Lineage-Specific Evolutionary Rate in Plants: Contributions of a Screening for Cereus (Cactaceae)

Authors: Romeiro-Brito, Monique, Moraes, Evandro M., Taylor, Nigel P., Zappi, Daniela C., and Franco, Fernando F.

Source: Applications in Plant Sciences, 4(1)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500074

Applications in Plant Sciences 2016 4(1): 1500074

Application Article

Lineage-specific evolutionary rate in plants: Contributions of a screening for *Cereus* (Cactaceae)[1]

Monique Romeiro-Brito[2], Evandro M. Moraes[2], Nigel P. Taylor[3], Daniela C. Zappi[4], and Fernando F. Franco[2,5]

[2]Departamento de Biologia, Universidade Federal de São Carlos, Rodovia João Leme dos Santos Km 110, 18052780 Sorocaba, São Paulo, Brazil; [3]National Parks Board, Singapore Botanic Gardens, 1 Cluny Road, Singapore 259569, Singapore; and [4]Conservation Department, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, United Kingdom

- *Premise of the study:* Predictable chloroplast DNA (cpDNA) sequences have been listed for the shallowest taxonomic studies in plants. We investigated whether plastid regions that vary between closely allied species could be applied for intraspecific studies and compared the variation of these plastid segments with two nuclear regions.
- *Methods:* We screened 16 plastid and two nuclear intronic regions for species of the genus *Cereus* (Cactaceae) at three hierarchical levels (species from different clades, species of the same clade, and allopatric populations).
- *Results:* Ten plastid regions presented interspecific variation, and six of them showed variation at the intraspecific level. The two nuclear regions showed both inter- and intraspecific variation, and in general they showed higher levels of variability in almost all hierarchical levels than the plastid segments.
- *Discussion:* Our data suggest no correspondence between variation of plastid regions at the interspecific and intraspecific level, probably due to lineage-specific variation in cpDNA, which appears to have less effect in nuclear data. Despite the heterogeneity in evolutionary rates of cpDNA, we highlight three plastid segments that may be considered in initial screenings in plant phylogeographic studies.

Key words: Cactaceae; *Cereus*; cpDNA; evolutionary rate heterogeneity; *isi1*; nDNA; *nhx1*.

Chloroplast DNA (cpDNA) sequences have historically been used for both phylogenetic and phylogeographic studies in plants (Soltis et al., 1997; Schaal et al., 1998; Schaal and Olsen, 2000; Shaw et al., 2005; Avise, 2009). Despite some advantages (i.e., uniparental inheritance, genome stability, and rare recombination; Clegg et al., 1994; Ravi et al., 2007; Pleines et al., 2008; Borsch and Quandt, 2009), there are also limitations in using cpDNA at the shallowest taxonomic level (i.e., low genetic variability and evolutionary rate heterogeneity between lineages; Korotkova et al., 2014; Shaw et al., 2014). Nuclear regions are usually combined with cpDNA data, considering that these regions usually have a higher evolutionary rate in these organisms (Wolfe et al., 1987). Low-copy or single-copy nuclear genes (LCGs) are promising markers to perform comparative studies in plants and, frequently, LCGs present higher levels of variability than the traditional plastid markers at the shallowest taxonomic level (e.g., Sang, 2002; Small et al., 2004; Naumann et al., 2011).

[1]Manuscript received 29 June 2015; revision accepted 1 December 2015.

We are particularly grateful to G. Olsthoorn for supplying some of the samples of taxa used in this study. We are also grateful to H. S. M. Utsunomiya for technical assistance and to the anonymous referees and I. A. S. Bonatelli for critical suggestions. This work was supported by grants from the Fundação de Amparo à Pesquisa do Estado da São Paulo (FAPESP) to M.R.B. (2013/07211-7), F.F.F. (2010/19557-7), and to E.M.M. (2005/55200-8).

