AGB.

MATERIAL AND METHODS

Plant material

The AGB of *L. sidoides* and *L. gracilis* is maintained in a Research Farm called "Campus Rural da UFS", located in São Cristóvão municipality, Sergipe State, Brazil, 1°00'S latitude and 37°12'W longitude. The AGB contains 12 accessions of *L. sidoides* (LS) and 7 accessions of *L. gracilis* (LG) (Table 1). To amplify the studied AGB, 40 *Lippia* sp plants were collected for molecular analyses using microsatellite markers only (Table 2). Of these, 34 plants were *L. gracilis* and were distributed as follows: 13 plant samples from Tobias Barreto municipality, 12 plant samples from Poço Verde municipality, and 9 plant samples from São Domingos municipality, all collected in the Sergipe State, Brazil. Universidade Federal do Ceará donated two of the six LS plants.

Table 1. Identification and geographic origin of *Lippia sidoides* and *L. gracilis* accessions conserved in the Active Germplasm Bank of Universidade Federal de Sergipe.

Code	Scientific name	Origin	Geographic data	Voucher No.
LSID002	*Lippia sidoides*	Mossoró, Rio Grande do Norte, Brazil	5° 07' 26.7" S; 37° 24' 14.6" W	8219
LSID003	*Lippia sidoides*	Mossoró, Rio Grande do Norte, Brazil	5° 08' 28.3" S; 37° 23' 58.0" W	8220
LSID004	*Lippia sidoides*	Quixeré, Ceará, Brazil	5° 05' 03.5" S; 37° 58' 43.9" W	8221
LSID005	*Lippia sidoides*	Limoeiro do Norte, Ceará, Brazil	5° 09' 47.8" S; 38° 06' 31.0" W	8222
LSID006	*Lippia sidoides*	Tabuleiro do Norte, Ceará, Brazil	5° 14' 05.4" S; 38° 11' 35.0" W	8223
LSID102	*Lippia sidoides*	Poço Redondo, Sergipe, Brazil	9° 58' 07.6" S; 37° 51' 49.2" W	8224
LSID103	*Lippia sidoides*	Poço Redondo, Sergipe, Brazil	9° 58' 08.6" S; 37° 51' 50.3" W	8225
LSID104	*Lippia sidoides*	Poço Redondo, Sergipe, Brazil	9° 58' 09.2" S; 37° 51' 50.3" W	8226
LSID105	*Lippia sidoides*	Poço Redondo, Sergipe, Brazil	9° 58' 12.9" S; 37° 51' 49.2" W	8227
LSID111	*Lippia sidoides*	Poço Redondo, Sergipe, Brazil	9° 58' 12.9" S; 37° 51' 50.7" W	9384
LSID112	*Lippia sidoides*	Poço Redondo, Sergipe, Brazil	9° 58' 09.6" S; 37° 51' 50.7" W	9385
LSID301	*Lippia sidoides*	Recife, Pernambuco, Brazil	8° 01' 05.0" S; 34° 56' 48.0" W	8228
LGRA106	*Lippia gracilis*	Tomar do Geru, Sergipe, Brazil	11° 19' 16.7" S; 37° 55' 09.2" W	14733
LGRA107	*Lippia gracilis*	Tomar do Geru, Sergipe, Brazil	11° 19' 20.1" S; 37° 55' 13.5" W	14737
LGRA108	*Lippia gracilis*	Tomar do Geru, Sergipe, Brazil	11° 19' 22.4" S; 37° 55' 12.6" W	14734
LGRA109	*Lippia gracilis*	Tomar do Geru, Sergipe, Brazil	11° 19' 20.7" S; 37° 55' 16.9" W	14735
LGRA110	*Lippia gracilis*	Tomar do Geru, Sergipe, Brazil	11° 19' 21.1" S; 37° 55' 14.9" W	14732
LGRA201	*Lippia gracilis*	Rio Real, Bahia, Brazil	11° 23' 38.7" S; 38° 00' 54.1" W	14736
LGRA202	*Lippia gracilis*	Rio Real, Bahia, Brazil	11° 23' 45.3" S; 38° 00' 51.3" W	14731

DNA extraction and quantification

Young leaves from each accession were properly labeled, wrapped in gauze, and immediately stored in liquid nitrogen. Next, the samples were freeze-dried, ground, and stored in a freezer at -20°C until use. Genomic DNA was extracted according to the cetyltrimethylammonium bromide (CTAB) protocol (Doyle and Doyle, 1990), with modifications. Approximately 100 ng powder was transferred to a 2-mL tube and stored in 700 µL CTAB buffer [2%, 1.4 mM sodium chloride (NaCl), 1 mM Tris, pH 8.0, 2% polyvinylpyrrolidone, 3.5% β-mercaptoethanol] and incubated in a water bath at 65°C for 1 h.

Following incubation, 600 µL chloroform/isoamyl alcohol (24:1) mixture was added, and samples were homogenized and centrifuged at 14,534 *g* for 10 min. The resulting supernatant was transferred to a new tube, and chloroform/isoamyl alcohol (24:1) purification was repeated. One-quarter of a 5 M NaCl solution was added for a final concentration of 1 M, and the mixture was centrifuged at 20,929 *g* for 12 min to precipitate proteins and polysaccharides.

The supernatant was transferred to another tube, containing 500 µL ice-cold isopropanol and stored in a freezer at -20°C for 1 h to precipitate the DNA. The samples were centrifuged at 20,929 *g* for 15 min, and the supernatant was discarded. The resulting pellet was washed with 500 µL 70% ethanol, followed by the addition of 300 µL absolute ethanol and then centrifuged at 20,929 *g* for 5 min. The pellet was dried overnight and diluted in 100 µL TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA), with the addition of 1 µL RNAse (10 mg/mL), and incubated in a water bath at 37°C for 3 h. The tubes were then stored in a freezer at -20°C.

The quality of the extracted DNA was assessed on a 1% agarose gel [1.0 g agarose, 100 mL 1X Tris-borate-EDTA] subjected to electrophoresis for 1 h at 80 V. DNA of the phage lambda (λ) (Invitrogen, Carlsbad, CA, USA) was used to estimate the resulting DNA concentration. The gel was stained using SYBR® safe (Invitrogen), visualized under ultraviolet light, and photographed using an imaging system.

Table 2. Identification and geographic origin of the collected *Lippia* sp accessions.

Plant code	Origin	Geographic data
LG01	Tobias Barreto, Sergipe, Brazil	11° 07' 22.8" S; 37° 53' 06.7" W
LG02	Tobias Barreto, Sergipe, Brazil	11° 07' 19.7" S; 37° 52' 49.9" W
LG03	Tobias Barreto, Sergipe, Brazil	11° 07' 25.6" S; 37° 53' 06.6" W
LG04	Tobias Barreto, Sergipe, Brazil	11° 02' 03.4" S; 38° 03' 23.2" W
LG05	Tobias Barreto, Sergipe, Brazil	11° 02' 00.2" S; 38° 03' 24.8" W
LG06	Tobias Barreto, Sergipe, Brazil	11° 01' 43.7" S; 38° 03' 30.4" W
LG07	Tobias Barreto, Sergipe, Brazil	11° 00' 11.3" S; 38° 04' 05.0" W
LG08	Tobias Barreto, Sergipe, Brazil	11° 07' 19.4" S; 37° 53' 10.5" W
LG09	Tobias Barreto, Sergipe, Brazil	11° 07' 42.8" S; 37° 53' 51.0" W
LG10	Tobias Barreto, Sergipe, Brazil	11° 08' 59.8" S; 37° 57' 04.8" W
LG11	Tobias Barreto, Sergipe, Brazil	11° 06' 06.2" S; 38° 01' 26.0" W
LG12	Tobias Barreto, Sergipe, Brazil	11° 01' 53.4" S; 38° 03' 21.5" W
LG13	Tobias Barreto, Sergipe, Brazil	11° 00' 37.2" S; 38° 03' 54.8" W
LG14	Poço Verde, Sergipe, Brazil	10° 44' 02.5" S; 38° 07' 03.4" W
LG15	Poço Verde, Sergipe, Brazil	10° 58' 39.1" S; 38° 04' 32.1" W
LG16	Poço Verde, Sergipe, Brazil	10° 44' 03.4" S; 38° 07' 02.7" W
LG17	Poço Verde, Sergipe, Brazil	10° 59' 03.5" S; 38° 04' 30.1" W
LG18	Poço Verde, Sergipe, Brazil	10° 44' 01.8" S; 38° 07' 00.3" W
LG19	Poço Verde, Sergipe, Brazil	10° 44' 02.5" S; 38° 06' 54.7" W
LG20	Poço Verde, Sergipe, Brazil	10° 44' 04.1" S; 38° 06' 48.2" W
LG21	Poço Verde, Sergipe, Brazil	10° 56' 31.2" S; 38° 05' 26.2" W
LG22	Poço Verde, Sergipe, Brazil	10° 44' 05.0" S; 38° 06' 47.0" W
LG23	Poço Verde, Sergipe, Brazil	10° 53' 15.8" S; 38° 06' 37.9" W
LG24	Poço Verde, Sergipe, Brazil	10° 44' 05.7" S; 38° 06' 43.0" W
LG25	Poço Verde, Sergipe, Brazil	10° 44' 11.4" S; 38° 06' 12.3" W
LG26	São Domingos, Sergipe, Brazil	10° 46' 09.3" S; 37° 58' 17.1" W
LG27	São Domingos, Sergipe, Brazil	10° 46' 10.0" S; 37° 58' 15.3" W
LG28	São Domingos, Sergipe, Brazil	10° 46' 09.7" S; 37° 58' 14.9" W
LG29	São Domingos, Sergipe, Brazil	10° 46' 10.7" S; 37° 58' 16.2" W
LG30	São Domingos, Sergipe, Brazil	10° 46' 10.8" S; 37° 58' 14.7" W
LG31	São Domingos, Sergipe, Brazil	10° 46' 11.4" S; 37° 58' 16.3" W
LG32	São Domingos, Sergipe, Brazil	10° 46' 24.8" S; 37° 59' 16.7" W
LG33	São Domingos, Sergipe, Brazil	10° 46' 19.2" S; 37° 59' 15.0" W
LG34	São Domingos, Sergipe, Brazil	10° 46' 08.9" S; 37° 59' 10.4" W
LS01	Poço Redondo, Sergipe, Brazil	9° 59' 38.5" S; 37° 48' 46.1" W
LS02	Poço Redondo, Sergipe, Brazil	9° 59' 35.5" S; 37° 48' 44.2" W
LS03	Poço Redondo, Sergipe, Brazil	9° 59' 43.6" S; 37° 49' 05.8" W
LS04	Poço Redondo, Sergipe, Brazil	9° 59' 34.2" S; 37° 49' 01.6" W
LS05	Donated by Universidade Federal do Ceará, Fortaleza, Ceará, Brazil	-
LS06	Donated by Universidade Federal do Ceará, Fortaleza, Ceará, Brazil	-

