CrossMark

ORIGINAL PAPER

Screening chloroplast regions in *Calophyllum brasiliense* (Calophyllaceae) for suitability for population genetic analyses

Cecilia B. Percuoco[1] · Lucas G. Giménez[2] · Liliana N. Talavera Stéfani[1] ·
Jorge V. Crisci[3] · Carina F. Argüelles[1]

Abstract The uniparental inheritance and conserved structure of the chloroplast genome among many angiosperms makes it suitable for constructing gene lineages and inferring population histories based on these relationships. However, the analysis of a single chloroplast region to solve population issues has not yielded the desired amount of variability due to the low mutation rate. Therefore, it is necessary to screen and carefully analyze chloroplast intergenic or intronic regions to select the most appropriate for studying the genetic variability in the chloroplast genome of a particular taxon. In this context, the aim of this work was to characterize 8 regions in *Calophyllum brasiliense* using 11 selected primer pairs and to evaluate their usefulness for phylogeographic analysis. Intergenic spacers *petA-psbJ*, *petG-trnP* and *rpl32-trnL* and the *trnL* intron demonstrated potentially informative sites. The combination of these 4 regions will allow the analysis of genetic population structure and the integration of historical aspects to inform strategies for conserving *C. brasiliense*, especially for highly impacted populations and those at risk of local extinction.

Keywords cpDNA · cpSSRs · Noncoding regions

✉ Cecilia B. Percuoco
  ceciliapercuoco@fceqyn.unam.edu.ar

[1] Laboratorio GIGA, Instituto de Biología Subtropical, Nodo Posadas, Universidad Nacional de Misiones (UNaM), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Jujuy 1745, N3300NFK Posadas, Misiones, Argentina

[2] Cátedra de Genética Molecular, Facultad de Ciencias Exactas Químicas y Naturales (FCEQyN), Universidad Nacional de Misiones (UNaM), Félix de Azara 1552, N3300LQH Posadas, Misiones, Argentina

[3] Laboratorio de Sistemática y Biología Evolutiva (LASBE) - Facultad de Ciencias Naturales y Museo (FCNyM), Universidad Nacional de La Plata-(UNLP), Avenida 122 y 60, 1900 La Plata, Buenos Aires, Argentina

Introduction

Calophyllum brasiliense (Calophyllaceae) is a native tree species from Central and South America. Its distribution range extends from southern Mexico to northeast Argentina (Reitz and Klein 1978; Rodríguez et al. 2009). It grows exclusively in hygrophyle forests and in lowlands that are almost permanently flooded (Marques et al. 2003). It offers many ecosystem services as a component of riparian forests, and several active compounds of medicinal interest have been recently isolated from different plant tissues. These compounds have inhibitory activity against some tropical pathogens that cause regional and worldwide diseases such as schistosomiasis, Chagas disease and tuberculosis (Gasparotto et al. 2005; Rea et al. 2013; Pires et al. 2014).

The chloroplast genome (cpDNA) is now the most widely used to understand phylogenetic issues among plant families and genera (Shaw et al. 2014). In the past, the use of cpDNA for population analysis was discouraged because of its low mutation rate within species. However, since

Springer

low-cost sequencing technologies have become routine, many phylogeographic studies in plant species have been based on a variable number of regions (Olmstead and Palmer 1994; Dumoulin-Lapègue et al. 1997; Huang et al. 2004). At first, the few cpDNA regions commonly used for higher taxonomic levels sometimes did not serve the desired purpose, but once the first complete genomes were sequenced, the conservative nature and cpDNA synteny (Palmer 1987; Dumoulin-Lapègue et al. 1997) allowed researchers to test many potential regions. Shaw et al. (2005, 2007, 2014) compared the variability rates of chloroplast regions among distinct clades and found that the cpDNA could be much more informative than earlier thought. Nevertheless, the most suitable chloroplast regions for a population genetic analysis usually will depend on the target species. In this context, the aim of this study was to characterize 8 cpDNA regions among *C. brasiliense* samples derived from populations from Argentina, Paraguay and Mexico by using 10 previously reported primer pairs and 1 specific primer pair to evaluate their utility in phylogeographic analysis, information that could also help in conservation decisions within the most reduced populations.