[5]Author for correspondence: franco@ufscar.br

doi:10.3732/apps.1500074

Cereus Mill. (Cactaceae) represents a Neotropical, long-lived, succulent taxon comprising approximately 30 species mainly distributed in South America and currently subdivided into four subgenera: *Cereus*, *Ebneria* (Backeb.) D. R. Hunt, *Mirabella* (F. Ritter) N. P. Taylor, and *Oblongicarpi* (Croizat) D. R. Hunt & N. P. Taylor (Hunt et al., 2006). *Cereus* presents some characteristics of Cactaceae taxa that make the genus a candidate for conducting evolutionary studies in relation to: (1) their specialization for xeric habitats (Arakaki et al., 2011; Hernández-Hernández et al., 2011; Bonatelli et al., 2014), being potentially informative on the causal effects of Pleistocene paleoclimatic changes under Neotropical biota (Majure et al., 2012; Bonatelli et al., 2014); (2) the diversity of environments where species occur, facilitating adaptation studies; and (3) their broad distribution across South American biomes (Taylor and Zappi, 2004; Hunt et al., 2006), permitting biogeographical studies on both continental and subcontinental scales.

As screening for variable markers has become the first mandatory step to perform phylogeographic studies in targeted species due to the high evolutionary rate heterogeneity between plant lineages (Shaw et al., 2014; Korotkova et al., 2014), our aim was to evaluate whether regions that are likely variable among closely allied species (i.e., those described by Shaw et al., 2007, 2014) could similarly be predicted as variable in intraspecific studies. We focused on three biological hierarchical levels to screen for potentially informative markers for *Cereus*: distinct clades in the phylogeny, different species of the same clade, and allopatric populations of the same species. We selected 16 plastid segments to collect empirical data, including those more likely to be variable as proposed by Shaw

et al. (2005, 2007, 2014): *rpS16*, *rpL16*, *trnL-trnF*, *trnT-trnL*, *3′trnK-matK* partial, *3′rps16–5′trnK*$^{(UUU)}$, *trnS-trnG*, *trnH-psbA*, *trnQ*$^{(UUG)}$*-5′rps16*, *rpl32-trnL*$^{(UAG)}$, *psbJ-petA*, *atpI-atpH*, *petL-psbE*, *psbD-trnT*$^{(GGU)}$, *3′trnV*$^{(UAC)}$*-ndhC*, and *ndhF-rpl32*. Furthermore, we included in our screening intronic regions of two nuclear genes: the impaired sucrose induction gene (*isi1*; Rook et al., 2006) and the Na^+/H^+ vacuolar antiporter gene (*nhx1*; Gaxiola et al., 1999), because these genes have been used successfully in comparative studies within Cactaceae (Franck et al., 2012, 2013). With this approach, we were able to identify regions suitable for intra- and interspecific studies in *Cereus*. In general, we found more variability in nuclear than in plastid regions. Moreover, our study suggests no correspondence between variation of plastid regions at the interspecific and intraspecific level, probably due to lineage-specific variation in cpDNA. This trend also appears to a lesser degree in nuclear sequences. The causes of lineage-specific evolutionary rates, especially in cpDNA, are discussed.

MATERIALS AND METHODS

***Sampling and DNA extraction*—**We screened for molecular variability across 12 species of three subgenera of *Cereus*, according to phylogenetic information currently available for the group (Table 1; Appendix S1). Genomic DNA from each sample was extracted from root tissue using the QIAGEN DNeasy Plant Mini Kit (Hilden, Nordrhein-Westfalen, Germany) and maintained at −20°C.

***DNA amplification and sequencing*—**Many relevant papers have contributed to the development of primers for the shallowest taxonomic level studies in plants (e.g., Taberlet et al., 1991; Savolainen et al., 1994; Hamilton, 1999; Muller and Borsch, 2005). However, efforts to identify and compare additional and predictive variable regions of the chloroplast genome throughout angiosperm groups were primarily undertaken by Shaw et al. (2005, 2007, 2014).

