Genetics and Molecular Biology, 34, 2, 244-251 (2011)

www.sbg.org.br

Research Article

Phylogenetic relationships within *Chamaecrista* sect. *Xerocalyx* (Leguminosae, Caesalpinioideae) inferred from the cpDNA *trnE-trnT* intergenic spacer and nrDNA ITS sequences

Davi Coe Torres[1], João Paulo Matos Santos Lima[2], Afrânio Gomes Fernandes[3], Edson Paula Nunes[3] and Thalles Barbosa Grangeiro[1]

[1]*Laboratório de Citogenética e Genética Molecular, Departamento de Biologia, Centro de Ciências, Universidade Federal do Ceará, Fortaleza, CE, Brazil.*
[2]*Departamento de Bioquímica, Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil.*
[3]*Herbário Prisco Bezerra, Departamento de Biologia, Centro de Ciências, Universidade Federal do Ceará, Fortaleza, CE, Brazil.*

Abstract

Chamaecrista belongs to subtribe Cassiinae (Caesalpinioideae), and it comprises over 330 species, divided into six sections. The section *Xerocalyx* has been subjected to a profound taxonomic shuffling over the years. Therefore, we conducted a phylogenetic analysis using a cpDNA *trnE-trnT* intergenic spacer and nrDNA ITS/5.8S sequences from Cassiinae taxa, in an attempt to elucidate the relationships within this section from *Chamaecrista*. The tree topology was congruent between the two data sets studied in which the monophyly of the genus *Chamaecrista* was strongly supported. Our analyses reinforce that new sectional boundaries must be defined in the *Chamaecrista* genus, especially the inclusion of sections *Caliciopsis* and *Xerocalyx* in sect. *Chamaecrista*, considered here paraphyletic. The section *Xerocalyx* was strongly supported as monophyletic; however, the current data did not show *C. ramosa* (microphyllous) and *C. desvauxii* (macrophyllous) and their respective varieties in distinct clades, suggesting that speciation events are still ongoing in these specimens.

Key words: phylogeny, ITS/5.8S, *Chamaecrista*, *Xerocalyx*, *trnE-trnT* intergenic spacer.

Received: December 8, 2010; Accepted: January 26, 2011.

Introduction

The subtribe Cassiinae (Leguminosae; Caesalpinioideae; Cassieae), formerly represented by *Cassia* s.l., is now subdivided into three genera, *Cassia sensu stricto*, *Senna* Mill. and *Chamaecrista* (Breyne) Moench (Irwin and Barneby, 1982). Molecular analyses, including *trnL* intron (Bruneau *et al.*, 2001), *matK/3'-trnK* (Bruneau *et al.*, 2008) and *rbcL* sequences (Doyle *et al.*, 2000; Kajita *et al.*, 2001), and morphological data (Tucker, 1996) suggest that this subtribe is not monophyletic. There is also disagreement about sister-group relationships in Cassiinae. One hypothesis considers *Chamaecrista* a clade distinct from its sister taxa *Senna* and *Cassia* (Bruneau *et al.*, 2001, 2008; Marazzi *et al.*, 2006; de Souza Conceição *et al.*, 2009). However, other studies (Doyle *et al.*, 1997; Kajita *et al.*, 2001; Herendeen *et al.*, 2003) indicate that *Senna* and *Chamaecrista* are sister taxa and that *Cassia* occurs in a distinct clade.

The genus *Chamaecrista*, formerly defined as *Cassia* subgenus *Lasiorhegma* (Irwin and Barneby, 1982), has been divided into six sections of very unequal sizes [*Absus* (Collad.) H.S. Irwin & Barneby, *Apoucouita* (Benth.) H.S. Irwin & Barneby, *Caliciopsis* H.S. Irwin & Barneby, *Chamaecrista* H.S. Irwin & Barneby, *Grimaldia* (Schrank) H.S. Irwin & Barneby and *Xerocalyx* (Benth.) H.S. Irwin & Barneby]. Including trees, shrubs and herbs, *Chamaecrista* comprises approximately 330 species, 266 of which are native to the Americas (Lewis, 2005). It has a significant ecological importance because it is the only genus within Cassiinae with concave extrafloral nectaries and roots bearing bacterial nodules (Irwin and Barneby, 1982).

The section *Xerocalyx* is easily recognizable and is distinguished by its parallel-nerved leaflets, strongly graduated and multistriate sepals and reduced chromosome number of 2n = 14. However, it has been subjected to considerable taxonomic reformulation (Irwin and Barneby,

Send correspondence to Davi Coe Torres. Laboratório de Citogenética e Genética Molecular, Departamento de Biologia, Bloco 906, Centro de Ciências. Universidade Federal do Ceará, Av. Humberto Monte s/n, 60451-970 Fortaleza, CE, Brazil. E-mail: torresdc@gmail.com.

1982). While it was included in the taxon *Cassia*, Irwin (1964) recognized 16 species within *Xerocalyx*, as defined by their morphological, chromosomal and chemical characteristics. Afterwards, based on an arbitrary classification of morphological characters, such as amplitude of foliage and length of petiole, Irwin and Barneby (1982) proposed a profound reorganization within *Xerocalyx*, recognizing only three species with 22 varieties. More recently, employing expressive organographic characteristics and corological aspects, Fernandes and Nunes (2005) rearranged this section into 10 species and 27 varieties.