Materials and methods

Material and study site

From 2 to 10 individuals from each of the populations in three regions were sampled and stored in Ziploc bags: (1) Misiones (San Ignacio POP—SI) and Corrientes (Puerto Valle POP—PV; Rincón Ombú POP—RO) Argentinean Provinces; (2) Misiones Paraguayan Department (Ayolas POP—AY and Isla Yacyretá POP—IY); and (3) a Mexican sample from the Ecological Reserve Jaguaroundí Park in Veracruz-Llave (Mexico POP—MX). Table 1 summarizes the geographic coordinates and the number of individual collected per site.

DNA extraction

Once leaves were completely dried, total genomic DNA was isolated from these tissues following the protocol of Stange et al. (1998) with modifications made by Percuoco et al. (2014), which involved the addition of antioxidant reagents (2% w/v polyvinylpyrrolidone [PVP], 5 mM ascorbic acid, 4 mM sodium diethyldithiocarbamatetrihydrate [DIECA] and 1.2% β-mercaptoethanol) to the digestion buffer and the extension of incubation steps (digestion step: 0.5 vs. 3 h; precipitation step: 0.5 vs. 1 h).

cpDNA amplification and sequencing

Seven chloroplast intergenic regions were amplified: *trnH-psbA*, *psbI-trnS* (*ccmp2*), *trnS-trnG-trnG*, *psbC-trnS*, *petA-psbJ*, *petG-trnP* and *rpl32-trnL*. The eighth region was the *trnL* intron. Eight heterologous and 2 universal primer pairs obtained from the literature and an intraspecific pair that we designed were used for the PCR reactions (Table 2). The PCR mix composition and cycling profiles are detailed in Tables 3 and 4, respectively. PCR products were purified using GFX PCR DNA #28-9034-70 (GE) and sequenced by Macrogen Inc. using both forward and reverse primers.

Sequence analysis

The quality control of each electropherogram and all sequence editing were done using BioEdit v.7.0.9.0 (Hall, 1999). Consensus sequences were created for each individual, and their identity was verified using BLASTn 2.2.25 (Zheng et al. 2000).

C. brasiliense cpDNA descriptive data

The percentage sequence coverage for *C. brasiliense* cpDNA was estimated using the genome of *Jatropha curcas* (NC_012224) as reference. We chose this genome arbitrarily among four available within the order Mapighiales. We also mapped the approximate physical position of

Table 1 Location of *Calophyllum brasiliense* populations sampled

POP ID	Locality	Geographic reference	N
San Ignacio (SI)	San Ignacio (Misiones—ARG)	27°16′34.4″S–55°34′11.9″W	11
Rincón Ombú (RO)	Ituzaingó (Corrientes—ARG)	27°24′54.42″S–56°29′43″W	11
Puerto Valle (PV)	Ituzaingó (Corrientes—ARG)	27°35′23.84″S–56°31′34.17″W	10
Yacyretá Island (IY)	Misiones (PY)	27°24v53.88″S–56°45′21.25″W	3
Ayolas (AY)	Misiones (PY)	27°24′6.48″S–56°51′17.75″W	2
Jaguaroundi Reserve (MX)	Coatzacoalcos (Veracruz Llave- MX)	18°7′20.27″N–94°21′45.15″W	6

POP ID identity of the population, *N* number of samples collected at each site

Table 2 Details of the primers used for the PCR amplification in *Calophyllum brasiliense* chloroplast