We selected 16 plastid molecular markers: 14 intergenic spacers (*trnT-trnL*, *3′trnK-matK* partial, *trnL-trnF*, *3′rps16-5′trnK*$^{(UUU)}$, *trnS-trnG*, *trnH-psbA*, *trnQ*$^{(UUG)}$*-5′rps16*, *rpl32-trnL*$^{(UAG)}$, *psbJ-petA*, *atpI-atpH*, *petL-psbE*, *psbD-trnT*$^{(GGU)}$, *3′trnV*$^{(UAC)}$*-ndhC*, and *ndhF-rpl32*) and two introns (*rpL16* and *rpS16*), which are considered the most variable regions on average for angiosperms as a whole in both Shaw et al. publications (Shaw et al., 2007, 2014: *3′rps16-5′trnK*$^{(UUU)}$, *trnQ*$^{(UUG)}$*-5′rps16*, *rpl32-trnL*$^{(UAG)}$, *psbJ-petA*, *atpI-atpH*, *petL-psbE*, *psbD-trnT*$^{(GGU)}$, *3′trnV*$^{(UAC)}$*-ndhC*, *ndhF-rpl32*), and also other variable markers previously used in Cactaceae (*trnT-trnL*, *3′trnK-matK* partial, *trnL-trnF*, *trnS-trnG*, *trnH-psbA*, and *rpL16*; Nyffeler, 2002; Bonatelli et al., 2013). The *rpS16* intron was included because it is among the most used in species-level studies due to the early development of universal primers for this region (Shneyer, 2009). The selected nuclear regions (*nhx1* and *isi1*) were previously used in Cactaceae (Franck et al., 2012, 2013), being potential regions for screening for variability. Each plastid and nuclear region was amplified using universal primers (Appendix S2).

PCR reactions for plastid regions were performed in a total volume of 15 μL containing 1 μL of genomic DNA (10–40 ng), 1× reaction buffer, 0.5–1 unit of Promega *Taq* DNA polymerase (Promega Corporation, Madison, Wisconsin, USA), and 200 μM dNTPs. The $MgCl_2$ and primer concentrations as well as the temperature conditions used are described in Appendix S3. For nuclear regions, PCR reactions were performed in a total volume of 30 μL containing 2 μL of genomic DNA, 1× reaction buffer containing $MgCl_2$ to 2 mM, and 1 unit of Q5 Hot Start High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, Massachusetts, USA). Amplification was performed using an Eppendorf Mastercycler (Eppendorf, Hamburg, Germany). PCR products were isolated on a 1% agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, California, USA) using the BioDoc-It 220 Imaging System (UVP, Upland, California, USA) and purified with ExoSap-IT PCR Product Cleanup (Affymetrix, Santa Clara, California, USA) or the illustra GFX Gel Band Purification Kit (GE Healthcare, Piscataway, New Jersey, USA). Sequencing was performed using an ABI PRISM 3730 DNA Analyzer (Applied Biosystems, Foster City, California, USA) with the BigDye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems). Forward and reverse sequences were compared and edited in Chromas Lite 2.0 software (Technelysium Pty Ltd., South Brisbane, Australia; www.technelysium.com.au). The sequences were aligned using ClustalW 1.8 software (Thompson et al., 1994) available in BioEdit (Hall, 1999).

***Descriptive analysis*—**We adopted three levels of variability assessment, with a minimum threshold of four potentially informative characters (PIC; Shaw et al., 2005) to progress to the next level of screening. We followed the respective levels of analyses according to phylogenetic information available for the group (Appendix S1):

- Clade level: species allocated into distinct clades of the *Cereus* phylogeny (Appendix S1), which roughly correspond to a comparison of distinctive subgenera: clade A (*C. fernambucensis* Lem., subgenus *Cereus*), clade C (*C. mirabella* N. P. Taylor, *C. albicaulis* (Britton & Rose) Luetzelb., subgenus *Mirabella*), clade B (*C. hankeanus* F. A. C. Weber ex K. Schum.,

TABLE 1. Species of *Cereus* and number of samples selected in the current study. Clades were delimited based on phylogenetic (Appendix S1) and taxonomic information. See Appendix 1 for GenBank accessions.