AFLP markers

Dominant AFLP markers were used to estimate the genetic variability of the *L. sidoides* and *L. gracilis* accessions conserved in the AGB according to the protocol described by Vos et al. (1995), with modifications. The AFLP method consists of four steps: DNA digestion, adapter ligation, pre-amplification, and selective amplification. Two restriction enzymes were used for the digestion: the rare cutter *Eco*RI and the frequent cutter *Mse*I, which generate three types of fragments: Mse-Mse, Eco-Mse, and Eco-Eco. A total of 300 ng DNA from each accession was used in the digestion reaction in a final volume of 20 µL. The mixture consisted of 12 µL 25 ng/µL DNA, 5.4 µL ultrapure water; 0.2 µL *Eco*RI (4 U) (New England

BioLabs, Ipswich, MA, USA), 0.4 µL *Mse*I (4 U) (New England BioLabs), and 2 µL One-Phor-All buffer plus-OPA 10X (OPA Amersham GE). The reaction was performed in a BIO-RAD thermocycler (Bio-Rad, München, Germany) at 37°C for 3 h, followed by an enzyme inactivation step at 70°C for 15 min.

The adapters were then ligated to the cohesive ends of the fragments generated by the restriction enzymes. The adapters serve as primer annealing sites for PCR amplification in the subsequent steps. Adapter solutions (*Eco*RI Forward: 5'-CTCGTAGACTGCGTACC-3'; *Eco*RI Reverse: 5'-AATTGGTACGCAGTCTAC-3', at a concentration of 5 pM; *Mse*I Forward: 5'-GACGATGAGTCCTGAG-3'; *Mse*I Reverse: 5'-TACTCAGGACTCAT-3', at a concentration of 50 pM) were prepared. The F and R adapters of each enzyme were separately denatured at 94°C for 2 min. The ligation reaction was performed in a final volume of 40 µL, containing 20 µL digestion reaction mix, 12 µL ultrapure water, 4 µL 5X T4 DNA ligase buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 1.0 mM ATP), 1 µL *Eco*RI adapters (5 µM), 1 µL *Mse*I adapters (50 µM), 1 µL ATP (10 mM) (New England BioLabs), and 1 U/µL T4 DNA ligase enzyme (Invitrogen). The ligation was performed in a thermocycler (Bio-Rad) at 37°C for 2 h, followed by 16 h at 16°C.

Primers complementary to the *Eco*RI and *Mse*I adapters, and complementary to the enzyme restriction sites, with a selective nucleotide at the 3'-end (*Eco*RI+N: 5'-AGACTGCGTACCAATTC-3'; *Mse*I+N: 5'-GATGAGTCCTGAGTAA-3'), were used for the pre-amplification reaction. Approximately 1/16th of the fragments generated in the digestion were amplified, therefore reducing the pattern of complexity. The reaction was performed in a final volume of 15 µL, with 2 µL from the ligation reaction diluted 6-fold in ultrapure water, 8.3 µL ultrapure water, 10X Taq DNA polymerase enzyme buffer [75 mM Tris-HCl, 20 mM $(NH_4)_2SO_4$, pH 8.8, 0.01% Tween 20 (v/v)], 2.5 U/µL Taq DNA polymerase (Fermentas), 1 µL 10 mM dNTP, and 1 µL each *Eco*RI and *Mse*I primer at 50 µM. The PCR was performed according to the following program: 29 cycles at 94°C for 30 s, 56°C for 1 min, and 72°C for 1 min.

In total, 14 primer combinations were tested in the final step of selective amplification. The number of pre-amplified fragments was reduced to approximately 1/64 in this step, providing a number suitable for visualization on gel electrophoresis. The *Eco*RI and *Mse*I selective primers each contained three selective nucleotides. The two *Eco*RI primers (E-ACA and E-ACC) were labeled with the fluorophores IRDye 700 and IRDye 800 (LI-COR Biosciences, Lincoln, NE, USA), respectively, which were combined with the *Mse*I primers (CAC, CTG, CAA, CTC, CAG, CTT, and CTA) (Table 3).

The selective amplification reaction was performed in a final volume of 10 µL, with 2 µL pre-amplification reaction mixture diluted 10-fold, 3.73 µL ultrapure water, 10X Taq DNA polymerase enzyme buffer (Fermentas) [75 mM Tris-HCl, 20 mM $(NH_4)_2SO_4$, pH 8.8, 0.01% Tween 20 (v/v)], 0.8 µL 25 mM $MgCl_2$, 1.07 µL 2.5 mM dNTP, 1 µL 1 µM *Eco*RI labeled primers and 0.25 µL 10 mM *Mse*I primers, and 0.15 µL (2.5 U/µL) Taq DNA polymerase (Fermentas). The reactions performed with various combinations of *Eco*RI and *Mse*I primers always included one labeled and one unlabeled primer. The PCR was performed under the following conditions: an initial cycle at 94°C for 30 s, 65°C for 30 s, 72°C for 1 min, followed by 12 cycles at 94°C for 30 s, 65°C for 30 s (decreasing 0.7°C per cycle), and 72°C for 1 min, and additional 23 cycles at 94°C for 30 s, 56°C for 30 s, and 72°C for 1 min.

SSR locus transferability

A set of 20 primers designed for *L. alba* (Santos et al., 2012; Rocha et al., 2015) was tested in accessions of *L. sidoides*, *L. gracilis*, and *Lippia* sp to obtain primers for the amplification of specific, polymorphic bands that can be used to characterize the accessions in this study. An *M13 tail (5'-CACGACGTTGTAAAACGAC-3'), complementary to the tail added to the fluorescent dyes IRDye 700 and IRDye 800, was added to the 5'-end of the forward primer for each locus. Information on the transferred markers and the accession numbers of the corresponding sequences deposited in the National Center for Biotechnology Information (NCBI) database are available in Santos et al. (2014).

The optimization of SSR primers involved DNA samples from four specimens of each species, *L. alba*, *L. sidoides*, and *L. gracilis* (data not shown). Sampling was crucial to assess the loci that are conserved and can be transferred between species, and to determine the optimal amplification conditions. *L. alba* specimens were used as transfer controls because loci that only amplified *L. alba*-specific bands were considered to be no longer conserved between species.