Region	Primers sequences (5'– > 3')	Reference data		Data for *C. brasiliense* (present study)	
		References	Expected size	Sequence length	GenBank Accession
trnH(GUG)-*psbA*	ACT GCC TTG ATC CAC TTG GC / CGA AGC TCC ATC TAC AAA TGG	Hamilton (1999)	495 bp (*Nicotiana tabacum*)	409	–[c]
psbI-trnS-(GCU)	GAT CCC GGA CGT AAT CCT G	Weising and Gardner (1999)	189 bp (*Nicotiana tabacum*, dicotyledons)	194	–[d]
trnS(GCU)-*trnG*	ATC GTA CCG AGG GTT CGA AT / GCC GCT TTA GTC CAC TCA GC / GAA CGA ATC ACA CTT TTA CCAC	Hamilton (1999)	844 bp (*N. tabacum*)	600[a]	–[e]
trnS(GCU)-*trnG*-intron-*trnG*	AAC TCG TAC AAC GGA TTA GCA ATC / GAA TCG AAC CCG CAT CGT TAG	Shaw et al. (2007)	1434 bp (angiosperms, Eurosid I)	1897	KJ093667
psbC-trnS(UGA)	GGT CGT GAC CAA GAA ACC AC / GGT CCG AAT CCC TCT CTC TC	Demesure et al. (1995)	1680 bp (*Quercus robur*, universal)	1514	JN368484, JN596278
trnL intron	CGA AAT CGG TAG ACG CTA CG / GGG GAT AGA GGG ACT TGA AC	Taberlet et al. (1991)	200-1000 bp (universal, angiosperms)	611	JQ219659
	AAA ATG GGC AAT CCT GAG C / ATG GGA CTC TAT CTT TAT TCT CG	Percuoco et al. (2015)	450 bp (*C. brasiliense*)	450	KF854307–KF854309
petA-psbJ	GGA GAT GCA GAG ATA GTA C / CTC TTT GGT TGA TAG GTA CTG	Huang et al. (2004)	752 bp (*Trochodendron araliodes*)	1100[b]	–[e]
	ATA GGT ACT GTA ACT GGT ATT / GGT CTA ATT CCT ATA ACT TTG GC	Shaw et al. (2007)	734–1261 bp (angiosperms)	1137–1174	KJ093668, KJ093669
petG-trnP	GGG ATG TGG CGC AGC TTG G / AAC AGT TTG AGA AGG TTC AAT T	Huang et al. (2004)	350 bp (*T. araliodes*)	449	KJ093665, KJ093666
rpl32-trnL(UAG)	CTG CTT CCT AAG AGC AGC GT / CAG TTC CAA AAA AAC GTA CTT C	Shaw et al. (2007)	543–1417 bp (angiosperm)	811–901	KJ093670, KJ093671
Total length				7022–7149	

For each cpDNA region, expected size and tested species, based on published sources, are provided. Also listed are sequences lengths and GenBank accession number for *C. brasiliense* regions

[a] Two other bands (1300, 1100 bp) were always observed. Primer not selected

[b] These primers amplified this region only from some individuals

[c] Sequencing was difficult with these primers

[d] Fragment was shorter than the minimum size required for submission (< 200 bp)

[e] Primers were not selected to sequence these regions

Table 3 PCR mix composition for each of the 8 noncoding cpDNA regions amplified in *Calophyllum brasiliense*

Components (initial concentrations)	Final concentrations					
	psbC-trnS	*trnH-psbA*	*trnS-trnG*	Spacers[a]	*trnL* intron	*psbI-trnS* (*ccmp2*)
Buffer *Taq* $SO_4(NH_4)_2$ (10 ×)	1×	1×	1×	1×	1×	1×
$MgCl_2$ (25 mM)	2.5 mM	2.5 mM	2 mM	2 mM	2 mM	2 mM
dNTPs (2 mM)	0.1 mM	0.2 mM	0.2 mM	0.2 mM	0.2 mM	0.2 mM
Primer F (10 μM)	0.2 μM	0.3 μM	0.4 μM	0.1 μM	0.2 μM	0.2 μM
Primer R (10 μM)	0.2 μM	0.3 μM	0.4 μM	0.1 μM	0.2 μM	0.2 μM
BSA (10 mg/mL)	5 μg/μL	2.5 μg/μL	2.5 μg/μL	–	–	–
*Taq*ADN pol (5 U/μL)	1 U/25 μL	1 U/25 μL	1 U/25 μL	1 U/25 μL	1 U/25 μL	1 U/25 μL
ADN (50 μg/μL)	5 μg/μL	5 μg/μL	5 μg/μL	5 μg/μL	5 μg/μL	5 μg/μL
Ultrapure water to	25 μL	25 μL	25 μL	25 μL	25 μL	25 μL

[a] *trnS-trnG-trnG*, *petA-psbJ*, *petG-trnP*, *rpl32-trnL*

the 8 reported regions. Three other chloroplast sequences of *C. brasiliense* available in GenBank were included in this cpDNA illustration generated with DNA Plotter, free on-line software available at http://www.sanger.ac.uk/science/tools/dnaplotter (Carver et al. 2009).