Species	Locality	Voucher[a]	Geographic coordinates	*N*	Clade
Subgenus *Cereus*					
C. hildmannianus K. Schum.	Salto de Itu, São Paulo, Brazil	SORO 2746	23°17′59″S, 47°19′50″W	1	A1
C. insularis Hemsl.	Fernando de Noronha, Pernambuco, Brazil	SORO 2677	03°50′15″S, 32°23′58″W	2	A2
C. fernambucensis Lem. subsp. *sericifer* (Ritter) N. P. Taylor & Zappi	Três Rios, Rio de Janeiro, Brazil	SORO 2662	22°00′38″S, 43°15′55″W	1	A2
	Água Branca, Espírito Santo, Brazil	SORO 2734	19°03′34″S, 40°41′07″W	1	A2
C. fernambucensis Lem. subsp. *fernambucensis*	Ubatuba, São Paulo, Brazil	SORO 2658	23°49′37″S, 45°25′21″W	1	A2
	São Mateus, Espírito Santo, Brazil	SORO 2669	18°45′18″S, 39°44′51″W	1	A2
	Maracajaú, Rio Grande do Norte, Brazil	SORO 2670	05°23′23″S, 35°18′42″W	1	A2
	Ilhéus, Bahia, Brazil	SORO 2675	15°06′17″S, 38°59′47″W	1	A2
Subgenus *Ebneria*					
C. hankeanus F. A. C. Weber ex K. Schum.	NA	SORO 2739	NA	1	B2
C. saddianus (Rizzini & A. Mattos) P. J. Braun	Cáceres, Mato Grosso, Brazil	SORO 3632	15°55′12″S, 57°30′03″W	1	C1
Subgenus *Mirabella*					
C. mirabella N. P. Taylor	São Domingos, Goiás, Brazil	SORO 2776	13°27′17″S, 46°19′12″W	1	D2
	Barreiras, Bahia, Brazil	SORO 2779	12°06′38″S, 45°15′34″W	1	D2
C. albicaulis (Britton & Rose) Luetzelb.	Una, Bahia, Brazil	SORO 2771	NA	1	D2
	Morro do Chapéu, Bahia, Brazil	SORO 3633	11°39′02″S, 41°17′20″W	1	D2

Note: *N* = number of samples; NA = not available.

[a]All plants were deposited in Centro de Ciências e Tecnologia para a Sustentabilidade (SORO) Herbarium.

subgenus *Ebneria*), and clade D (*C. saddianus* (Rizzini & Mattos) P. J. Braun, subgenus *Ebneria*), totaling five sequences analyzed in this level;
- Subclade level: species allocated in clade A (Appendix S1) and belonging to subgenus *Cereus* (*C. fernambucensis* subsp. *fernambucensis*, *C. fernambucensis* subsp. *sericifer* (Ritter) N. P. Taylor & Zappi, *C. hildmannianus* K. Schum., *C. insularis* Hemsl.), totaling four sequences analyzed in this level. This level roughly corresponds to a comparison within the same subgenus;
- Species level: allopatric populations of *C. fernambucensis* (subspecies: four samples from different populations of *C. fernambucensis* subsp. *fernambucensis*, two samples from different populations of *C. fernambucensis* subsp. *sericifer*), totaling six sequences analyzed in this level.

Standard indices of variability were considered with all sequences obtained for each level of analysis. Nucleotide diversity (π), number of haplotypes (h), polymorphic sites (S), and average number of nucleotide differences (k) were calculated for each level of analysis in DnaSP version 5 (Librado and Rozas, 2009). The percentage of variability was calculated using the formula (PIC/L) × 100, where L is the total length of the sequence. We used a simple model of genetic distance (average *p*-distance and between group *p*-distance) among the levels calculated in MEGA 5.1 (Tamura et al., 2011). More complex models of genetic distance were also tested in our sampling and presented similar results to the simple model used in this study (data not shown).

RESULTS

Three of the 16 segments screened in *Cereus* (*3′trnV*$^{(UAC)}$*-ndhC*, *ndhF-rpl32*, and *rpl32-trnL*$^{(UAG)}$) could not be amplified, even after several attempts to modify the PCRs (Appendix S3). This is not an uncommon result for universal primers such as *ndhF-rpl32* and *rpl32-trnL*$^{(UAG)}$ (Prince, 2015) due the high variability of these regions, which prevents the design of universal primers (Shaw et al., 2014). For some Cactaceae, for example, to use this segment it has been necessary to design a new primer set for focal taxa (Calvente et al., 2011; Majure et al., 2012).

The results for the remaining 13 cpDNA segments are as follows: (1) the *trnL-trnF*, *trnT-trnL*, and *3′trnK-matK* partial showed no or low variation at the clade level; (2) the segments *trnH-psbA*, *rpS16*, and *3′rps16-5′trnK*$^{(UUU)}$ were variable only at the clade level; (3) the segment *atpI-atpH* was variable only at the clade and subclade levels; and (4) the *rpL16* intron and the segments *trnQ*$^{(UUG)}$*-5′rps16*, *petL-psbE*, *psbD-trnT*$^{(GGU)}$, *psbJ-petA*, and *trnS-trnG* were variable at all three levels of analysis (Table 2). The most variable segment at the clade and subclade levels was the *trnQ*$^{(UUG)}$*-5′rps16* segment, whereas *trnS-trnG* showed the highest PIC value among plastid regions at the species level (Table 2, Fig. 1).