The amplification reactions were performed in a total volume of 20 µL, containing 2 µL 10 ng DNA, 10 µL ultrapure water, 2 µL Taq DNA polymerase buffer (75 mM Tris-HCl), 0.5 µL (2.5 mM) bovine serum albumin, 1.0 µL dNTP (2.5 mM), 0.2 µL (2.5 U/µL) Taq DNA polymerase (Fermentas), 0.35 µL each primer (forward + reverse at 10 mM), and 0.2 µL (10 mM) of each tail labeled with the fluorescent dyes IRDye 700 or IRDye 800 (LI-COR Biosciences). The loci were amplified using the touchdown program under the following conditions: an initial cycle at 94°C for 5 min, followed by 10 touchdown cycles at 94°C for 40 s for denaturation, the specific annealing temperature of each primer for 40 s (-1°C), 72°C for 1 min to extend the fragments, an additional 30 cycles at 94°C for 40 s, 40°C for 40 s, 72°C for 1 min, and a final extension cycle of 72°C for 10 min.

AFLP and SSR locus genotyping

The reaction products were separated in an automated DNA sequencer, LI-COR model 4300 (LI-COR Biosciences), equipped with two infrared lasers that can simultaneously detect two wavelengths (700-800 nm). Mixtures were prepared at a 1.0:1.0 µL reaction ratio for this purpose using a primer labeled with the fluorescent dye IRDye 700 and another primer labeled with IRDye 800, placed in the same reaction plate with 5 µL water and 3 µL bromophenol blue, with 90% formamide. Next, this mixture was denatured in a thermocycler at 95°C for 5 min and maintained at 4°C. A volume of 0.6 µL of each sample was loaded on a 6.5% polyacrylamide gel pre-heated for 30 min. A 0.25-µL sample of markers containing 50-700-bp fragments for the AFLP markers and 50-350-bp fragments for the microsatellite markers labeled with the fluorescent dyes IRDye 700 and IRDye 800 (LI-COR Biosciences) were also loaded onto the gel. Digital gel images were automatically generated for each fluorescence filter during electrophoresis. These images were subsequently used to determine the size of the alleles obtained at each loci using SAGA^{MX} version 3.3 (LI-COR Biosciences). The genotyping of AFLP markers was scored as follows: (1) for the presence and (0) for the absence of bands.

Statistical and genetic analyses

AFLP markers

The STRUCTURE V. 2.3.4 program (Pritchard et al., 2000) was used to determine the number of the most likely groups (K) of samples using Bayesian statistics to assess the genetic structure of AGB, visualized from the color pattern. The model used was haploid, without mixing, independent alleles, with K values ranging from 1 to 6, with 10 runs for each K value, with 200,000 burn-ins and 500,000 Markov chain Monte Carlo simulations. The most likely K was determined based on the ΔK values, which were obtained from the website STRUCTURE HARVESTER (http://taylor0.biology.ucla.edu/structureHarvester/) following the criteria reported by Evanno et al. (2005).

A simplified representation of genetic similarity was completed by designing a dendrogram using the neighbor-joining clustering method, followed by principal component analysis (PCA) using DARwin version 5.0 (Perrier and Jacquermoud-Collet, 2006).

Analyses of molecular variances (AMOVAs) were conducted by breaking the total variation into components, between and within species and clusters, using the squared interpoint distances, as reported by Excoffier et al. (1992), using Arlequin version 3.0 (Excoffier et al., 2005).

POPGENE version 1.32 (Yeh et al., 1999) (Population Genetic Analysis) was used to estimate the genetic diversity using parameters for diploid dominant data. This software assumes that all loci are in Hardy-Weinberg equilibrium when calculating estimates of allele frequencies, given the dominant nature of the data. The estimates obtained for each group were the number and percentage of polymorphic loci, expected heterozygosity (H_E), total heterozygosity (H_T), genetic differentiation index (G_{ST}) (Nei, 1978), and Shannon's diversity index (I).

Microsatellite markers

STRUCTURE was used to determine the structure of the AGB using microsatellite markers. The specimens were presumed diploid in this model considering the aforementioned parameters (Pritchard et al., 2000; Evanno et al., 2005).

Allele frequency, observed heterozygosity (H_O), H_E, allelic range, and polymorphic information content (PIC) were assessed using the MStools software (Park, 2001). The Genetic Data Analysis (GDA) software provided the private alleles for each cluster and species (Lewis and Zaykin, 1999). The number of alleles and the heterozygosity estimators, as proposed by Nei (1978), were assessed using FSTAT 2.9.3 (Goudet, 1995). The most genetically divergent accessions of *L. sidoides* and *L. gracilis* from the AGB and of the collected *Lippia* sp accessions were identified using PowerCore version 1.0.

The modified Rogers' genetic distance matrix (Wright, 1978) was assessed using the TFPGA software (Miller, 1997). The matrix generated was then entered into DARwin version 5.0, which generated a dendrogram according to the neighbor-joining clustering criterion and PCA (Perrier and Jacquermoud-Collet, 2006). Bootstrapping was performed to assess the consistency of each cluster using 10,000 resamplings through the Booch Object Oriented Design (BOOD) application (Coelho, 2000).

Correlation between markers

Correlation analysis was conducted between the Jaccard dissimilarity matrix generated using the data from the AFLP markers and the genetic distance matrix generated using data from the microsatellite markers. Pearson's correlation and the Mantel test with 1000 permutations were performed using NTSYS pc2.1 (Rohlf, 2000).

RESULTS

AFLP markers

Fourteen combinations of primers were assessed. Of these, 12 combinations generated amplificons forming a high-quality, polymorphic pattern (Table 3) with fragments ranging from 50 to 350 bp (Figure 1).

Table 3. Estimates of diversity for each primer combination of *Lippia sidoides* (Clusters I and II) and *L. gracilis* (Cluster III), number of loci (L), polymorphic loci (PL), and percentage of polymorphic loci (PPL%).

	Lippia sidoides						*Lippia gracilis*		
Primers	Cluster I			Cluster II			Cluster III		
	L	PL	PPL (%)	L	PL	PPL (%)	L	PL	PPL (%)
ACA/CAG	63	14	22.22	63	49	77.78	63	55	87.30
ACA/CAC	65	12	18.46	65	46	70.77	65	55	84.62
ACA/CTT	60	13	21.67	60	43	71.67	60	58	88.33
ACA/CTG	74	10	13.51	74	35	47.30	74	70	94.59
ACA/CAA	74	9	12.16	74	36	48.65	74	40	54.05
ACA/CTA	73	16	21.92	73	33	45.21	73	65	89.04
ACA/CTC	58	12	20.69	58	25	43.10	58	55	94.83
ACC/CAC	50	8	16.00	50	26	52.00	50	42	84.00
ACC/CTG	51	17	33.33	51	36	70.59	51	44	86.27
ACC/CTC	74	8	10.80	74	43	58.11	74	63	85.14
ACC/CTT	62	9	14.52	62	28	45.16	62	58	93.55
ACC/CTA	85	19	22.35	85	54	63.53	85	71	83.53

The genetic structure was determined using STRUCTURE and a Bayesian approach. The most likely K value was assessed using STRUCTURE HARVESTER, which generated $\Delta K = 3$ (Figure 2). Each structural value of K (clusters) is represented by one color (Figure 3).

Accessions of *L. sidoides* and *L. gracilis* from the AGB were clustered into three groups with likely genetic structure. The 12 accessions of *L. sidoides* formed two different clusters according to their origin. The *L. sidoides* accessions LSID002 and LSID003 (Rio Grande do Norte State), LSID004, LSID005, and LSID006 (Ceará State), and LSID301 (Pernambuco State) represented Ccluster I (red; Figure 3), while the *L. sidoides* accessions LSID102, LSID103, LSID104, LSID105, LSID111, and LSID112, collected in Sergipe State, formed Cluster II (blue). The 7 *L. gracilis* accessions LGRA106, LGRA107, LGRA108, LGRA109, LGRA110, LGRA201, and LGRA202 formed Cluster III (green; Figure 3).

Genotypes in the dendrogram were identified using the color determined by STRUCTURE for the cluster of each specimen. The *L. sidoides* accessions were clustered into two different clusters: Cluster I and Cluster II. Cluster III was represented by the *L. gracilis* accessions.