Results

The 8 regions amplified showed single fragments of approximately the expected size in a variable number of individuals and populations. The editing and analysis of the obtained sequences allowed the characterizing of different noncoding cpDNA regions in *C. brasiliense*. Three intergenic spacers (see below) and the *trnL* intron had putative informative sites for *C. brasiliense*. The length of the sequences obtained and the GenBank accessions for each of the 8 regions analyzed are described in Table 2.

(1) *trnH-psbA* From the 500-bp products separated by agarose gel electrophoresis, a 409-bp consensus sequence was obtained with several microsatellite motifs included in it. These mono- and dinucleotide tandem repeats made the sequencing reaction difficult in some samples, even using microsatellite capillary electrophoretic conditions. The only alignment obtained showed 100% identity between 2 individuals from different Argentinean populations (SI and RO).

(2) *psbI-trnS* (*ccmp2*) Amplicons of ≈ 200-bp were obtained from samples from 2 Argentinean populations (SI, RO), whereas a smaller band size was observed in the Mexican samples. A 194-bp sequence was edited based on the SI and RO populations. This intergenic fragment corresponded to a chloroplast microsatellite that had an interrupted

motif in *C. brasiliense*. The sequence shared 63% identity with *Nicotiana tabacum* (Z00044) and maximum identity (73%) with *Mallotus paniculatum* (AY159390). The sequence could not be submitted to GenBank due to the minimal length of 200-bp required by the database (see http://www.ncbi.nlm.nih.gov/genbank/submit_types).

(3) *trnS-trnG* and *trnS-trnG* intron-*trnG* Two primer pairs were tested to amplify this region: (a) Hamilton's (1999) primers amplified the expected band size of ≈ 600-bp and two extra bands; and (b) The primers of Shaw et al. (2007) that amplified a unique, expected band of ≈ 2000-bp. From the latter, an 1897-bp consensus sequence was edited based on an individual from the Isla Yacyretá population, which shared 86% identity with *Jatropha curcas*. However, most accessions in the BLAST alignment showed coverage only up to 43% of the query sequence, except when aligned *with Licania heteromorpha* (KJ414481), another tree species from the family Chrysobalanaceae within the order Malpighiales, with which coverage was 66%. Four more samples (9%) were sequenced and partially aligned, but no polymorphic site was identified.

(4) *psbC-trnS* A unique amplicon of ≈ 1500-bp was amplified in all individuals from SI and RO (22/51%) but failed to be amplified among samples derived from the other populations (21/49%). Once the reliable electropherograms were edited, we obtained a 1514-bp consensus sequence, from which 1307 bp matched the *psbC* gene 3′ end encoding photosystem II CP43 chlorophyll apoprotein, and only 207-bp corresponded to the noncoding intergenic portion in *C. brasiliense*. Simultaneously, we carried out an additional digestion assay in this

Table 4 PCR profiles for 8 noncoding cpDNA regions amplified in *Calophyllum brasiliense*

Step	*psbC—trnS*		*trnH-psbA*		*trnS-trnG*		*rpl32-trnL*		Spacers[a]		*trnL* intron		*psbI-trnS* (*ccmp2*)	
	T	t	T	t	T	t	T	t	T	t	T	t	T	t
Initial denaturation	94	4 min	96	5 min	96	5 min	96	5 min	80	5 min	94	5	94	5 min
Denaturation	94	45 s	94	30 s	94	45 s	94	45 s	95	1 min	93	1	94	1 min
Annealing	57	45 s	62	30 s	55	1 min	55	1 min	50	1 min	62	1	60	1 min
Elongation	72	2 min	72	30 s	72	1 min	72	1.5 min	65	4 min	72	2	72	2 min
No. of cycles	30		35		30		30		35		30		35	
Final elongation	72	10 min	–	–	72	10 min	72	10 min	65	10 min	72	4	72	4 min
Final hold	4	∞	4	∞	4	∞	4	∞	4	∞	4	∞	4	∞

T temperature (°C), *t* time

[a] *trnS-trnG-trnG*, *petA-psbJ, petG-trnP, rpl32-trnL*

region with MboI restriction endonuclease and compared the profiles after standard agarose electrophoresis. In this assay, 30 individuals of the SI and RO POP were analyzed, showing a unique profile for both populations.