The main sources of variation in cpDNA markers were insertion-deletions and nucleotide substitutions. The exception was the *rpS16* intron, which showed an inversion in the *C. saddianus* sample. Inversions are difficult to analyze because they are not well recognized in the alignment, especially in small sample sizes (Borsch and Quandt, 2009). Recent studies have discovered that this mutation is usually common in noncoding regions (Borsch and Quandt, 2009), even at population level or in lineage-specific analyses (Quandt et al., 2003; Borsch and Quandt, 2009; Korotkova et al., 2014). Inversions are usually associated with inverted repeat sequences, yielding a hairpin secondary structure such as predicted in our sampling (Appendix S4). To investigate the presence of this mutation in *C. saddianus* populations, we increased the number of samples. We observed some intraspecific variation in *C. saddianus* populations, detecting a 35-bp inversion as well as other variable characters (Appendix S4).

The nuclear regions were variable on all three levels of analysis, being potentially informative for phylogenetic and phylogeographic studies, at least in the genus *Cereus* (Fig. 1, Table 2). Furthermore, one of the nuclear regions (*isi1*) presented higher variability than plastid data in almost all levels of analysis (except for the first level of analysis where *trnQ*$^{(UUG)}$*-5′rps16* was the most variable segment).

Genetic distance data calculated for those segments that were variable at the three levels of analysis suggest heterogeneity across levels and occasionally presented higher genetic distance within species than among species. This result clearly suggests that evolutionary rates vary among closely related species as well as within population units (Fig. 2). Furthermore, the plastid markers presented a higher evolutionary rate heterogeneity than nuclear regions, suggesting that plastid regions may likely be more affected than nuclear regions.

DISCUSSION

The three levels of analysis established in this study allowed us to analyze the potential utility of each region according to the distinct hierarchical levels defined in the *Cereus* phylogeny. From the 13 screened plastid regions successfully amplified in our sample, we identified three potentially informative markers for interspecific studies in *Cereus* (*atpI-atpH*, *trnH-psbA*, and *3′rps16-5′trnK*$^{(UUU)}$) and eight potentially informative markers for interspecific and intraspecific studies (six plastid regions: *trnS-trnG*, *psbD-trnT*$^{(GGU)}$, *petL-psbE*, *trnQ*$^{(UUG)}$*-5′rps16*, *psbJ-petA*, and *rpL16*; and two nuclear regions: *nhx1* and *isi1*).

Most of these regions were previously used in higher taxonomic studies in Cactaceae (e.g., Nyffeler, 2002; Calvente et al., 2011; Hernández-Hernández et al., 2011). Although most of the segments that revealed variability in the three levels of analysis were previously used in cactus studies to establish evolutionary relationships for phylogeographic analyses (*trnQ*$^{(UUG)}$*-5′rps16*, *psbJ-petA*, and *rpL16* [Korotkova et al., 2011]; *trnS-trnG* [Bonatelli et al., 2013]), some of the plastid regions previously used at the intraspecific level did not present enough variation in this work (*trnT-trnL* and *psbA-trnH* [Korotkova et al., 2011; Bonatelli et al., 2013]).

The segments *rpl32-trnL*$^{(UAG)}$, *ndhF-rpl32*, and *trnV-ndhC* were considered among the best ranked for angiosperms in general, but could not be amplified in this work. Segment *rpl32-trnL*$^{(UAG)}$ has been successfully used in evolutionary studies with Cactaceae (Larridon et al., 2015), but due to its high variability in the flanking region (Shaw et al., 2014) it is usually necessary for a new set of primers to be developed for a respective target group (e.g., Calvente et al., 2011; Majure et al., 2012). It is likely that *rpl32-trnL*$^{(UAG)}$ should be highly variable and useful for comparative studies, such as shallow phylogenetic (Miller et al., 2009; Calvente et al., 2011; Ornelas and Rodriguez-Gomez, 2015) and phylogeographic studies (Jiménez-Mejías et al., 2012; Aguirre-Liguori et al., 2014; Sramkó et al., 2014). However, due to the lack of suitable universal primers (Prince, 2015) its potential has not been explored.

