Molecular Identification and Phylogenetic Analysis of *Egletes viscosa* (L.) Less., a Traditional Medicinal Plant from Northeastern Brazil

José M. M. Batista,* Davi C. Torres,*,a Daniel M. de M. Jorge,* Carlos E. A. Soares, b Walderly M. Bezerra, a Edilberto R. Silveira c and Thalles B. Grangeiro a

Departamento de Biologia and Departamento de Química Orgânica e Inorgânica, Centro de Ciências, Universidade Federal do Ceará (UFC), Fortaleza-CE, Brazil

Unidade Acadêmica de Ciências Biológicas, Universidade Federal de Campina Grande (UFCG), Patos-PB, Brazil

Macela (*Egletes viscosa* (L.) Less.) é uma erva medicinal utilizada para tratar distúrbios digestivos e intestinais. Estudos prévios revelaram a existência de dois quimiotipos de macela, cujos óleos essenciais são caracterizados pela presença de acetato de *trans*-pinocarveíla (quimiotipo A) ou acetato de *cis*-isopinocarveíla (quimiotipo B). A análise de sequências de DNA do espaçador transcrito interno (ITS/5.8S) revelou oito polimorfismos de base única (SNPs) e um sítio inserção/deleção que diferenciam os dois quimiotipos. Além disso, foi possível identificar espécimes selvagens de *E. viscosa* pertencendo a ambos quimiotipos e amostras comerciais, obtidas em mercados locais, pertencendo somente ao quimiotipo A. A análise filogenética confirmou a posição taxonômica desses dois quimiotipos na tribo Astereae. Portanto, a análise de sequências de DNA da região ITS/5.8S proporciona uma efetiva e acurada estratégia para identificar os quimiotipos A e B de *E. viscosa*.

“Macela” (*Egletes viscosa* (L.) Less.) is a medicinal herb used to treat digestive and intestinal problems. Previous studies revealed the existence of two chemotypes of “macela”, whose essential oils are characterized by the presence of *trans*-pinocarveyl acetate (chemotype A) or *cis*-isopinocarveyl acetate (chemotype B). Herein, DNA sequence analysis of the internal transcribed spacer region (ITS/5.8S) showed eight single nucleotide polymorphisms (SNPs) and one insertion/deletion site that clearly distinguished the chemotypes of *E. viscosa*. Moreover, it was possible to identify wild specimens of *E. viscosa* belonging to both chemotypes and commercial samples, purchased at local markets, belonging only to chemotype A. The phylogenetic analysis confirmed their taxonomic position within the tribe Astereae. Therefore, analysis of ITS/5.8S DNA sequences provides an effective and accurate strategy for identifying the chemotypes A and B from *E. viscosa*.

**Keywords**: phylogeny, ITS/5.8S, *Egletes viscosa*, Asteraceae, “macela”

**Introduction**

Asteraceae (Compositae) is the largest family of flowering plants, comprising nearly 25,000 species in approximately 1,620 genera, of which 180 are found in Brazil.² Erect tropical daisy, *Egletes viscosa* (L.) Less., is a small annual herb mostly native to intertropical Americas. In the Northeast of Brazil, it easily grows in local flood plains near the banks of small ponds and streams.³ In this region, it is popularly known as “macela” or “marcela”, and sometimes as “macela-da-terra”, to differentiate it from *Achyrocline satureioides* (Lam.) D.C., another Asteraceae species common in the South and Southeast of Brazil, where it is also known as “macela” or “marcela”.

Dried flower buds of *E. viscosa* have been widely used in Brazilian traditional medicine for the treatment of digestive and intestinal disorders. Several biologically active flavonoids and terpenoids have been reported in this plant. Pharmacological investigations have revealed that *E. viscosa* essential oil has antinociceptive, anticonvulsant and antibacterial properties,³ whereas its flavonoid compounds possess anti-inflammatory, antianaphylactic, antithrombotic, hepatoprotective, gastroprotective, uroprotective, antidiarrhoeal and moderate anti-HIV...
activities. Moreover, E. viscosa diterpenoids have been reported to have antispasmodic, anti-inflammatory, anti-edematogenic, gastroprotective and antinociceptive effects, as well as antiproliferative action on cultured cells.