(5) *trnL intron* An amplicon of $\approx$ 700-bp was obtained in all populations. A first 611-bp consensus sequence was obtained for this region, but difficulties in sequencing of most of the amplicons required us to design a new specific primer pair for the species, which improved the electropherogram quality. The new consensus sequence was 450-bp long, and the alignment of the 43 sequences (100%) showed 4 informative sites, 2 SNPs, a 2-bp inversion and 1 indel 7-bp long. Three haplotypes were then identified. SSRs previously reported for this intron in other species (Hale et al. 2004) were interrupted and monomorphic in *C. brasiliense*.

(6) *petA-psbJ* Two primer pairs were assayed to amplify this region. Both pairs amplified the expected band of $\approx$ 1200-bp, but the primers designed by Huang et al. (2004) yielded no amplicon in several individuals (7/18%); thus, we chose the primers published by Shaw et al. (2007). The primer *petA* yielded overlapped products and low signal after 140-bp in all individuals analyzed, although the reverse primer *psbJ* allowed us to obtain 1200-bp readings. Once editing was completed, sequences from 1137 to 1174-bp long were aligned. Two indels, a substitution, and a $(T)_n$ microsatellite allowed discriminating 2 haplotypes exclusive to the southernmost populations and Mexican populations respectively.

(7) *petG-trnP* An amplification product of $\approx$ 500-bp was obtained from all samples; 19 (44%) of these were sequenced and resulted in a 449-bp consensus sequence. An 11-bp inversion was the only

polymorphism found, which discriminated 2 mono-haplotypic clusters, 1 from the Mexican population samples and the others from the other 5 populations.

(8) *rpl32-trnL* Bands of $\approx$ 1000-bp were observed in all populations. Once the sequences were edited, we obtained a fragment in the range of 811–890-bp. The variation in length was partially due to the inferior quality of the consensus sequence obtained and to a variable number of the $(AAATT)_n$ motifs that differentiated the Argentinean and Paraguayan from the Mexican populations.

Three other partial *C. brasiliense* coding sequences for cpDNA are available in GenBank. These correspond to *ndhF* (HQ331853), *matK* (HQ331550) and 8 accessions of *rbcL* (JQ591092–JQ591094; KC493366–KC493368; KC570911; KF981208). The alignment including all 8 *rbcL* sequences showed complete identity among them.

The 8 sequences reported here totaled 7149-bp, which increased to 9529 bp when *ndhF*, *matK* and *rbcL* (JQ591094) available sequences were added. Considering that the *Jatropha curcas* genome is 163,856-bp, approximately 6% of *C. brasiliense* cpDNA is now accessible, of which $\approx$ 4.5% corresponds to noncoding regions. The 11 regions (our 8 regions plus the 3 previous deposited in GenBank) now available for *C. brasiliense* were mapped over the reference genome (Fig. 1).

Discussion

All 8 regions analyzed in this work have been previously used in studies of several tree species (Dumoulin-Lapègue et al. 1997; Huang et al. 2004; Ramos et al. 2007), showing different informative rates. Eight heterologous, 2 universal and 1 specific primer pair developed in this work allowed amplicons to be obtained from many of the *C. brasiliense*

samples analyzed. Three intergenic regions (*petA-psbJ, petG-trnP, rpl32-trnL*) and the *trnL* intron were informative for the species. Particularly, two indels, a substitution and a mononucleotide cpSSR in *petA-psbJ,* an 11-bp inversion in *petG-trnP* and a 5-bp repeat motif within *trnL-rpl32* were the polymorphisms that allowed 2 haplotypes to be identified, 1 common to the southernmost *C. brasiliense* populations and the other restricted to the samples derived from Mexico. We can thus affirm that they are suitable chloroplast noncoding regions for population genetic approaches in *C. brasiliense*.