Furthermore, Irwin and Barneby (1982) moved, with some confidence, several specimens of *C. diphylla* (L.) Greene, previously classified by Irwin (1964), to *C. rotundifolia* (Pers.) Greene. This demonstrates the occasional confusion between the identification of *C. diphylla* and *C. rotundifolia* in herbaria. The two species are similar in number and sometimes also the form of their leaflets, but they can be distinguished by the venation of their leaflets, the presence of petiolar glands, the strongly graduated and multistriate calyx-lobes, and by the decandrous androecium (Irwin and Barneby, 1982). These disagreements concerning the classification of *Xerocalyx* raise the question whether *C. diphylla* could be closer related to *C. rotundifolia* (sect. *Chamaecrista*) than to the other *Xerocalyx* members.

The taxonomic incongruence within *Xerocalyx* remains. Because most studies have been done by analysis of herbarium specimens that could not discern truly discrete units, it is worthwhile employing alternative tools to resolve this taxonomic instability. Recently, using sequence data from nuclear ITS and plastid *trnL-F* DNA spacers and representatives of all six sections of *Chamaecrista*, de Souza Conceição *et al.* (2009) analyzed the phylogeny of this genus and supported the monophyly of sect. *Xerocalyx*. However, the phylogenetic relationships within *Xerocalyx* were not discussed in detail.

To the best of our knowledge, we report in the present work the first sequences from the $trnE^{UUC}$-$trnT^{GGU}$ intergenic spacer region (*trnE-trnT*) of the cpDNA for the subfamily Caesalpinioideae (Leguminosae). This intergenic spacer is located within the $trnD^{GUC}$-$trnT^{GGU}$ region, which has relatively high rates of substitution compared to other chloroplast regions (Hahn, 2002; Shaw *et al.*, 2005) and has been effectively used in phylogenetic studies at lower taxonomic levels (Friesen *et al.*, 2000; Lu *et al.*, 2001). We aimed to evaluate the usefulness of *trnE-trnT* spacer sequences to provide further information on the phylogeny of Cassiinae, focusing in the taxon *Chamaecrista* sect. *Xerocalyx*. However, the comparison of phylogenetic hypotheses derived from different sequences from both nuclear and chloroplast genomes is crucial to obtain additional resolution to represent true organismal relationships (Kuzoff *et al.*, 1998). Therefore, we also obtained several sequences from the internal transcribed spacer (ITS)/5.8S region of the nuclear ribosomal DNA (nrDNA) cistron, which comprises the first spacer (ITS1), the 5.8S rRNA gene and the second spacer (ITS2), to further investigate whether these molecular characteristics infer the true relationships within Cassiinae.

Material and Methods

Taxonomic sampling

Twelve specimens from the genus *Chamaecrista* were obtained, of which accessions from sect. *Xerocalyx* are the main target in the study. In order to also represent the phylogenetic diversity outside *Chamaecrista*, the ingroup also included other species from Cassiinae. Samples were collected from different locations of six states (Ceará, Piauí, Bahia, Tocantis, Goiás and Minas Gerais) from Brazil. Voucher specimens were deposited in the Herbarium Prisco Bezerra, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil. The list of taxa, locality data, voucher specimens, and GenBank accession numbers of the sequences are shown in Table 1. This study comprises two datasets, including the ITS/5.8S region of the nrDNA and the *trnE-trnT* complete intergenic spacer sequence from the cpDNA.

The nrDNA dataset included, in addition to the sequences determined in this work, those generated by de Souza Conceição *et al.* (2009) (FJ009815-FJ009869) and the sequences from *C. belemii* H.S. Irwin & Barneby (DQ787389), *Senna tora* L. (FJ572046), *Senna alata* (L.) H.S. Irwin & Barneby (FJ980412), and *Cassia javanica* L. subsp. *nodosa* (FJ980413), which were retrieved from GenBank. Four outgroups [*Bauhinia ungulata* L., *Copaifera coriacea* Mart., *Hymenaea courbaril* L., and *Martiodendron mediterraneum* (Mart.ex Benth) Köeppen] were chosen based on previously published phylogenies of Caesalpinioideae (Bruneau *et al.*, 2001; de Souza Conceição *et al.*, 2009; Kajita *et al.*, 2001).

The cpDNA dataset included 12 specimens from *Chamaecrista*, represented by four of the six sections with emphasis on *Xerocalyx*, six species from *Senna* and one from *Cassia*. The tribe Cercideae has already been demonstrated to be the sister group of the remainder Leguminosae in molecular analyses (Bruneau *et al.*, 2001; Kajita *et al.*, 2001). Therefore, *Bauhinia pentandra* (Bong) Vog. ex. Steua (Cercideae) was selected as the outgroup for the *trnE-trnT* intergenic spacer analysis.

DNA extraction

Total genomic DNA was extracted from plant (leaf) material sampled from herbarium specimens. Samples (0.3 g) were ground in liquid nitrogen and digested for 1 h at 60 °C in CTAB extraction buffer (2% w/v CTAB, 100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 1.4 M NaCl, and 0.2% v/v 2-mercaptoethanol). Further processing of the samples was done as described by Foster and Twell (1996).

Table 1 - List of taxa (subtribe Cassiinae and outgroups) included in the phylogenetic analysis.