More recently, extensive analyses of the essential oils from flower buds of specimens sampled in different localities in the state of Ceará, northeastern Brazil, have demonstrated the existence of two E. viscosa chemotypes. These chemotypes have been designated as “macela” A and “macela” B, whose essential oils are characterized by trans-pinocarveyl acetate and cis-isopinocarveyl acetate as their major constituents, respectively. Furthermore, lyophilized teas from flower buds of both chemotypes also showed different chemical profiles. The trans-pinocarveyl chemotype (“macela” A) gave ternatin, centipedic acid, 12-acetoxy-hawtriwaic lactone and the new diterpene 12-acetoxy-7-hydroxy-3,13(14)-clerodandien-18,19:15,16-diolide. From the cis-isopinocarveyl chemotype (“macela” B) the same diterpene, 12-epi-bacchotricuneatin and scopoletin were isolated, in addition to ternatin. These two E. viscosa chemotypes also have distinct morphological features. “Macela” A presents pinnatifid and broad leaves with a less distinct petiole, while “macela” B leaves are much more jagged, resembling secondary leaves, with a prominent petiole, which is a very characteristic feature of the Asteraceae family.

The rational use of the “macela” medicinal herb by the population may be compromised due to difficulties in distinguishing the species and their chemotypes. The accurate identification of this herb is a pre-requisite for quality control and conservation procedures. Moreover, an effective method to distinguish them is important for the healthy development of the herbal industry. Although the leaves and petioles of the E. viscosa chemotypes are clearly different, the dried plants usually sold in herbal stores are very similar in morphological appearance. Furthermore, many commercial E. viscosa products are in the form of dried flower buds, rendering their authentication by traditional methods very difficult.

Nowadays, DNA polymorphism-based assays have been developed for the molecular identification of medicinal plants. In this study, we have explored the sequencing of the internal transcribed spacer region (ITS/5.8S) of the nuclear ribosomal DNA (nrDNA) to classify E. viscosa chemotypes from the state of Ceará, northeastern Brazil, and also to determine the phylogenetic relationships of this taxon within the family Asteraceae. In addition, we have also sequenced commercial “macela” samples in order to identify the chemotypes from E. viscosa sold in the common market. Herein, we have successfully applied a well known DNA based technique for identifying two E. viscosa chemotypes.

**Experimental**

**Plant materials**

Several domesticated E. viscosa specimens from the Laboratório de Análise de Sementes, Universidade Federal do Ceará (UFC), Fortaleza, Ceará, Brazil, were grown in greenhouses. These specimens were from chemotype A (n = 9) and chemotype B (n = 10). Moreover, dried flower buds of “macela” (n = 4) were purchased from local herbal stores in Fortaleza. To protect the manufacturers’ identities, the sample sources are labeled CM1-4. We also obtained fresh leaves from wild specimens of “macela” (n = 6), previously identified as E. viscosa, from distinct regions of the state of Ceará. Voucher specimens are deposited in the Herbarium Prisco Bezerra, UFC (Água Verde - voucher EAC 37740, Chorozinho - EAC 34763, Crato - EAC 34743 and Várzea Alegre - EAC 34744).

**DNA extraction**

Total genomic DNA was extracted from fresh plant (leaf) material or dried flower buds. The samples (0.6 g) were ground in liquid nitrogen and then incubated for 1 h at 60 °C in CTAB extraction buffer (2% m/v CTAB, 100 mmol L⁻¹ Tris-HCl, pH 8.0, 20 mmol L⁻¹ EDTA, 1.4 mol L⁻¹ NaCl, 0.2% v/v 2-mercaptoethanol). Further processing of the samples was done as previously described. DNA concentration was determined by measuring the absorbance at 260 nm (A₅₀₀) of a ten-fold dilution of each sample. The quality of all DNA preparations was checked by 0.8% agarose gel electrophoresis.