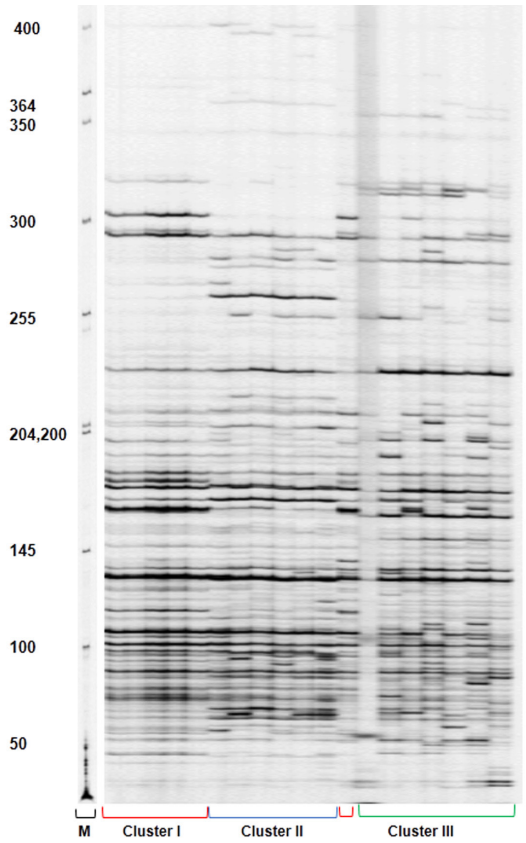

Figure 1. Amplification pattern obtained using amplified fragment length polymorphism (AFLP) markers for the primer combination (*Eco*-ACC/*Mse*-CTA) generated using an automated LI-COR Biosciences sequencer. *Lane M* represents the standard ladder size, Clusters I (red) and II (blue) represent the 12 *Lippia sidoides* accessions, and Cluster III (green) the seven *L. gracilis* accessions of the Active Germplasm Bank, Federal University of Sergipe.

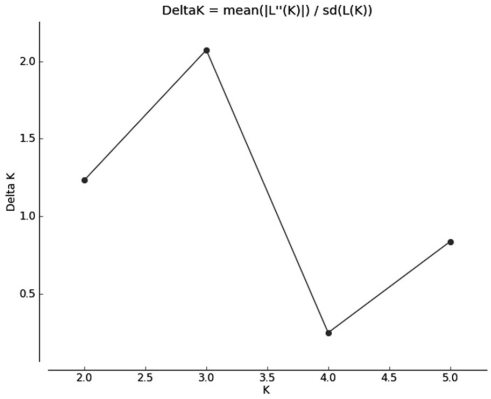

Figure 2. ΔK values obtained using STRUCTURE HARVESTER and AFLP marker combinations. K = 3 was the most likely value, suggesting the presence of three different clusters for the 12 *Lippia sidoides* and 7 *L. gracilis* accessions.

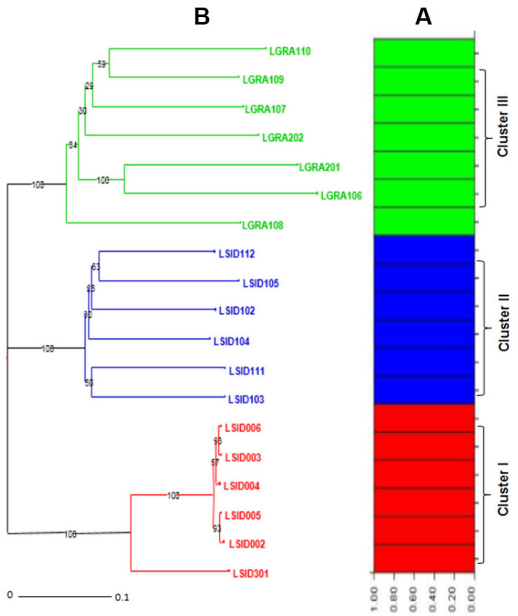

Figure 3. Genetic structure pattern obtained for 12 *Lippia sidoides* and 7 *L. gracilis* accessions from the Active Germplasm Bank (AGB) using 12 AFLP marker combinations. **A.** Plot generated using the STRUCTURE software according to Bayesian statistics. **B.** Dendrogram based on the neighbor-joining clustering criterion. Clusters I and II represent the 12 *L. sidoides* accessions and Cluster III represents the 7 *L. gracilis* accessions.

Cluster I of *L. sidoides* (red) was divided into two subclusters: Subcluster I was represented by accessions LSID002 and LSID003 (Rio Grande do Norte) and LSID004, LSID005, and LSID006 (Ceará), which were genetically similar. Accession LSID301 (Pernambuco) was the most divergent from Cluster I, forming Subcluster II alone (Figure 3). Cluster II (blue) was divided into two subclusters; Subcluster I was represented by accessions LSID111 and LSID103, while accessions LSID102, LSID104, LSID105, and LSID112 formed Subcluster II. Accessions LSID105 and LSID103 were the most divergent (39.7%), and accessions LSID102 and LSID104 were the most similar (68.6%; Figure 3 and Table 4).

The seven *L. gracilis* accessions formed Cluster III (green), which contained three subclusters; Subcluster I was represented by LGRA107, LGRA109, LGRA110 and LGRA202, and Subcluster II consisted of accessions LGRA106 and LGRA201, while accession LGRA108 formed Subcluster III, diverging from the others (Figure 3). Accessions LGRA106 and LGRA109 formed the most divergent pair (58.05%), and LGRA107 and LGRA109 were the least divergent (63.20%; Figure 3 and Table 4).

The dissimilarity between Clusters I and II ranged from 56 to 58.7%; accessions LSID002 and LSID111 were the most divergent (58.7%). The accessions from Clusters I and II of *L. sidoides* diverged from the *L. gracilis* accessions, with variation ranging from 58.67 to 72.87% and from 55.33 to 70.80%, respectively (Table 4).

The first two principal components in the PCA accounted for 61.16% of the total variation found, according to axes 1 and 2 (Figure 4). The PCA permitted the identification of three different clusters according to the results previously shown in the STRUCTURE program and dendrogram.

Table 4. Jaccard similarity matrix generated from 12 amplified fragment length polymorphism (AFLP) primer combinations for the *Lippia sidoides* and *L. gracilis* accessions of the Active Germplasm Bank of the Universidade Federal de Sergipe.

Accessions	LSID002	LSID003	LSID004	LSID005	LSID006	LSID102	LSID103	LSID104	LSID105	LSID111	LSD112	LSID301	LGRA106	LGRA107	LGRA108	LGRA109	LGRA110	LGRA201	LGRA202
LSID002	0.0000																		
LSID003	0.0284	0.0000																	
LSID004	0.0219	0.0153	0.0000																
LSID005	0.0088	0.0240	0.0131	0.0000															
LSID006	0.0330	0.0066	0.0132	0.0219	0.0000														
LSID102	0.5771	0.5732	0.5693	0.5716	0.5725	0.0000													
LSID103	0.5813	0.5820	0.5759	0.5759	0.5813	0.3542	0.0000												
LSID104	0.5810	0.5795	0.5736	0.5758	0.5788	0.3248	0.3511	0.0000											
LSID105	0.5841	0.5863	0.5847	0.5847	0.5879	0.3592	0.3971	0.3647	0.0000										
LSID111	0.5870	0.5809	0.5792	0.5838	0.5779	0.3566	0.3559	0.3563	0.3866	0.0000									
LSID112	0.5621	0.5652	0.5612	0.5612	0.5668	0.3301	0.3574	0.3369	0.3408	0.3441	0.0000								
LSID301	0.2629	0.2469	0.2449	0.2554	0.2454	0.5645	0.5784	0.5668	0.5967	0.5699	0.5657	0.0000							
LGRA106	0.6998	0.6980	0.7008	0.6986	0.6998	0.7055	0.6919	0.6723	0.7080	0.7025	0.6880	0.7287	0.0000						
LGRA107	0.6021	0.6006	0.5990	0.6012	0.6000	0.6030	0.5812	0.5642	0.6006	0.6131	0.5943	0.6175	0.5557	0.0000					
LGRA108	0.5944	0.5928	0.5935	0.5935	0.5944	0.6033	0.5859	0.5533	0.6138	0.6134	0.5990	0.6185	0.5454	0.4659	0.0000				
LGRA109	0.5883	0.5867	0.5874	0.5874	0.5883	0.6042	0.6121	0.5716	0.6033	0.6200	0.6052	0.6056	0.5805	0.3789	0.4400	0.0000			
LGRA110	0.6125	0.6086	0.6092	0.6092	0.6079	0.6428	0.6581	0.6016	0.6623	0.6405	0.6409	0.6317	0.4944	0.4497	0.4864	0.3797	0.0000		
LGRA201	0.6516	0.6428	0.6434	0.6458	0.6422	0.6869	0.6919	0.6442	0.6763	0.7020	0.6649	0.6607	0.4867	0.5477	0.5491	0.4607	0.4822	0.0000	
LGRA202	0.6091	0.6031	0.6015	0.6059	0.6003	0.6230	0.6209	0.6009	0.6292	0.6226	0.6035	0.6272	0.5480	0.4518	0.5120	0.4167	0.4700	0.5415	0.0000

LSID represents the *L. sidoides* accessions, and LGRA represents the *L. gracilis* accessions.