Taxa	Origin of samples	Voucher specimen numbers[1]	GenBank accession No.	
			ITS/5.8S	*trnE-trnT*
Bauhinia pentandra	Juazeiro do Norte, Ceará	EAC 34765	-	GU175320
Cassia fistula	Campus do Pici-UFC, Fortaleza, Ceará	EAC 31694	GU175310	GU175321
Chamaecrista				
Sect. *Absus*				
C. hispidula	Novo Horizonte, Jardim, Ceará	EAC 34769	-	GU175328
Sect. *Caliciopsis*				
C. calycioides	Parque Botânico do Ceará, Caucaia, Ceará	EAC 26229	GU175311	GU175322
Sect. *Chamaecrista*				
C. flexuosa	Chapada da Diamantina, Bahia	EAC 26280	-	GU175327
C. rotundifolia	Jacobina, Bahia	EAC 29154	-	GU175331
C. tenuisepala	Chapada da Ibiapaba, Tianguá, Ceará	EAC 29068	-	GU175332
C. trichopoda	Jacobina, Bahia	EAC 29106	GU175318	GU175333
Sect. *Xerocalyx*				
C. desvauxii var. glauca	Parque Nacional do Araguaia, Lagoa da Confusão, Tocantins	EAC 28602	GU175312	GU175323
C. desvauxii var. linearis	Formosa do Rio Preto, Bahia	EAC 28607	GU175314	GU175324
C. desvauxii var. mollissima	Formoso, Minas Gerais	EAC 28604	GU175313	GU175325
C. diphylla	Jericoacoara, Jijoca, Ceará	EAC 29521	GU175315	GU175326
C. ramosa var. lucida	Niquelândia, Goiás	EAC 28606	GU175316	GU175330
C. ramosa var mollissima	Chapada das Mangabeiras, Barreiras do Piauí, Piauí	EAC 24258	GU175317	GU175329
Senna				
Sect. *Chamaefistula*				
S. macranthera	Sítio Jaburu, Ubajara, Ceará	EAC 34832	-	GU175335
S. obtusifolia	Campus do Pici-UFC, Fortaleza, Ceará	EAC 31702	GU175319	GU175336
S. occidentalis	Juazeiro do Norte, Ceará	EAC 34764	-	GU175337
S. rizzinii	Parque Botânico do Ceará, Caucaia, Ceará	EAC 26510	-	GU175338
Sect. *Peiranisia*				
S. trachypus	Campus do Pici-UFC, Fortaleza, Ceará	EAC 18345	-	GU175339
Sect. *Senna*				
S. alata	Horto de Plantas Medicinais (LPM)-UFC, Fortaleza, Ceará	EAC 31591	-	GU175334

[1] Voucher specimens were deposited in the Herbarium Prisco Bezerra-UFC, Fortaleza-Ceará, Brazil.

DNA concentration was determined by measuring the absorbance at 260 nm (A_{260}) of a ten-fold dilution of each sample. The quality of all DNA preparations was checked by 0.8% agarose gel electrophoresis according to Sambrook *et al*. (1989).

PCR amplification and DNA sequencing

Amplification of the $trnE^{UUC}$-$trnT^{GGU}$ intergenic spacer region of the cpDNA was performed using the primers trnET-F (5'-ATCGGATTTGAACCGATGAC-3') and trnET-R (5'-CCCAGGGGAAGTCGAATC-3'). These primers were designed based on the *Lotus japonicus* chloroplast genome sequence (GenBank accession number: NC_002694; Kato *et al.*, 2000). For the internal transcribed spacers (ITS1 and ITS2) and the 5.8S rRNA coding region of the nrDNA, the primers ITS4 (5'-TCCTCCGCTTATT GATATGC-3') and ITS5 (5'-GCAAGTAAAAGTCGTA ACAAGA-3') were used, as suggested by Becerra and Venable (1999). Both amplification reactions were performed in a final volume of 25 µL containing: 800-1000 ng of genomic DNA (template); 20 mM Tris-HCl, pH 8.4; 50 mM KCl; 1.5 mM $MgCl_2$; 100 mM 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. For the *trnE-trnT* spacer, the cycling parameters included an initial denaturation step (4 min at 94 °C) followed by 35 cycles of 1 min at 94 °C, 1 min at 58 °C for primer annealing, and 1 min and 30 s at 72 °C for extension. The PCR cycling parameters for the amplification of the ITS/5.8S region com-

prised an initial denaturation step of 94 °C for 4 min, followed by 35 cycles with 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 2 min. The last cycle for both reactions was followed by a final incubation step of 9 min at 72 °C, and then the PCR products were 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, PCR products were purified from the remaining reactions using the 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 then analyzed in a MegaBACE 1000 automatic sequencer (GE Healthcare Life Sciences). Each PCR product was sequenced three times in both directions using the same primers employed in the amplification reaction. Sequencing of the ITS/5.8S region from several samples was not successful. The fact that most of DNA samples isolated were from herbarium specimens might explain this issue.

Sequence alignment and phylogenetic analyses

The quality of the DNA sequences was checked and overlapping fragments were assembled using the Phred/Phrap/Consed package (Ewing *et al.*, 1998; Ewing and Green, 1998; Gordon *et al.*, 1998). BLASTn searches (Zhang *et al.*, 2000) were conducted in GenBank to detect potential contaminant sequences. For the *trnE-trnT* spacer, positions of coding and noncoding borders were determined by comparison with *Lotus japonicus* cpDNA sequence (NC_002694), while the ITS/5.8S regions were determined by comparison with the nrDNA sequence from *Senna tora* (FJ572046) using a method based on Hidden Markov Models (HMMs) to delimit the ITS2 region (Keller *et al.*, 2009). Any uncertain base positions, generally located close to priming sites, were excluded from the phylogenetic analyses.