In the case of the *trnL* intron a 7-bp indel, one 2-bp inversion and 2 transversions were analyzed. These discriminated 3 haplotypes among *C. brasiliense* populations, 1 common haplotype distributed in Argentina and Paraguay, and 2 other haplotypes, exclusive to Paraguayan and Mexican populations respectively. Therefore, this region was used in a phylogeographic analysis (Percuoco et al. 2015). The *trnL* intron was used for several phylogenetic analyses (Hale et al. 2002; Hollingsworth et al. 2011) although some authors have discarded it for intraspecific studies due to its low resolution at low taxonomic levels (Shaw et al. 2005). This region was also considered as a putative barcode region in plants, because of its high conservation rate within species and the robustness of the amplification process (Taberlet et al. 2007).

The primers used to amplify the *psbC-trnS* intergenic spacer in *C. brasiliense* showed that this region contains a long *psbC* fragment, while the spacer is very short for the species. Although we did not sequence the fragment in many samples, the enzymatic digestion revealed no difference among the *C. brasiliense* samples, in accordance with the expected low level of variation because of the selection pressure exerted over coding regions. However, Ramos et al. (2007) used *psbC-trnS* in an intraspecific study of 17 populations of *Hymenaea stigonocarpa* from Brazil and found 23 haplotypes based on 13 substitutions and 4 indels in the alignment of a 524-bp fragment. Unfortunately, these sequences were not deposited in GenBank, and the electropherograms were not included as supplementary material for comparison with our results to enrich the present analysis.

The *trnH-psbA* intergenic spacer showed low-quality electropherograms due to several polyA/T repeats. There is, however, only one published phylogeographic analysis of *C. brasiliense* populations from Brazil where the use of this intergenic spacer seems to have been successful (Salgueiro et al. 2011). Although these authors used the same primers that we tested in this work, they reported a 263-bp sequence, a fragment much shorter than the one we obtained (409 bp). Surprisingly, they identified 28 informative sites within the 263 bp of the *trnH-psbA* intergenic spacer based on 192 individuals of *C. brasiliense*; however,
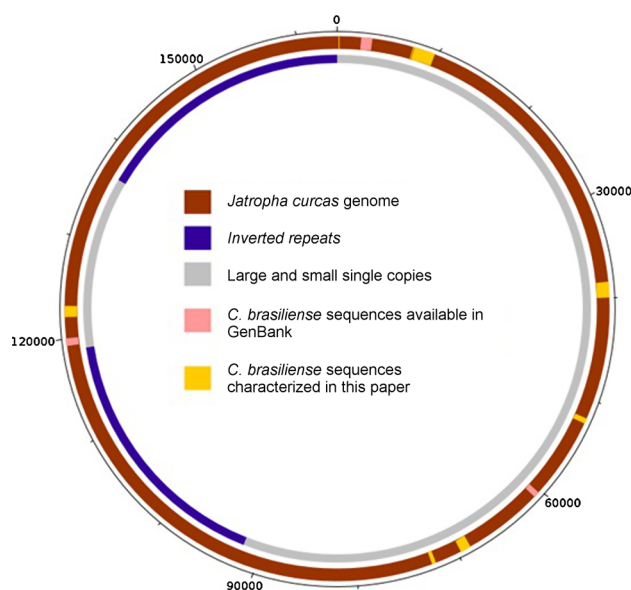

Fig. 1 Sequences of *C. brasiliense* cpDNA herein described mapped over the *Jatropha curcas* genome (NC_012224) used as reference

again, the electropherograms are not available to compare with our data.

The high number of repeats observed in the *trnH-psbA* region of *C. brasiliense* is in accordance with the common mono- and dinucleotide SSRs usually found in cpDNA, but these informative polymorphisms could easily arise as parallelism or homoplasies (Petit and Vendramin 2007), leading to wrong phylogeographic conclusions. Therefore, this type of mutation is often discarded in phylogeographic analysis (Huang et al. 2004; Lihová et al. 2010).