**PCR amplification and DNA sequencing**

Amplification of the ITS/5.8S region was done by PCR using the primers ITS4 (5′-TCCTCCGCTTATGATATGC-3′) and ITS5 (5′-GCAAGTAAAAGTCGTAACAGG-3′). Amplification reactions were performed in a final volume of 25 µL containing 800-1000 ng of genomic DNA (template); 20 mmol L⁻¹ Tris-HCl, pH 8.4; 50 mmol L⁻¹ KCl; 1.5 mmol L⁻¹ MgCl₂; 100 mmol L⁻¹ of each dATP, dCTP, dGTP, and dTTP (GE Healthcare Life Sciences, Piscataway, NJ, USA); 12.5 pmol of each primer; and 0.5 units of Taq DNA polymerase (GE Healthcare Life Sciences). PCR reactions were carried out in a MJ-research (Watertown, MD, USA) PTC-200 thermocycler. The cycling parameters were an initial denaturation step of 95 °C for 3 min, followed by 35 cycles of 95 °C for 1 min, 50 °C for 1 min, 72 °C for 1 min, and a final incubation step of 10 min at
72 °C. PCR products were then stored at 4 °C until used. Control samples containing all reaction components except DNA were always used to test that no self-amplification or DNA contamination occurred.

Once the specificity of the amplifications was confirmed by 1% agarose gel electrophoresis, PCR products were purified from the remaining reactions using GFX PCR DNA and gel band purification kit (GE Healthcare Life Sciences). DNA sequencing was performed with the DYEnamic ET terminators cycle sequencing kit (GE Healthcare Life Sciences), following the protocol supplied by the manufacturer. Sequencing reactions were analyzed in a MegaBACE 1000 automatic sequencer (GE Healthcare Life Sciences). Each PCR product was sequenced four times in both directions using the same primers employed in the amplification reaction. The nucleotide sequence data reported in this paper are available in the GenBank database (http://www.ncbi.nlm.nih.gov/genbank) under the accession numbers HQ184420-HQ184431.

Sequence alignment and phylogenetic analyses

The quality of the DNA sequences was checked and overlapping fragments were assembled using the Phred/Phrap/Consed package. BLASTn searches were conducted in GenBank to detect potential contaminant sequences. Any uncertain base positions, generally located close to priming sites, were excluded.

High quality (phred > 20) assembled sequences were aligned using ClustalX version 2.0.9,18 with default gap penalties, and manually corrected using the software BioEdit version 7.0.9.19 to produce an alignment with the fewest number of changes (indels or nucleotide substitutions). Alignment file is available upon request to the corresponding author.

Phylogenetic analyses were performed in PAUP* version 4.0b1020 and MrBayes 3.1.2.21 In addition to the generated sequences, we included several others from Asteroidae taxa, which were retrieved from the GenBank database. Barnadesioidae taxa (Barnadesia odorata Griseb., Dasyphyllum popayanense (Hieron.) Cabrera and Huarpea andina Cabrera) were chosen as outgroup based on previous data,22 which strongly supported this subfamily as sister to all other major lineages of the Asteraceae family. Imported DNA sequences are listed in the Supplementary Information with the taxonomic classification and corresponding GenBank accession numbers.

Maximum parsimony (MP) analysis was conducted using heuristic search with tree-bisection-reconnection (TBR) branch-swapping, ACCTRAN character optimization, and the Multtrees option in effect, holding a maximum of ten most parsimonious trees per replicate of 500 random addition replicates, in an attempt to sample multiple islands of most parsimonious trees. A maximum of 10,000 trees was allowed to accumulate, which is sufficient to capture topological variation.23 Characters were weighted equally and their state changes were treated as unordered. Indels were treated as missing data. Bootstrap support (BS) values for the optimal trees were calculated using 1,000 replicates with heuristic search settings identical to those applied in the original search.