Assembled sequences with high quality (phred >20) comprising the two datasets mentioned above were separately aligned using ClustalX version 2.0.9 (Larkin *et al.*, 2007), with default gap penalties, and manually corrected using the software BioEdit version 7.0.3 (Hall, 1999) to produce an alignment with the fewest number of changes (indels or nucleotide substitutions). Alignment files are available upon request to the corresponding author.

Phylogenetic analyses were performed independently for each dataset in PAUP* version 4.0b10 (Swofford, 2002) and MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001). Maximum parcimony (MP) analyses were conducted using heuristic searches with tree-bisection-reconnection (TBR) branch-swapping, ACCTRAN character optimization, and the Multrees option in effect, holding a maximum of ten most parcimonious 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 (Sanderson and Doyle, 1993). In all phylogenetic analyses, 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 for the original search.

The selection of the most suitable model for the Bayesian inferences was calculated using the Akaike Information Criterion (AIC) by MrModeltest, version 2.3 (Nylander, 2004), which presents several important advantages over other strategies of model selection (Posada and Buckley, 2004). Two independent analyses with five million generations were run to estimate parameters related 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) for each node. Trees generated were visualized by TreeView (Page, 1996). The proportion of variable sites and the GC content were calculated using the MEGA software, version 4.0 (Tamura *et al.*, 2007).

Results

DNA sequence characteristics

PCR amplification was not uniformly successful for all *loci* across the sampled taxa. While we were able to generate good quality sequences (phred >20) for the *trnE-trnT* intergenic spacer from all taxa, the ITS/5.8S region was not sequenced for several specimens under study. We hypothesize that, despite the multiple copies of the ITS/5.8S region presented in the nuclear genome, the method of DNA extraction chosen for herbarium specimens was not feasible to preserve good quality genomic DNA in contrast to chloroplast DNA.

The sequence characteristics for each DNA data set are summarized in Table 2. Complete ITS/5.8S sequences showed difficulties in alignment, mostly in the ITS1 region, among the three genera studied. The length of the ITS/5.8S region ranged from 612 bp (*Cassia fistula* L.) to 663 bp [*C. trichopoda* (Benth.) H.S. Irwin & Barneby], and the GC content ranged from 57% [*S. obtusifolia* (L.) H.S. Irwin & Barneby] to 62.8% [*C. desvauxii var. mollissima* (Benth.) H.S. Irwin & Barneby]. In most taxa examined, the 5.8S rRNA gene sequence had a constant length of 158 bp, with one exception for *C. ramosa var. lucida* (Benth.) H.S. Irwin & Barneby, whose sequence was 159 bp long. Concerning the *trnE-trnT* intergenic spacer, the sequences ranged from 796 bp [*S. occidentalis* (L.) Link] to 837 bp (*B. pentandra*), and the GC content ranged from 29.9% (*B. pentandra*) to

Table 2 - Sequence alignment information and summary of the maximum parsimony (MP) analyses.

	ITS/5.8S	*trnE-trnT*
Number of sequences	64	20
% GC content	61.1	32.4
Aligned length (bp)	749	898
Conserved characters	194 (26%)	671 (75%)
Parsimony Informative Sites (PIS)	453 (60.5%)	61 (7%)
Number of MP trees	886	4
Length of MP trees	2164	207
Consistency index (CI)	0.5	0.89
Retention index (RI)	0.82	0.91

33.5% (*Cassia fistula*). The small GC content observed in the cpDNA alignment is mainly due to polyA or polyT regions. This same observation has already been noticed in Asteraceae *trnD-trnT* sequences (Shaw *et al.*, 2005).

The ITS/5.8S dataset contained the highest ratio of parsimony informative sites (PIS) to aligned characters (60.5%), while the *trnE-trnT* intergenic spacer presented 7%. However, the ITS/5.8S region presented the lowest consistency and retention indexes. Values closer to 1 indicate a low amount of homoplasy. This convergent event has been interpreted as undesirable for phylogenetic data (Lyons-Weiler *et al.*, 1996; Swofford *et al.*, 1996), because one character may mislead the true branching history. Nevertheless, data that are homoplastic may still imply phylogenetic resolution, sometimes better than internally consistent data sets (Källersjö *et al.*, 1999; Wenzel and Siddall, 1999).

Phylogenetic analyses

Descriptive values for the MP trees resulting from the two datasets studied are summarized in Table 2. The 50% majority rule consensus trees from the Bayesian analyses for the ITS/5.8S and *trnE-trnT* intergenic spacer data sets are shown in Figures 1 and 2, respectively. Both MP and Bayesian analyses were mostly congruent. The monophyly of the *Chamaecrista* genus is well supported and a sister relationship between *Senna* and *Cassia* was also observed, although none of the data sets gave a robust confidence value.

The monophyly of section *Xerocalyx* is strongly supported in all analyses (BS and PP above 98). Bayesian analysis of the ITS/5.8S fragment provided better resolution within *Xerocalyx*. However, none of the regions studied provided enough resolution to clearly resolve the relationships among the specimen varieties from *C. ramosa* (Vogel) H.S. Irwin & Barneby and *C. desvauxii* (Collad.) Killip (Figures 1 and 2).