The *trnH-psbA* spacer has also been proposed as a potential DNA barcoding region (Kress et al. 2005; Chase et al. 2007), its high variability in length across larger angiosperm families and the ambiguity added by the obligatory manual editing because of multiple cpSSRs, as those observed in *C. brasiliense,* make it a problematic if it were the standard plant barcode alone (Chase et al. 2007).

Primer mismatches, a frequent problem when heterologous or universal primers are used, likely were responsible for the difficulties in the sequencing reaction for the *trnH-psbA*, *psbI-trnS*, *trnS-trnG-trnG* and *psbC-trnS* intergenic spacers amplified in *C. brasiliense*.

The chloroplast genome remains the most appropriate target in plant phylogeography studies due to its uniparental inheritance and its haploid nature. In this sense, cpDNA "records" historical fluctuations in populations for each species with high fidelity, providing a historical glance to improve conservation decisions. Although many cpDNAs have been completely sequenced (Malé et al. 2014), various nonmodal tropical tree species remain to be studied. Moreover, universal primers are frequently tested in several species but only in a few individuals within each

species; therefore, the intraspecific performance and features of the primers are unknown. Particularly for our native tree, all cpDNA regions were amplified with diverse levels of difficulty in the subsequent analysis of the sequences obtained. We improved our results when we designed a pair of primers for *C. brasiliense* as we showed for the *trnL* intron. Thus, we strongly recommend designing specific primer pairs to ensure the quality of the data even for those regions difficult to analyze. Finally, using more than one cpDNA region increases the number of informative sites and takes advantage of genome fitness. Thus, the 8 regions described in this work contribute to the knowledge of *C. brasiliense* cpDNA, but based on our results, we suggest *petA-psbJ*, *petG-trnP*, *rpl32-trnL* and the *trnL* intron analyzed in this work as suitable noncoding regions that could be useful for phylogeographic studies of this species.

Conclusions

All 8 chloroplastic regions were amplified, and a sequence of 7149-bp, which represents approximately 6% of *C. brasiliense* cpDNA is now available in GenBank. Three intergenic regions (*petA-psbJ*, *petG-trnP*, *rpl32-trnL*) and the *trnL* intron were informative and will be useful for population analyses of *C. brasiliense*.

References

Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J (2009) DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25:119–120

Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madriñán S, Petersen G, Seberg O, Jørgsensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M (2007) A proposal for a standardised protocol to barcode all land plants. Taxon 56:295–299

Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4:129–131

Dumoulin-Lapègue S, Demesure B, Fineschi S, Le Come V, Petit RJ (1997) Phylogeographic structure of white oaks throughout the European continent. Genetics 146:1475–1487

Gasparotto A Jr, Benzan MA, Piloto IC, Cortez DAZ (2005) Estudo fitoquímico da atividade moluscicida do *Calophyllum brasiliense* Camb. (Clusiaceae). Quim Nova 28:575–578

Hale ML, Squirrel J, Borland AM, Wolff K (2002) Isolation of polymorphic microsatellite loci in the genus *Clusia* (Clusiaceae). Mol Ecol Notes 2:506–508

Hale ML, Borland AM, Gustafsson MHG, Wolff K (2004) Causes of size homoplasy among chloroplast microsatellites in closely related *Clusia* sp. J Mol Evol 58:182–190

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

Hamilton MB (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol Ecol 8:513–525

Hollingsworth PM, Graham S, Little DP (2011) Choosing and using a plant DNA barcode. PLoS ONE 6:e19254

Huang S, Hwang S, Wang J, Lin T (2004) Phylogeography of *Trochodendron aralioides* (Trochodendraceae) in Taiwan and its adjacent areas. J Biogeogr 31:1251–1259

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci USA 102:8369–8374

Lihová J, Kudoh H, Marhold K (2010) Genetic structure and phylogeography of a temperate-boreal herb, *Cardamine scutata* (Brassicaceae), in northeast Asia inferred from AFLPs and cpDNA haplotypes. Am J Bot 97:1058–1070