Accessions in *L. sidoides* Cluster I were clustered on the positive side of axis 1. Cluster II of *L. sidoides* was located on the positive side of axis 2. Cluster III, formed by *L. gracilis* accessions, was clustered on the negative side of axes 1 and 2 (Figure 4).

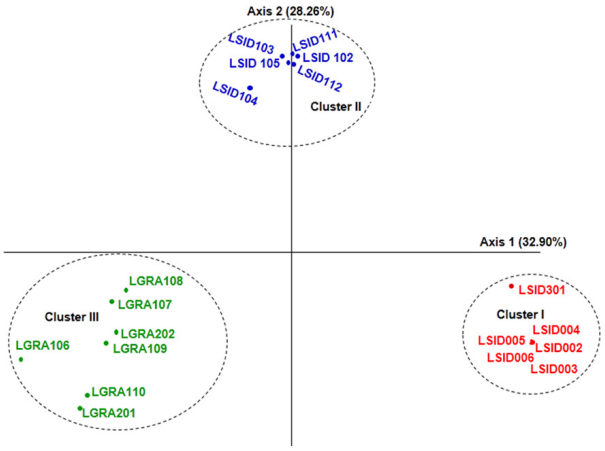

Figure 4. Principal component analysis (PCA) of the 12 *Lippia sidoides* accessions and 7 *L. gracilis* accessions using AFLP markers. Clusters I (red) and II (blue) represent the 12 *L. sidoides* accessions and Cluster III (green) represents the seven *L. gracilis* accessions.

Regarding the genetic diversity structure, AMOVA was performed twice with the accessions subdivided *a priori*. Two populations representative of each species were considered in the first AMOVA (Table 5). The second AMOVA was performed considering subdivision into three clusters (Cluster I + Cluster II + Cluster III), which corresponds to the clustering suggested by STRUCTURE. A total of 54.85% of the genetic variation found was among the three clusters assessed (Table 5).

Table 5. Analysis of molecular variance (AMOVA) performed using 12 *Lippia sidoides* and 7 *L. gracilis* accessions based on 12 amplified fragment length polymorphism marker combinations.

Source of variation	d.f.	Sum of squares	Variation (%)	Statistical Φ	P
Among species	1	713.82	34.73	$\Phi_{ST} = 0.34$	$P \leq 0.000$
Within species	17	2127.33	65.27		
Total	18	2841.15			
Among groups	2	1477.82	54.85	$\Phi_{ST} = 0.54$	$P \leq 0.000$
Within groups	16	1363.33	45.15		
Total	18	2841.15			

d.f. = degrees of freedom.

The combination ACC/CTG was the most polymorphic in Cluster I of *L. sidoides*, with only 33.33% of the 51 loci accessed (Table 3). Conversely, the highest rate of polymorphism in Cluster II was recorded for the combination ACA/CAG, with 77.78%. The genotypes of *L. gracilis* were the most polymorphic of all the combinations analyzed. The highest rate of polymorphism, 94.83%, resulted from the combination ACA/CTC (Table 3).

A total of 789 AFLP markers were identified using 12 primer combinations. Cluster I of *L. sidoides* only exhibited 147 polymorphic loci, which accounted for 18.63% of

polymorphic loci. Cluster II of *L. sidoides* was the most polymorphic, with 454 polymorphic loci, representing 57.54% of the total (Table 6).

Table 6. Estimates of the mean diversity for all combinations inferred for previously assessed clusters and species.

Population	L	PL	PPL (%)	*I*	G_{ST}
Lippia sidoides: Cluster I	789	147	18.63	0.09	-
Lippia sidoides: Cluster II	789	454	57.54	0.29	-
Lippia gracilis: Cluster III	789	690	87.45	0.42	-
Lippia sidoides	789	622	78.83	0.42	0.51
Lippia sidoides + *L. gracilis*	789	770	97.08	0.54	0.51

Number of loci (L), polymorphic loci (PL), percentage of polymorphic loci (PPL%), (*I*) Shannon diversity index, and Nei's differentiation index (G_{ST}).

A total of 622 polymorphic loci were observed in the 12 *L. sidoides* accessions, with 78.83% polymorphism. The seven *L. gracilis* accessions exhibited a total of 690 polymorphic loci, which accounted for approximately 87.45% of the loci analyzed. The combined analysis of *L. sidoides* and *L. gracilis* accessions recorded 770 polymorphic loci, which accounted for more than 97% of the loci analyzed (Table 6).

Genetic diversity was estimated using the Shannon diversity index (*I*). Cluster I of *L. sidoides* had an extremely low diversity value ($I = 0.09$), whereas Cluster II showed a moderate diversity index ($I = 0.29$; Table 6). Combined analysis of the 12 *L. sidoides* accessions generated the same value as that obtained for *L. gracilis* ($I = 0.42$). The Shannon diversity index was higher ($I = 0.54$) when both species were analyzed together. The rate of differentiation between Clusters I and II of *L. sidoides* and between the species was high, $G_{ST} = 0.51$ (Table 6).

SSR markers

Of the 20 SSR primers used in the molecular characterization of *L. sidoides* and *L. gracilis* accessions in the AGB, 12 amplified specific, consistent bands. The rate of transfer was 60% of the primers tested. Regarding the 12 primers transferred the loci LA02 and LAG04 were monomorphic in both species, and therefore, were not informative in this study. LAD03 was monomorphic in *L. sidoides* (Figure 5), while LAE09 was monomorphic in *L. gracilis*. Thus, a total of 9 markers for each species, 10 for the two species of the AGB, and 11, including the *Lippia* sp collected plants, were polymorphic.

The STRUCTURE plot and dendrogram designed using 10 SSR markers clustered the AGB into three different clusters (Figure 6). The clustering pattern was similar to that assessed using the AFLP markers. The most likely K value, assessed using the STRUCTURE HARVESTER, was also K = 3 (Figure 7). The *L. sidoides* accessions collected in Sergipe State formed Cluster II (blue), which was distinct from Cluster I (red), with accessions collected in the States of Ceará, Rio Grande do Norte, and Pernambuco. The *L. gracilis* accessions represented Cluster III (green).

Accession LSID301 was the most divergent in Cluster I of *L. sidoides* according to the dendrogram. Accessions LSID002 and LSID006 were identified as clones of the loci accessed (Figure 6). Accessions LSID111 and LSID112 were the most similar, and accessions LSID102 and LSID105 were the most divergent in Cluster II. Accessions LGRA109 and LGRA202 were the most similar, and accessions LGRA107 and LGRA110 were the most divergent in Cluster III (Figure 6).

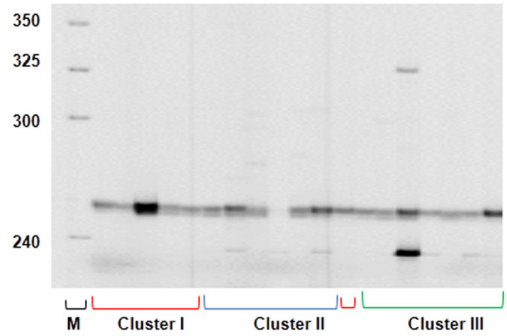

Figure 5. Amplification pattern obtained for the microsatellite locus LAD03 generated using the automated LI-COR Biosciences sequencer. *Lane M* represents the standard ladder size, Cluster I (red) and II (blue) represent the 12 *Lippia sidoides* accessions, and Cluster III (green) represents the 7 *L. gracilis* accessions.

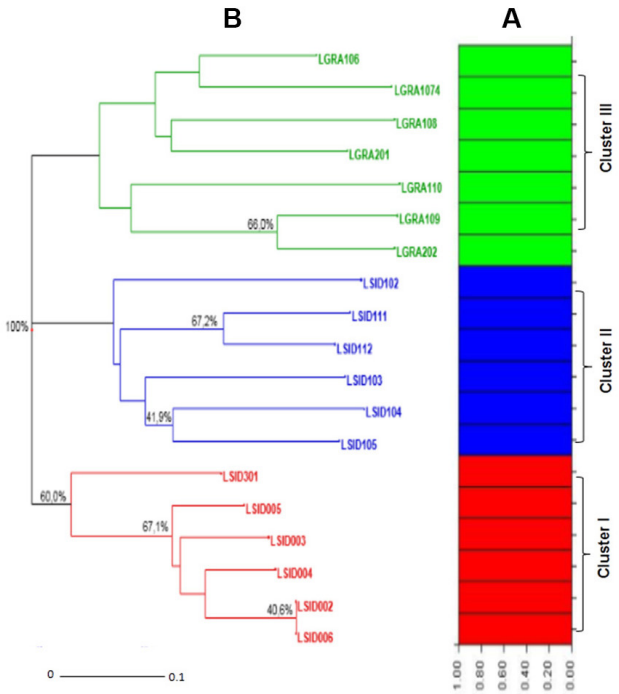

Figure 6. Genetic structure pattern obtained for the 12 *Lippia sidoides* and 7 *L. gracilis* accessions using 10 microsatellite markers. **A.** Plot generated using the STRUCTURE software according to Bayesian statistics. **B.** Dendrogram based on the neighbor-joining clustering criterion. Clusters I and II represent the 12 *L. sidoides* accessions and Cluster III represents the 7 *L. gracilis* accessions.