The most suitable model for the Bayesian inference (BI) was selected using the Akaike Information Criterion (AIC) by MrModeltest, version 2.3,24 which presents several important advantages over other strategies of model selection.25 Two independent analyses with five million generations were run to estimate parameters relating to sequence evolution and likelihood probabilities using a Markov chain Monte Carlo (MCMC) method. Trees were collected every 100th generation. After removing 25% of the generations as burn in, a 50% majority rule consensus tree was calculated to generate a posterior probability (PP) value for each node. The trees were visualized by TreeView.26

Results and Discussion

Sequence variation

The complete ITS/5.8S region (ITS1+5.8S+ITS2) of the nrDNA was amplified by PCR from 19 E. viscosa cultivated specimens, comprising 9 specimens from chemotype A (“macela” A) and 10 specimens from chemotype B (“macela” B). The sequences of the amplified DNA fragments were determined and deposited in the GenBank database (accession numbers HQ184420-HQ184431). There was no nucleotide variation within each chemotype specimens sequenced. The length of the ITS/5.8S region was 627 bp for “macela” A and 628 bp for “macela” B. Strikingly, ITS/5.8S sequence analysis allowed the identification of polymorphic sites that could discriminate the sequences from the two chemotypes. The identification was based on nine diagnostic sites, comprising eight single nucleotide substitutions (four in the ITS1, one in the 5.8S region and three in the ITS2 sequence) and one site containing a 1 bp indel (in the ITS1), as highlighted in the multiple sequence alignment depicted in Figure 1.

Characterization of wild-type accessions and commercial samples

Complete ITS/5.8S sequences were also determined from wild specimens of E. viscosa collected in different
Figure 1. Multiple alignment of ITS/5.8S nrDNA sequences, comprising 628 sites, from *E. viscosa* chemotypes (“macela” A and B), commercial samples (CM1–4) and wild-type samples harvested from distinct regions of the state of Ceará, Brazil (Água Verde, Chorozinho, Crato, Irauçuba, São Gonçalo and Várzea Alegre). The light gray color regions highlight the variable sites, while the dark gray delimites the 5.8S nrDNA region.
localities from the state of Ceará, northeastern Brazil. Taking into account the nine diagnostic sites identified in the ITS/5.8S sequences from the cultivated specimens, two wild accessions (from Crato and Irauçuba) were identified as *E. viscosa* chemotype A, whereas the other four sampled accessions (from Água Verde, Chorozinho, São Gonçalo and Várzea Alegre) were classified as *E. viscosa* chemotype B (Figure 1). In addition, commercial samples (CM1-4) were purchased from local herbal stores in Fortaleza (Ceará) and then sequenced. These commercial products are presumed to be flower buds of *E. viscosa*, which are dried and ground to a coarse powder. All commercial samples were identified as *E. viscosa* chemotype A (Figure 1). Moreover, the sequencing electropherograms obtained from these commercial samples showed sharp peaks with little or no background noise (data not shown), which is characteristic of a single template in the sequencing reaction or an excess of a unique template. This observation was taken as evidence that the samples were not adulterated.

**Phylogenetic analysis**

To determine the phylogenetic relationships of *E. viscosa* within the Asteraceae family, the DNA sequences generated in this study were aligned to ITS/5.8S sequences from several distinct species belonging to subfamily Asteroideae. The aligned sequences of the entire ITS/5.8S region had 677 characters, with 235 conserved sites (34.7%), 431 variable sites (63.6%), and 11 sites (1.7%) unique to individual taxa. The MP analysis included 357 parsimony-informative sites and produced three most parsimonious trees of 1647 steps, with a consistency index of 0.479 and a retention index of 0.643. Both MP and BI analyses were mostly congruent. As shown in Figure 2, every major tribe represented within the subfamily Asteroideae was resolved in the MP strict consensus tree. Moreover, the *E. viscosa* chemotypes were placed in the tribe Astereae with robust support (BS and PP of 100 and 1.00, respectively).