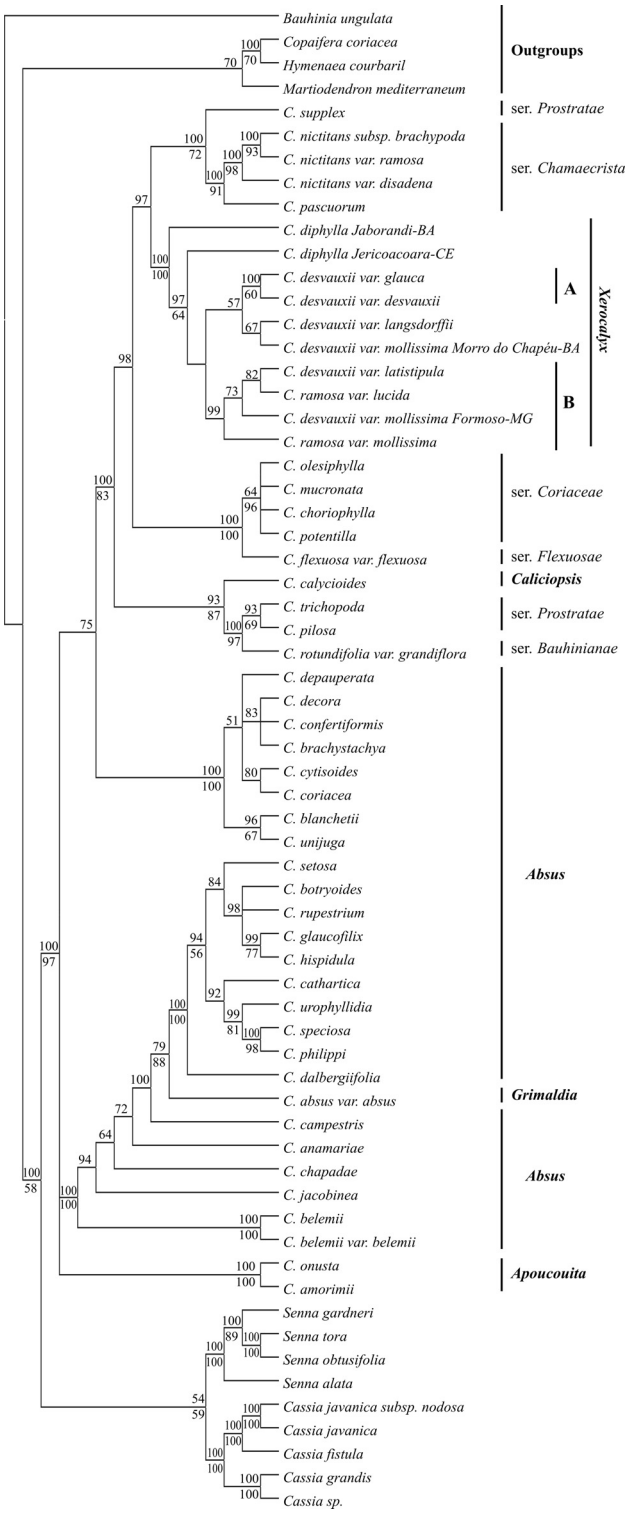

Figure 1 - Majority rule consensus tree based on the Bayesian analysis of the ITS/5.8S data set. *Bauhinia ungulata*, *Copaifera coriacea*, *Hymenaea courbaril* and *Martiodendron mediterraneum* were set as outgroups. Numbers above lines are Bayesian posterior probability values and below lines are bootstrap (1,000 replicates) support values from maximum parsimony analysis. Sections and series of *Chamaecrista* follow Irwin & Barneby (1982).

Discussion

The genus *Cassia* s.l., formerly comprised of 600 species, was submitted to several taxonomic treatments that led to the segregation of this large genus into three taxa (*Cassia* s. str., *Chamaecrista* and *Senna*), which were fur-

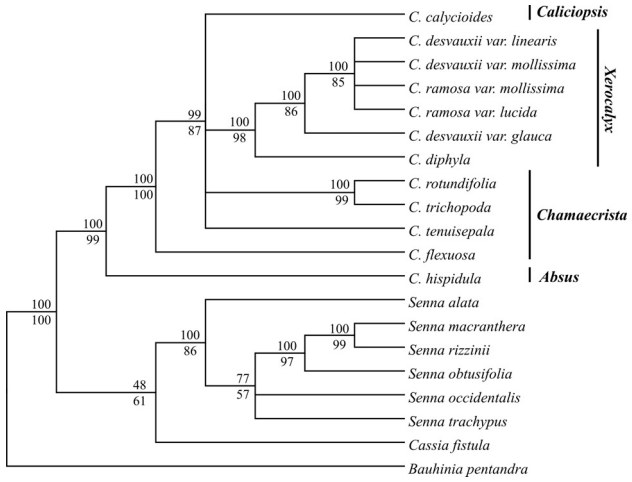

Figure 2 - Majority rule consensus tree based on the Bayesian analysis of the *trnE-trnT* intergenic spacer data set. *Bauhinia pentandra* was set as outgroup. Numbers above lines are Bayesian posterior probability values and below lines are bootstrap (1,000 replicates) support values from maximum parsimony analysis. Sections of *Chamaecrista* follow Irwin & Barneby (1982).

ther ascribed to subtribe Cassiinae (Irwin and Barneby, 1981, 1982). This separation was further confirmed by floral development (Tucker, 1996) and phenetic studies (Boonkerd *et al.*, 2005).