Malé PG, Bardon L, Besnard G, Coissac C, Delsuc F, Engel J, Lhuillier E, Scotti-Saintagne C, Tinaut A, Chave J (2014) Genome skimming by shotgun sequencing helps resolve the phylogeny of a pantropical tree family. Mol Ecol Resour 14:966–975

Marques MC, Silva SM, Salino A (2003) Florística e estrutura do componente arbustivo-arbóreo de uma floresta higrófila da bacia do rio Jacaré-Pepira, SP, Brasil. Acta Bot Bras 17:495–506

Olmstead RG, Palmer JD (1994) Chloroplast DNA systematics: a review of methods and data analysis. Am J Bot 81:1205–1224

Palmer JD (1987) Chloroplast DNA and biosystematic uses of chloroplast DNA variation. Am Nat 130:S6–S29

Percuoco CB, Bich GA, Talavera Stéfani LN, Cardozo AE, Rodríguez ME, González NL, Sorol CB, Crivello JF, Crisci JV, Argüelles CF (2014) Assessment of genetic differentiation among relict populations of *Calophyllum brasiliense* Camb. (Calophyllaceae) from Northeast Argentina. J Biodivers Enviro Sci 5:87–98

Percuoco CB, Talavera Stéfani LN, Rodríguez ME, González NL, Crivello JF, Crisci JV, Argüelles CF (2015) Looking inside noncoding chloroplast regions of *Calophyllum brasiliense* (Calophyllaceae) to understand its southernmost population distribution. J Plant Sci 3(6):310–319

Petit RJ, Vendramin GG (2007) Plant phylogeography based on organelle genes: an introduction. In: Weiss S, Ferrand N (eds) Phylogeography of southern European refugia. Springer, Dordrecht, pp 23–97

Pires CTA, Brenzan MA, de Lima Scodro RB, Garcia Cortez DA, Dias Ghiraldi Lopes L, Dias Siqueira VL, Fressatti Cardoso R (2014) Anti-*Mycobacterium tuberculosis* activity and cytotoxicity of *Calophyllum brasiliense* Cambess. (Clusiaceae). Mem Inst Oswaldo Cruz 109:324–329

Ramos ACS, Lemos-Filho JP, Ribeiro RA, Santos FR, Lovato MB (2007) Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: caesalpinioideae) and the influence of Quaternary climate changes in the Brazilian Cerrado. Ann Bot 100:1219–1228

Rea A, Tempone AG, Pinto EG, Mesquita JT, Rodrigues E, Grus M, Silva L, Sartorelli P, Lago JHG (2013) Soulamarin isolated from *Calophyllum brasiliense* (Clusiaceae) induces plasma membrane

permeabilization of *Trypanosoma cruzi* and mytochondrial dysfunction. PLoS Negl Trop Dis 7:e2556. doi:10.1371/journal.pntd.0002556

Reitz R, Klein RM (1978) Projeto madeira de Santa Catarina. Sellowia 30:28–30

Rodríguez ME, Cardozo A, Krauczuk E, Fontana J, Iriart Y (2009) *Calophyllum brasiliense* (CLUSIACEAE) nuevo registro para la flora de la Argentina. Bol Soc Argent Bot 44:361–366

Salgueiro F, Neri J, Alves-Ferreira M, Scarano F (2011) Phylogeographic patterns of *Calophyllum brasiliense* Camb. (Calophyllaceae) based on the psbA-trnH cpDNA locus. BMC Proc 5(7):17

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, SchillingEE 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

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am J Bot 94:275–288

Shaw J, Shafer HL, Leonard OR, Kovach MJ, Schorr M, Morris AB (2014) Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. Am J Bot 101:1987–2004

Stange C, Prehn D, Arce-Johnson P (1998) Isolation of *Pinus radiate* genomic DNA suitable for RAPD analysis. Plant Mol Biol Rep 1:1–8

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplifications of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. NuclAcids Res 35:e14

Weising K, Gardner R (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42:9–19

Zheng Z, Schwartz L, Wagner L, Miller W (2000) A greedy algorithm for aligning DNA sequence. J Comput Biol 7:203–214