Figure 2. Majority rule consensus tree from Maximum Parsimony analysis using the ITS/5.8S nrDNA sequences under study. *Barnadesia odorata*, *Dasyphyllum popayanense* and *Huarea andina* were set as outgroup. Bootstrap values (1,000 replicates) for maximum parsimony analysis are presented above the branches. Posterior probability values are presented under the branches. The major represented subtribes of Asteroideae are assigned on the right.
Discussion

Identification of the origin and breed of medicinal herbs not only ensures the purity of the supplying source, but also serves as a good monitoring tool for production quality. DNA markers have become quite popular regarding the identification of plant species. They are not tissue-specific, thus can be detected at any phase of the organism’s development, and only a small amount of sample is enough for analysis using PCR. The ITS/5.8S region of the nrDNA is widely applied in authentication and phylogenetic analyses of medicinal herbs. 11,27,28 Many by copies of the nrDNA are present in each plant genome and the rRNA genes are highly homologous within individuals and species, due to homologous recombination and gene conversion. 29

In this report, we have shown that the ITS/5.8S region is useful in providing reliable and effective means for the identification of two *E. viscosa* chemotypes. Our results agree with other studies in which ribosomal RNA genes have been successfully used as a tool for molecular authentication. Moreover, we showed the capability of the ITS1 (which exhibited 56% of the polymorphic sites) to be used as a unique marker for authentication and differentiation of plant species, given its greater variability, as it was also previously observed among *Dendrobium officinale* Kimura and Migo specimens. 27

Based on their ITS/5.8S sequences, the wild *E. viscosa* specimens from Crato and Irauçuba were identified as chemotype A. Indeed, previous phytochemical analysis of the same samples used in this study showed that the essential oils extracted from the specimens collected in Crato and Irauçuba were rich in *trans*-pinocarveyl acetate, a characteristic feature of the chemical profile of *E. viscosa* chemotype A. 10

Despite the samples from Água Verde, Chorozinho, São Gonçalo and Várzea Alegre were molecularly identified as *E. viscosa* chemotype B, the standard phytochemical analysis only showed the presence of *cis*-isopinocarveyl acetate, the characteristic constituent for chemotype B, in the essential oils extracted from specimens collected in Água Verde and Chorozinho. 10 Therefore, our results regarding the molecular identification of the São Gonçalo and Várzea Alegre specimens do not corroborate with their expected chemotype patterns. Nevertheless, the samples from Várzea Alegre did not show detectable amounts of *cis*-isopinocarveyl acetate, 10 which could indicate a misidentification of the chemotype pattern. This can also be justified by the fact that the chemical composition of essential oils can vary according to climatic conditions and harvesting, which could have misled the final analysis of their components. Thus, the molecular data presented here support the application of the ITS/5.8S region as an authentication tool for wild specimens of “macela”. Furthermore, it demonstrates the possible use for future biogeographic studies on *E. viscosa* in Brazil, as described for other plants of medicinal interest such as ginseng (*Panax ginseng* C.A. Mey.). 30

All commercial samples of “macela” were identified as *E. viscosa* chemotype A and no evidence of adulterants was found among the analyzed samples. The exclusive identification of chemotype A may be due to seasonal features in the available stocks for sale in the local markets and should not be considered as a predominant pattern. A larger sample size is needed in order to establish which specimen (chemotype A or B) is more frequently commercialized in the local markets of Fortaleza and other localities.

The phylogenetic analyses confirmed the taxonomic position of *E. viscosa* in the tribe Astereae, as previously classified based on morphological data. 2,31 Moreover, the monophyly of Astereae is strongly supported (BS = 100, PP = 1.00), corroborating previous molecular studies. 22,32 Although our phylogenetic analyses recognized 8 lineages at the tribe level, they failed to resolve tribal relationships within Asteroideae subfamily.

Conclusions

Molecular identification and authentication of medicinal plants such as “macela” provide an important tool for guiding pharmacological, phylogenetic, taxonomic and biogeographic studies. Our data suggest that “macela” chemotypes A and B are indeed genetically distinct, which is in agreement with their striking morphological and chemical differences already pointed out in previous studies. 5,33 Thus, the hypothesis that these *E. viscosa* chemotypes may represent distinct botanical varieties or even different species should be further evaluated in future studies.

Supplementary Information

GenBank accession numbers for the sequences used in the present study are available free of charge at http://jbcs.sbq.org.br as PDF file.

Acknowledgments

This work was supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP).
References


Submitted: October 6, 2011
Published online: March 29, 2012