In the present study, it was observed that *Senna* and *Cassia* are monophyletic, corroborating previous molecular phylogenetic studies (Bruneau *et al.*, 2001, 2008; Herendeen *et al.*, 2003; Marazzi *et al.*, 2006). The absence of taxon sampling outside Cassiinae did not allow us to make any conclusive remarks concerning the monophyly and generic relationships within the subtribe, although our results favor the sister relationship between *Cassia* and *Senna* (Bruneau *et al.*, 2001, 2008; Marazzi *et al.*, 2006) rather than between *Chamaecrista* and *Senna* (Doyle *et al.*, 1997; Kajita *et al.*, 2001; Herendeen *et al.*, 2003; De-Paula and Oliveira, 2008).

Numerous peculiarities in the inflorescence structure (Tucker, 1996) and the presence of root nodules (Sprent, 2000) make *Chamaecrista* quite an interesting taxon within Cassieae. More recently, biochemical and genetic studies have been conducted in order to elucidate the variability within *Chamaecrista* (*e.g.* Conceição *et al.*, 2008a, b; Costa *et al.*, 2007; Silva *et al.*, 2007; de Souza Conceição *et al.*, 2009). The present study tried to elucidate some phylogenetic relationships within this genus, focusing on sect. *Xerocalyx*.

Our Bayesian analysis of the ITS/5.8S sequences is highly congruent with previous results based on a combined dataset of ITS/5.8S and plastid *trnL-F* regions (de Souza Conceição *et al.*, 2009) in which sections *Apoucouita* and *Xerocalyx* were supported as monophyletic while sections *Absus* and *Chamaecrista* were found to be paraphyletic. Moreover, *C. calyciodes* (sect. *Caliciopsis*) also appeared as a sister group of members of sect. *Chamaecrista* based on our ITS/5.8S dataset. *C. calyciodes* presents ambiguous characteristics relative to sections *Chamaecrista* and *Xerocalyx* (Irwin, 1964). As member of *Caliciopsis*, it resembles herbaceous specimens from sect. *Chamaecrista* in the morphology and chromosome number, while a resemblance to specimens from *Xerocalyx* is evident in the close parallel striate venation of the sepals. Whether the members from *Caliciopsis* evolved independently from the other two mentioned sections or represent a recombination of genetic material from both of them is unknown (Irwin and Barneby, 1982).

The sect. *Chamaecrista* comprises the largest number of species and is subdivided into six series. The controversial taxonomic classification of this group, probably due to an explosive evolutionary radiation, was pointed out earlier by Irwin and Barneby (1982). The less numerous ser. *Flexuosae* H.S. Irwin & Barnaby is represented here by *C. flexuosa* (L.) Greene, one very distinct species relative to the other American *Chamaecrista*, essentially by the presence of peculiar characteristics, like venulation of the leaflets, stem and angulate leaf-stalks (Irwin and Barneby, 1982). The basal node position of *C. flexuosa* relative to sect. *Chamaecrista*, *Caliciopsis* and *Xerocalyx* in the *trnE-trnT* spacer tree topology is congruent with previous work (de Souza Conceição *et al.*, 2009), which might be a reflection of the morphological features mentioned above. However, this topology was not supported by our ITS/5.8S analyses.

As observed by de Souza Conceição *et al.* (2009), ser. *Prostratae* (Benth.) H.S. Irwin & Barnaby also appeared polyphyletic in our phylogenetic analyses. Furthermore, *C. rotundifolia* [ser. *Bauhinianae* (Collad.) H.S. Irwin & Barnaby] grouped with *Prostratae* taxa [*C. trichopoda* and *C. pilosa* (L.) Greene] with robust node support. A resemblance among members of the series *Bauhinianae* and *Prostratae* has already been suggested, despite determined differences, like the glandless petioles and the reduced pair of leaves observed only in *Bauhinianae* (Irwin and Barneby, 1982). Moreover, we suggest that the common features observed between *C. diphylla* and *C. rotundifolia*, such as the number and sometimes form of the leaflets, that causes confusion in herbarium specimens identification, cannot be regarded as a great evolutionary step. Thus, other characteristics, like the venation of the leaflets, presence of a petiolar gland and chromosome number (2n = 14), must be interpreted as synapomorphies for sect. *Xerocalyx*.

Sect. *Xerocalyx* has suffered continuous taxonomic reorganization. After its reformulation by Irwin and Barneby (1982), this section was considered as a macrospecies in which evolutionary processes are still under development to give rise to truly discrete units at the subgeneric level. It forms an extremely distinct type, segregated into three species, distinguished by the number (diphyllous: *C. diphylla*; tetraphyllous: *C. ramosa* and *C. desvauxii*) and

size of the leaflets (microphyllous: *C. ramosa*; macrophyllous: *C. desvauxii*). More recently, Fernandes and Nunes (2005) discussed the classification proposed by Irwin and Barneby (1982) and, considering them as extremely subjective, proposed the elevation of several varieties to the species level.