The divergence between Clusters I and II of the *L. sidoides* accessions ranged from 0.50 to 0.59. Accessions LSID002 and LSID006 were the most divergent from LSID103 of Cluster II (0.59). The accessions of Clusters I and II of *L. sidoides* diverged from the *L. gracilis* accessions, ranging from 0.408 to 0.621 and from 0.524 to 0.741, respectively (Table 7).

Table 7. Genetic distance matrix based on the Rogers' distance modified by Wright generated using 10 SSR markers for the *Lippia sidoides* and *L. gracilis* accessions from the Active Germplasm Bank of the Universidade Federal de Sergipe.

Accessions	LSID002	LSID003	LSID004	LSID005	LSID006	LSID102	LSID103	LSID104	LSID105	LSID111	LSDI112	LSID301	LGRA106	LGRA107	LGRA108	LGRA109	LGRA110	LGRA201	LGRA202
LSID002	0.0000																		
LSID003	0.1581	0.0000																	
LSID004	0.1581	0.2236	0.0000																
LSID005	0.2236	0.1581	0.1581	0.0000															
LSID006	0.0000	0.1581	0.1581	0.2236	0.0000														
LSID102	0.5701	0.5477	0.5477	0.5244	0.5701	0.0000													
LSID103	0.5916	0.5701	0.5701	0.5477	0.5916	0.4743	0.0000												
LSID104	0.5916	0.5701	0.5701	0.5477	0.5916	0.4743	0.4472	0.0000											
LSID105	0.5701	0.5477	0.5477	0.5244	0.5701	0.5000	0.3536	0.3536	0.0000										
LSID111	0.5270	0.5528	0.5000	0.5270	0.5270	0.4082	0.4410	0.4410	0.4714	0.0000									
LSID112	0.5000	0.5244	0.4743	0.5000	0.5000	0.5244	0.4472	0.4472	0.4183	0.2357	0.0000								
LSID301	0.3953	0.3536	0.3536	0.3062	0.3953	0.5303	0.4677	0.4677	0.5000	0.5000	0.5000	0.0000							
LGRA106	0.4714	0.4410	0.4410	0.4082	0.4714	0.6667	0.5528	0.6455	0.5270	0.6236	0.5774	0.5000	0.0000						
LGRA107	0.6009	0.5270	0.5774	0.5000	0.6009	0.6667	0.6455	0.7265	0.6236	0.7500	0.7071	0.5976	0.3062	0.0000					
LGRA108	0.6124	0.5477	0.5916	0.5244	0.6124	0.6519	0.6325	0.7071	0.6124	0.7071	0.6892	0.6124	0.4082	0.4082	0.0000				
LGRA109	0.5244	0.5000	0.5477	0.5244	0.5244	0.7416	0.6892	0.6892	0.6708	0.7265	0.6519	0.5590	0.4410	0.4410	0.5916	0.0000			
LGRA110	0.6009	0.6009	0.5774	0.5774	0.6009	0.6236	0.6872	0.6872	0.6667	0.6455	0.6455	0.5669	0.5774	0.6847	0.6236	0.5528	0.0000		
LGRA201	0.5863	0.5590	0.5590	0.5303	0.5863	0.6124	0.5863	0.6847	0.5590	0.6374	0.6124	0.5000	0.3536	0.4677	0.3953	0.5863	0.5303	0.0000	
LGRA202	0.6009	0.5774	0.5774	0.5528	0.6009	0.7265	0.6667	0.6667	0.6455	0.6872	0.6455	0.5669	0.4410	0.4330	0.5774	0.2357	0.5000	0.5303	0.0000

LSID represents the *L. sidoides* accessions and LGRA represents the *L. gracilis* accessions.

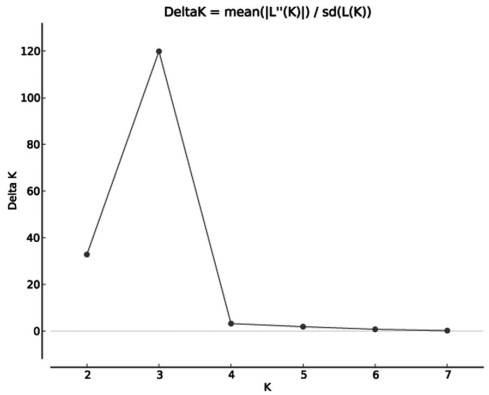

Figure 7. ΔK values obtained using the STRUCTURE HARVESTER based on combinations of microsatellite markers. ΔK = 3 was the most likely value, suggesting the presence of three different clusters for the 12 *Lippia sidoides* accessions and 7 *L. gracilis* accessions.

PCA was also performed using 10 SSR loci. The first two principal components accounted for 51.82% of the total variation found according to axes 1 and 2 (Figure 8). This analysis permitted the separation of the accessions into three different clusters, as observed using the AFLP markers, corroborating the results presented thus far.

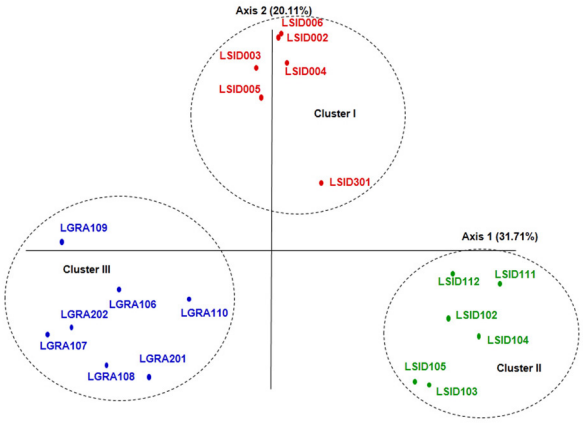

Figure 8. PCA of the 12 *Lippia sidoides* and 7 *L. gracilis* accessions using simple sequence repeat (SSR) markers. Clusters I (red) and II (blue) represent the 12 *L. sidoides* accessions and Cluster III (green) represents the 7 *L. gracilis* accessions.

Analyses of genetic diversity for the clusters of *L. sidoides*, *L. gracilis*, and *Lippia* sp accessions collected are outlined in Table 8. A total of 23, 22 and 36 alleles were identified, averaging 2.5, 2.4, and 3.27 alleles per locus, respectively. The average PICs were moderately informative, with values of 0.39, 0.35, and 0.34, respectively. The mean H_E was 0.49, 0.47, and 0.47 while the mean H_O was 0.48, 0.43, and 0.38 for the *L. sidoides*, *L. gracilis*, and *Lippia* sp accessions, respectively. G_{ST} ranged from 0.23 to 0.35 for the accessions conserved in the AGB and the accessions of the AGB combined with the collected accessions.

Table 8. Estimates of mean diversity for all microsatellite loci assessed in the Active Germplasm Bank accessions (*Lippia sidoides* and *L. gracilis*) alone and in combination with previously assessed collected *Lippia* sp accessions.

Population	Accessions	PL	N_A	Mean N_A	PIC	H_O	H_E	G_{ST}
Lippia sidoides	12	9	23	2.50	0.39	0.48	0.49	-
Lippia gracilis	7	9	22	2.40	0.35	0.43	0.47	-
Lippia sp	40	11	36	3.27	0.34	0.38	0.47	-
Lippia sidoides + *Lippia gracilis*	19	10	-	-	-	-	-	0.23
AGB + collected samples	59	11	-	-	-	-	-	0.35

Number of polymorphic loci (PL), number of alleles (N_A), polymorphic information content (PIC), observed heterozygosity (H_O), expected heterozygosity (H_E), Nei's differentiation index (G_{ST}).

The distribution of allele frequencies indicated the presence of private, rare alleles within the species and accession clusters. For *L. sidoides*, nine alleles (39.13%) were private, including four found in Cluster I, and five in Cluster II (Figure 9). For *L. gracilis*, 11 alleles (50%) were private. For the collected *Lippia* sp plants, termed LG plants, five alleles were private (Figure 9).