The monophyly of *Xerocalyx* is strongly supported by our phylogenetic analyses. Considering the ITS/5.8S data set and the tetraphyllous group, the Bayesian analysis revealed two clades A and B (Figure 1) with robust node support. Clade B suggests that the size of the leaflets cannot be considered a truly discrete unit to distinguish *C. ramosa* and *C. desvauxii* specimens. However, only clade A was congruent with the MP analysis (weak branch support). In the plastid data set analyses only *C. desvauxii var. glauca* separated from the other tetraphyllous, remaining at the basis node. According to Fernandes and Nunes (2005), this *C. desvauxii* variety, well determined for its great size among *Xerocalyx* and for considerable morphological variation, like the leaflets and stipules glaucescent, should be recognized at the species status as *C. latistipula*. Moreover, considering ecological, geographical, morphological, reproductive and genetic data, Costa *et al.* (2007) proposed that two other varieties, *C. desvauxii var. latistipula* and *C. desvauxii var. graminea*, should be treated as distinct species. However, none of the trees analyzed had enough resolution in discriminating the microphyllous and the macrophyllous groups.

Another interesting feature found in sect. *Xerocalyx* is the paraphyletic relationship observed between *C. desvauxii var. mollissima* from Morro do Chapéu, BA, and Formoso, MG. The phylogenetic position of the specimen from Formoso, MG is clearly discriminated in Clade B (PP = 99); however, the phylogenetic resolution obtained for the other specimen was not clear. Although the results don't allow us to draw conclusive remarks, we suggest that cryptic species might have emerged within sect. *Xerocalyx*.

In the present work, our analyses reinforce the need for new sectional boundaries in the genus *Chamaecrista*, especially the inclusion of sections *Caliciopsis* and *Xerocalyx* in sect. *Chamaecrista* as suggested by de Souza Conceição *et al.* (2009). None of the trees analyzed showed the microphyllous and the macrophyllous groups as distinct clades. Thus, we hypothesize that speciation events are still ongoing in the tetraphyllous group, which is congruent with the macro-species hypothesis suggested by Irwin and Barneby (1982). On the other hand, it is premature to draw any final conclusions on species circumscriptions in the tetraphyllous complex. A more extensive revision and phylogenetic study of this group are necessary to further establish the current taxonomic shuffling involved in section *Xerocalyx*.

Acknowledgments

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

References

Becerra JX and Venable DL (1999) Nuclear ribosomal DNA phylogeny and its implications for evolutionary trends in Mexican *Burserea* (Burseraceae). Am J Bot 86:1047-1057.

Boonkerd T, Pechsri S and Baum BR (2005) A phenetic study of *Cassia sensu lato* (Leguminosae-Caesalpinioideae, Cassieae, Cassiinae) in Thailand. Plant Syst Evol 232:153-165.

Bruneau A, Forest F, Herendeen PS, Klitgaard BB and Lewis GP (2001) Phylogenetic relationships in the Caesalpinioideae (Leguminosae) as inferred from chloroplast *trnL* intron sequences. Syst Bot 26:487-514.

Bruneau A, Mercure M, Lewis GP and Herendeen PS (2008) Phylogenetic patterns and diversification in the caesalpinioid legumes. Botany 86:697-718.

Conceição AS, Queiroz LP and Borba EL (2008a) Natural hybrids in *Chamaecrista* sect. *Absus* subsect. *Baseophyllum* (Leguminosae-Caesalpinioideae): Genetic and morphological evidence. Plant Syst Evol 271:19-27.

Conceição AS, Queiroz LP, Lambert SM, Pereira ACS and Borba EL (2008b) Biosystematics of *Chamaecrista* sect. *Absus* subsect. *Baseophyllum* (Leguminosae-Caesalpinioideae) based on allozyme and morphometric analyses. Plant Syst Evol 270:183-207.

Costa CB, Lambert SM, Borba EL and de Queiroz LP (2007) Post-zygotic reproductive isolation between sympatric taxa in the *Chamaecrista desvauxii* complex (Leguminosae-Caesalpinioideae). Ann Bot 99:625-35.

De-Paula OC and Oliveira DMT (2008) Multiple pleurograms in *Chamaecrista* Moench (Leguminosae, Caesalpinioideae). Bot J Linn Soc 157:487-492.

de Souza Conceição A, Paganucci de Queiroz L, Lewis GP, Gomes de Andrade MJ, Machado de Almeida PR, Schnadelbach AS and van den Berg C (2009) Phylogeny of *Chamaecrista* Moench (Leguminosae-Caesalpinioideae) based on nuclear and chloroplast DNA regions. Taxon 58:1168-1180.

Doyle JJ, Chappill JA, Bailey CD and Kajita T (2000) Towards a comprehensive phylogeny of legumes: Evidence from *rbcL* sequences and nonmolecular data. In: Herendeen PS and Bruneau A (eds) Advances in Legume Systematic. Part 9. Royal Botanic Gardens, Kew, pp 1-20.

Doyle JJ, Doyle JL, Ballenger JA, Dickson EE, Kajita T and Ohashi H (1997) A phylogeny of the chloroplast gene *rbcL* in the Leguminosae: Taxonomic correlations and insights into the evolution of nodulation. Am J Bot 84:541-554.

Ewing B and Green P (1998) Base-calling of automated sequencer traces using Phred. II. Error probabilities. Genome Res 8:186-194.

Ewing B, Hillier L, Wendl MC and Green P (1998) Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. Genome Res 8:175-185.

Fernandes A and Nunes EP (2005) Registros Botânicos. Edições Livro Técnico, Fortaleza, 112 pp.

Foster GR and Twell D (1996) Plant Gene Isolation. Principles and Practice. John Wiley & Sons Ltd., West Sussex, 426 pp.

Friesen N, Fritsch RM, Pollner S and Blattner FR (2000) Molecular and morphological evidence for an origin of the aberrant

genus *Milula* within the Himalayan species of *Allium* (Alliaceae). Mol Phylogenet Evol 17:209-218.

Gordon D, Abajian C and Green P (1998) Consed: A graphical tool for sequence finishing. Genome Res 8:195-202.

Hahn WJ (2002) A phylogenetic analysis of the Arecoid line of palms based on plastid DNA sequence data. Mol Phylogenet Evol 23:189-204.

Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95-98.

Herendeen PS, Bruneau A and Lewis GP (2003) Phylogenetic relationships in caesalpinioid legumes: A preliminary analysis based on morphological and molecular data. In: Klitgaard BB and Bruneau A (eds) Advances in Legume Systematics. Part 10, Higher Level Systematics. Royal Botanic Gardens, Kew, pp 37-62.

Huelsenbeck JP and Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754-755.

Irwin HS (1964) Monographic studies in *Cassia* (Leguminosae – Caesalpinioideae) I. Section *Xerocalyx*. Mem N Y Bot Gard 12:1-114.

Irwin HS and Barneby RC (1981) Tribe 2. Cassiae Bronn (1822). In: Pohlhill RM and Raven PH (eds) Advances in Legume Systematics. Part 1. Royal Botanic Gardens, Kew, pp 97-106.

Irwin HS and Barneby RC (1982) The American Cassinae: A synoptical revision of Leguminosae tribe Cassieae subtribe Cassinae in the New World. Mem N Y Bot Gard 35:1-918.

Kajita T, Ohashi H, Tateishi Y, Bailey CD and Doyle JJ (2001) *rbcL* and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and Allies. Syst Bot 26:515-536.

Källersjö M, Albert VA and Farris JS (1999) Homoplasy increases phylogenetic structure. Cladistics 15:91-93.

Kato T, Kaneko T, Sato S, Nakamura Y and Tabata S (2000) Complete structure of the chloroplast genome of a legume, *Lotus japonicus*. DNA Res 7:323-330.

Keller A, Schleicher T, Schultz J, Müller T, Dandekar T and Wolf M (2009) 5.8S-28S rRNA interaction and HMM-based ITS2 annotation. Gene 430:50-57.

Kuzoff RK, Sweere JA, Soltis DE, Soltis PS and Zimmer EA (1998) The phylogenetic potential of entire 26S rDNA sequences in plants. Mol Biol Evol 15:251-263.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al.* (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947-2948.

Lewis GP (2005) Tribe Cassieae. In: Lewis GP, Schrire B, MacKinder B and Lock M (eds) Legumes of the World. Royal Botanical Gardens, Kew, pp 111-124.

Lu SY, Peng C, Cheng YP, Hong KH and Chiang TY (2001) Chloroplast DNA phylogeography of *Cunningiamnia konishlii* (Cupressaceae), an endemic conifer of Taiwan. Genome 44:797-807.

Lyons-Weiler J, Hoelzer GA and Tausch R (1996) Relative apparent synapomorphy analysis (RASA). I. The statistical measure of phylogenetic signal. Mol Biol Evol 13:749-757.

Marazzi B, Endress PK, Queiroz LP and Conti E (2006) Phylogenetic relationships within *Senna* (Leguminosae, Cassiinae) based on three chloroplast DNA regions: Patterns in the evolution of floral symmetry and extrafloral nectaries. Am J Bot 93:288-303.

Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Page RDM (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357-358.

Posada D and Buckley TR (2004) Model selection and model averaging in phylogenetics: Advantages of the AIC and Bayesian approaches over likelihood ratio tests. Syst Biol 53:793-808.

Sambrook J, Fritsch EF and Maniatis T (1989) Molecular Cloning. A Laboratory Manual. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanderson MJ and Doyle JJ (1993) Phylogenetic relationships in North American *Astragalus* (Fabaceae) based on chloroplast DNA restriction site variation. Syst Bot 18:395-408.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE and Small RL (2005) The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92:142-166.

Silva RM, Fernandes GW and Lovato MB (2007) Genetic variation in two *Chamaecrista* species (Leguminosae), one endangered and narrowly distributed and another widespread in the Serra do Espinhaço, Brazil. Can J Bot 85:629-636.

Sprent JI (2000) Nodulation as a taxonomic tool. In: Herendeen PS and Bruneau A (eds) Advances in Legume Systematics. Part 9. Royal Botanic Gardens, Kew, pp 21-43.

Swofford DL (2002) PAUP* (v. 4.0b10). Phylogenetic analysis using parsimony (* and other methods). Sinauer Associates, Sunderland.

Swofford DL, Olsen GJ, Waddell PJ and Hillis DM (1996) Phylogenetic inference. In: Hillis DM, Moritz C and Mable BK (eds) Molecular Systematic. Sinauer, Sunderland, pp 407-514.

Tamura K, Dudley J, Nei M and Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software v. 4.0. Mol Biol Evol 24:1596-1599.

Tucker SC (1996) Trends in evolution of floral ontogeny in *Cassia* sensu stricto, *Senna*, and *Chamaecrista* (Leguminosae, Caesalpinioideae, Cassieae, Cassiinae); a study in convergence. Am J Bot 83:687-711.

Wenzel JW and Siddall ME (1999) Noise. Cladistics 15:51-64.

Zhang Z, Schwartz S, Wagner L and Miller W (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol 7:203-214.

Associate Editor: Marcio C. Silva Filho