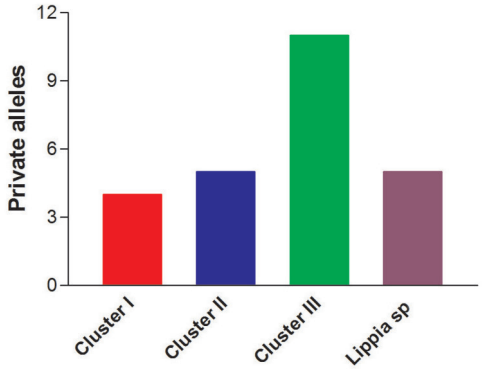

Figure 9. Number of private alleles of *Lippia sidoides* (Clusters I and II) and *L. gracilis* (Cluster III) observed in 10 polymorphic microsatellite loci.

The statistical package PowerCore was used to identify accessions encompassing 100% of the allelic diversity for each cluster studied based on the polymorphic loci. The *L. sidoides* collection of the AGB may be represented by six of its 12 accessions, LSID104, LSID006, LSID103, LSID102, LSID112, and LSID301. Conversely, the *L. gracilis* collection of the AGB was represented by accessions LGRA108, LGRA110, LGRA202, LGRA109, and LGRA201. Of the 40 collected plants, the following may represent the set: LG33, LG29, LG26, LG05, LG27, LG22, LG25, LS06, LG18, and LG19.

DISCUSSION

The 19 accessions of *L. sidoides* and *L. gracilis* composing the AGB were separated into three clusters with likely genetic structure. Similar results were found using AFLP and SSR markers. The results observed in the STRUCTURE plot, dendrogram, and PCA corroborate the existence of three genetically diverged clusters, including two different clusters represented

by *L. sidoides* accessions. These are different analyses, which obtained similar results with both types of markers used. Together, these analyses strengthen the reliability of the presented results.

The reliability of clustering may be assessed using the bootstrap values of the nodes obtained. The bootstrap value of the node separating the three clusters here was the maximum value of 100, suggesting that 100% genetic similarity exists between Clusters I, II, and III, which was confirmed with the AFLP and SSR markers (Figure 3). The bootstrap values of the AFLP dendrogram were high and consistent, primarily because of the high number of loci accessed. This most likely allowed for wide coverage, without previous knowledge thereof, providing more reliable and robust results, which is typical of AFLP markers (Meudt and Clarke, 2007).

The high intra- and inter-species genetic differentiation index observed with the SSR markers (G_{ST} = 0.23), and especially with the AFLP markers (G_{ST} = 0.51 and F_{ST} = 0.54), may be explained by the reproductive system of *L. sidoide*s and *L. gracilis*, which is predominantly vegetative propagation, as their seeds rarely germinate (Matos and Oliveira, 1998). Minimal gene flow and recombination, plus the accumulation of mutations in somatic cells of accessions maintained in germplasm banks, contribute to the high value of differentiation observed between groups. Accordingly, in the absence of recombination, mutations are maintained and accumulate, subsequently increasing the genetic divergence between genotype clusters (Allendorf and Luikart, 2007).

Recently, our research group performed a morphological characterization of *L. sidoides* accessions, and different characteristics were identified among clusters (Fontes SS and Blank AF, unpublished data). The accessions of Cluster I of *L. sidoides* have light brown stems, whereas the accessions of Cluster II have brown stems. Furthermore, Cluster II accessions have rounder leaves and a smaller leaf length/width ratio than leaves of Cluster I.

A set of 18 RAPD markers provided 490 fragments and permitted the analysis of inter- and intra-species variability in nine species of the genus *Lippia* (Viccini et al., 2004). Those authors highlighted *L. sidoides* because it shows the highest intra-species variability among the studied species. The seven *L. gracilis* accessions were previously characterized using 14 RAPD primers. High genetic variability was observed with 75% dissimilarity between the LGRA108 and LGRA110 accessions (Pinto et al., 2011). Species of the genus *Lippia* typically show high genetic variability, but the mechanisms that enable species of this genus to maintain high intra-species variability are not clear, although the fact that some species avoid endogamy is interesting (Viccini et al., 2004).

Genetic divergence between genotypes of Clusters I and II of *L. sidoides* was observed using morphological traits and confirmed using AFLP and SSR markers. Thus, the hypothesis that Cluster II belongs to a species other than *L. sidoides* (Cluster I) or that these species represent different genotypes is proposed. In the case of the latter hypothesis, this cluster is extremely important because it contains private alleles and characteristics that are quite different to the other genotype characteristics of the species. Regardless, future studies, including cytogenetic, leaf anatomy, and genomic investigations, should be conducted to assist in the accurate taxonomic classification of these materials.

Analyses of molecular variance suggest that higher intra-species variability is observed (twice the variation found between the two species) when the species are analyzed as two separate populations. However, higher variability between the two species would be expected. This finding may be explained by *L. sidoides* having two clusters with genetic structures that are distinct from the cluster formed by the *L. gracilis* accessions. Higher variation between clusters was observed under the conditions of the second AMOVA. Comparison with the

previous AMOVA indicated that all three clusters exhibit high genetic divergence and likely genetic structure. An overview of the results of AMOVA suggested that most of the genetic variation observed results from Clusters I, II, and III, and not only between species. The set of accessions contained in the AGB also formed clusters that are divergent from each other, and considerable genetic variability is maintained within the clusters.

Our results indicate high rates of polymorphism at the species and cluster levels, except for Cluster I of *L. sidoides*. The *L. gracilis* genotypes were previously characterized using RAPD markers, and high rates of polymorphism were observed (69.5%; Pinto et al., 2011), although they were lower than the rates observed for *L. sidoides* (78.83%) and *L. gracilis* (87.45%) using AFLP markers. Values similar to those assessed in the present study were reported for *L. alba* accessions using RAPD (86.8%) and ISSR (82.1%) markers (Manica-Cattani et al., 2009).

Three combinations of AFLP primers were used for *Helichrysum italicum* (Roth) G. Don ssp *italicum*, which exhibited a low percentage of polymorphic loci (31.66%; Morone-Fortunato et al., 2010); however, according to the authors, a single combination was sufficient to differentiate all genotypes. A polymorphism rate higher than that observed for *L. sidoides* and lower than that observed for *L. gracilis* was observed when a set of rice germplasm was characterized using AFLP markers. These results demonstrate the ability of AFLP markers to differentiate genotypes even when using few combinations and with low polymorphism, demonstrating the reliability and robustness of these markers.

The number of alleles and the heterozygosity values observed using the SSR markers confirmed that the AGB of *L. sidoides* and *L. gracilis* have considerable genetic variability, which underlines the importance of conserving genotypes. Furthermore, those data indicated that the accessions collected increased the genetic differentiation between species. Some of the *Lippia* sp accessions collected were highly divergent, indicating their importance in complementing these collections.

The mean PICs assessed using SSR markers for *L. sidoides*, *L. gracilis*, and the collected accessions were moderately informative, ranging from 0.35 to 0.38. A higher mean PIC value (0.7) was observed when 82 *Ricinus communis* accessions were analyzed (Pecina-Quintero et al., 2013). Nevertheless, these results suggest that the microsatellite loci used efficiently characterized the accessions from the AGB as well as those collected.

The rates of genetic diversity used suggest that Cluster I of *L. sidoides* has low genetic variability because its accessions are genetically similar. However, the presence of private and rare alleles was observed in this cluster, which was found to diverge from Cluster II, underlining the importance of conserving these accessions. When characterized at the species level, the AGB exhibited intermediate genetic diversity.

The SSR markers were useful for selecting the most representative genotypes of the AGB, providing information on the structure of a future core collection. They were also useful for selecting the collected accessions, LG33, LG29, LG26, LG05, LG27, LG22, LG25, LS06, LG18, and LG19, to complement the *L. sidoides* and *L. gracilis* AGB. In fact, the SSR markers represent an adequate tool to help establish, characterize, and expand germplasm banks.

AFLP and SSR markers were shown to be highly efficient tools for assessing the inter- and intra-species genetic variability, because *L. sidoides* exhibited two clusters that were different from each other and from *L. gracilis*. They were similarly clustered by both markers, confirming the intermediate genetic variability of the genotypes of these species conserved in the AGB.

The consistency of the results when assessed using AFLP and SSR markers was also confirmed by the high correlation of the matrices, $r = 0.78$ ($P \geq 0.001$), as shown by the

Mantel test. The high correlation observed using the AFLP and SSR markers indicates that they were equally informative when characterizing these accessions. Although each marker has its own peculiarities, when used in combination, the results were consistent. Furthermore, these markers enabled greater coverage of the genome, rendering the results more robust and reliable. High correlation values (ranging from 0.83 to 0.95) were also observed between the ITS, RAPD, and ISSR markers (Adams et al., 2003).

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors thank the Sergipe State Research and Technological Innovation Foundation (Fundação de Apoio à Pesquisa e à Inovação Tecnológica do Estado de Sergipe - FAPITEC/SE), the National Counsel for Technological and Scientific Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq) and the Coordination for the Improvement of Higher Education Personnel (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES) for funding this study.

REFERENCES

Adams RP, Schwarzbach AE and Pandey RN (2003). The concordance of terpenoid, ISSR and RAPD markers, and ITS sequence data sets among genotypes: an example from *Juniperus*. *Biochem. Syst. Ecol.* 31: 375-387. http://dx.doi.org/10.1016/S0305-1978(02)00157-6

Allendorf FW and Luikart G (2007). Conservation and Genetics of Populations. 1st edn. Wiley-Blackwell Publishing, New Jersey.

Botelho MA, Nogueira NAP, Bastos GM, Fonseca SGC, et al. (2007). Antimicrobial activity of the essential oil from *Lippia sidoides*, carvacrol and thymol against oral pathogens. *Braz. J. Med. Biol. Res.* 40: 349-356. http://dx.doi.org/10.1590/S0100-879X2007000300010

Carvalho AFU, Melo VMM, Craveiro AA, Machado MIL, et al. (2003). Larvicidal activity of the essential oil from *Lippia sidoides* Cham. against *Aedes aegypti* linn. *Mem. Inst. Oswaldo Cruz* 98: 569-571. http://dx.doi.org/10.1590/S0074-02762003000400027

Cavalcanti SCH, Niculau E dos S, Blank AF, Câmara CAG, et al. (2010). Composition and acaricidal activity of *Lippia sidoides* essential oil against two-spotted spider mite (*Tetranychus urticae* Koch). *Bioresour. Technol.* 101: 829-832. http://dx.doi.org/10.1016/j.biortech.2009.08.053

Coelho ASG (2000). BOOD: Avaliação de dendrogramas baseados em estimativas de distâncias/similaridades genéticas através do procedimento de bootstrap. Goiânia, UFG.

Cruz EMO, Costa Jr LM, Pinto JAO, Santos DdeA, et al. (2013). Acaricidal activity of *Lippia gracilis* essential oil and its major constituents on the tick *Rhipicephalus (Boophilus) microplus*. *Vet. Parasitol.* 195: 198-202. http://dx.doi.org/10.1016/j.vetpar.2012.12.046

Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Evanno G, Regnaut S and Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620. http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier L, Smouse PE and Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Excoffier L, Laval G and Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47-50.

Goedert CO (2007). Recursos genéticos vegetais. In: Histórico e avanços em recursos genéticos no Brasil. 1st edn. Embrapa Recursos Genéticos e Biotecnologia, Brasília, 23-60.

Goudet J (1995). FSTAT version 1.2: a computer program to calculate F-statistics. *J. Hered.* 86: 485-486.

Guimarães CT, Magalhães JV, Lanza MA and Schuster I (2009). Marcadores moleculares e suas aplicações no melhoramento genético. *Informe Agropec.* 30: 24-33.

Kuleung C, Baenziger PS and Dweikat I (2004). Transferability of SSR markers among wheat, rye, and triticale. *Theor. Appl. Genet.* 108: 1147-1150. http://dx.doi.org/10.1007/s00122-003-1532-5

Kumar P, Gupta VK, Misra AK, Modi DR, et al. (2009). Potential of molecular markers in plant biotechnology. *Plant Omics* 2: 141-162.

Lewis P and Zaykin D (1999). Genetics Date Analysis: computer program for the analysis of allelic date (software), Version 1.0. Available at [http://alleyn.eeb.uconn.edu/gda/]. Accessed August 11, 2013.

Lorenzi H and Matos FJA (2002). Plantas medicinais no Brasil: nativas e exóticas. 1st edn. Instituto Plantarum de Estudos da Flora, Nova Odessa.

Manica-Cattani MF, Zacaria J, Pauletti G, Atti-Serafini L, et al. (2009). Genetic variation among South Brazilian accessions of *Lippia alba* Mill. (Verbenaceae) detected by ISSR and RAPD markers. *Braz. J. Biol.* 69: 375-380. http://dx.doi.org/10.1590/S1519-69842009000200020

Matos FJA and Oliveira F (1998). *Lippia sidoides* Scham - farmacognosia, química e farmacologia. *Rev. Bras. Farm.* 79: 84-87.

Meudt HM and Clarke AC (2007). Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends Plant Sci.* 12: 106-117. http://dx.doi.org/10.1016/j.tplants.2007.02.001

Miller MP (1997). Tools for Population Genetic Analysis (TFPGA), Version 1.3: A windows program for the analysis of allozyme and molecular population data. Available at [http://www.ccg.unam.mx/~vinuesa/.../TFPGADOC.PDF]. Accessed August 9, 2013.

Morone-Fortunato I, Montemurro C, Ruta C, Perrine R, et al. (2010). Essential oils, genetic relationships and *in vitro* establishment of *Helichrysum italicum* (Roth) G. Don ssp. *italicum* from wild Mediterranean germplasm. *Ind. Crops Prod.* 32: 639-649. http://dx.doi.org/10.1016/j.indcrop.2010.07.023

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

Park S (2001). MStools V.3 (Excel Spreadsheet Toolkit for Data Conversion). Smurfit Institute of Genetics, Trinity College, Dublin.

Pecina-Quintero V, Anaya-Lopéz JL, Núñez-Colín CA, Zamarripa-Colmenero A, et al. (2013). Assessing the genetic diversity of castor bean from Chiapas, México using SSR and AFLP markers. *Ind. Crops Prod.* 41: 134-143. http://dx.doi.org/10.1016/j.indcrop.2012.04.033

Perrier X and Jacquermoud-Collet JP (2006). DARwin software. Available at [http://darwin.cirad.fr/darwin]. Accessed August 10, 2013.

Pessoa ODL, de Carvalho CBM, Silvestre JOVL, Lima MCL, et al. (2005). Antibacterial activity of the essential oil from *Lippia* aff. *gracillis*. *Fitoterapia* 76: 712-714. http://dx.doi.org/10.1016/j.fitote.2005.07.004

Pinto JAO, Blank AF, Cruz EMO, Goes IB, et al. (2011). Caracterização molecular (RAPD) de acessos de alecrim-de-tabuleiro (*Lippia gracilis* Schauer). *Sci. Plena* 7: 1-6.

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

Rocha DS, Santos CP, Bajay MM, Campos JB, et al. (2015). Development of a novel set of microsatellite markers for *Lippia alba* (Verbenaceae). *Genet. Mol. Res.* 14: 971-974. http://dx.doi.org/10.4238/2015.February.3.4

Rohlf FJ (2000). NTSYSpc: Numerical taxonomy and multivariate analysis system. Version 2.1. Department of Ecology and Evolution, State University of New York, New York.

Santos CP, Rocha DS, Bajay MM, Santos FRC, et al. (2014). Cross-species transferability of microsatellite markers in the genus *Lippia*. *Genet. Mol. Res.* 13: 9846-9850. http://dx.doi.org/10.4238/2014.November.27.11

Santos FRC, Lima PF, Priolli RH, Siqueira WJ, et al. (2012). Isolation and characteristics of eight novel polymorphic microsatellite loci in *Lippia alba* (Verbenaceae). *Am. J. Bot.* 99: e301-e303. http://dx.doi.org/10.3732/ajb.1100578

Silva WJ, Dória GAA, Maia RT, Nunes RS, et al. (2008). Effects of essential oils on *Aedes aegypti* larvae: alternatives to environmentally safe insecticides. *Bioresour. Technol.* 99: 3251-3255. http://dx.doi.org/10.1016/j.biortech.2007.05.064

Viccini LF, Costa DCS, Machado MA and Campos AL (2004). Genetic diversity among nine species of *Lippia* (Verbenaceae) based on RAPD markers. *Plant Syst. Evol.* 246: 1-8. http://dx.doi.org/10.1007/s00606-003-0079-x

Viccini LF, Pierre PMO, Praça MM, Costa DCS, et al. (2006). Chromosome numbers in the genus *Lippia* (Verbenaceae). *Plant Syst. Evol.* 256: 171-178. http://dx.doi.org/10.1007/s00606-005-0351-3

Vos P, Hogers R, Bleeker M, Reijans M, et al. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23: 4407-4414. http://dx.doi.org/10.1093/nar/23.21.4407

Wright S (1978). Evolution and the genetics of populations: variability within and among natural populations. 4th edn. University of Chicago Press, Chicago.

Yeh F, Yang RC, Mao J, Ye Z, et al. (1999). Popgene, the user-friendly shareware for Population Genetic Analysis. University of Alberta, Edmonton, Alberta.