It is worth noting that nuclear regions presented high variability when compared with plastid DNA (Fig. 1), thus becoming promising regions to perform inter- and intraspecific studies in *Cereus*. This is in agreement with the premise that nuclear markers frequently show more variation in plants than plastid markers (Sang, 2002; Small et al., 2004; Zimmer and Wen, 2012). For most of the cpDNA variable regions, we did not observe the expected increase in genetic differentiation among higher taxonomic evolutionary levels. In contrast, for the *rpS16* intron, for example, we detected higher variation at the population

Table 2. Variation in plastid and nuclear regions at each level of analysis.

Variability indices	*trnT-trnL*	*trnL-trnF*	*3′trnK-matK*	*trnH-psbA*	*rpS16*	*atpI-atpH*	*3′rps16-5′trnK(UUU)*	*psbD-trnT(GGU)*	*petL-psbE*	*rpL16*	*psbJ-petA*	*trnS-trnG*	*trnQ(UUG)-5′rps16*	*nhx1*	*isi1*
Clade level															
N	4	5	4	5	5	5	5	5	5	5	5	5	4	5	4
S	1	1	2	3	3	3	4	11	13	15	32	51	93	21	41
PIC	1	2	3	4	5	5	6	13	13	19	35	54	109	24	43
h/N	1/4	2/5	3/4	4/5	4/5	5/5	4/5	4/5	5/5	4/5	5/5	5/5	4/4	5/5	5/5
π	0	0.001	0.001	0.004	0.001	0.002	0.005	0.007	0.12	0.009	0.04	0.025	0.08	0.018	0.019
k	0	0.4	1.17	1.2	1.2	1.3	2.6	5	6.2	7.92	16.3	23.57	49.17	8.67	17
%	0	0.48	0.25	1.2	0.625	0.928	1.1	1.94	2.57	2.29	8.56	5.66	18.2	4.43	4.81
Subclade level															
N	—	—	—	4	4	4	4	4	4	4	4	4	3	4	3
S	—	—	—	1	0	6	0	3	4	13	5	12	16	14	33
PIC	—	—	—	2	1	6	0	4	5	16	7	13	18	15	35
h/N	—	—	—	3/4	2/4	3/4	1/4	3/4	4/4	4/4	4/4	4/4	3/3	4/4	3/3
π	—	—	—	0.002	0	0.006	0	0.002	0.004	0.008	0.008	0.007	0.019	0.018	0.020
k	—	—	—	0.67	0	3	0	1.5	2.17	6.5	3.5	6.5	11.67	8.5	17.67
%	—	—	—	0.6	0.13	1.13	0	0.6	0.99	1.94	1.63	1.36	2.99	4.44	3.9
Species level															
N	—	—	—	—	—	6	—	6	5	6	6	6	4	6	4
S	—	—	—	—	—	2	—	7	7	6	3	12	8	15	28
PIC	—	—	—	—	—	2	—	7	7	10	5	12	11	15	31
h/N	—	—	—	—	—	3/6	—	4/6	4/6	4/6	6/6	5/6	4	4/4	4/4
π	—	—	—	—	—	0.002	—	0.003	0.005	0.004	0.003	0.005	0.001	0.018	0.020
k	—	—	—	—	—	0.87	—	2.3	2	3.4	1.27	5.2	5.33	8.5	17.83
%	—	—	—	—	—	0.37	—	1.04	1.39	1.23	1.18	1.26	2.38	4.44	3.47

Note: *N* = sample used for each level; *S* = polymorphic sites; PIC = potentially informative characters (Shaw et al., 2005); *h/N* = number of haplotypes by number of samples; π = nucleotide diversity; *k* = average number of nucleotide differences; % = percentage of variability.

level than in subclade or clade levels (Fig. 2). These results agree with the high heterogeneity in molecular evolutionary rates in plants (Korotkova et al., 2014; Shaw et al., 2014); however, they also suggest that lineage-specific variation in cpDNA seems to be more accentuated than in nuclear DNA. Furthermore, these data indicate that rate heterogeneity in cpDNA evolution appears to increase in early stages of population differentiation (Duchene and Bromham, 2013; Bromham et al., 2015).

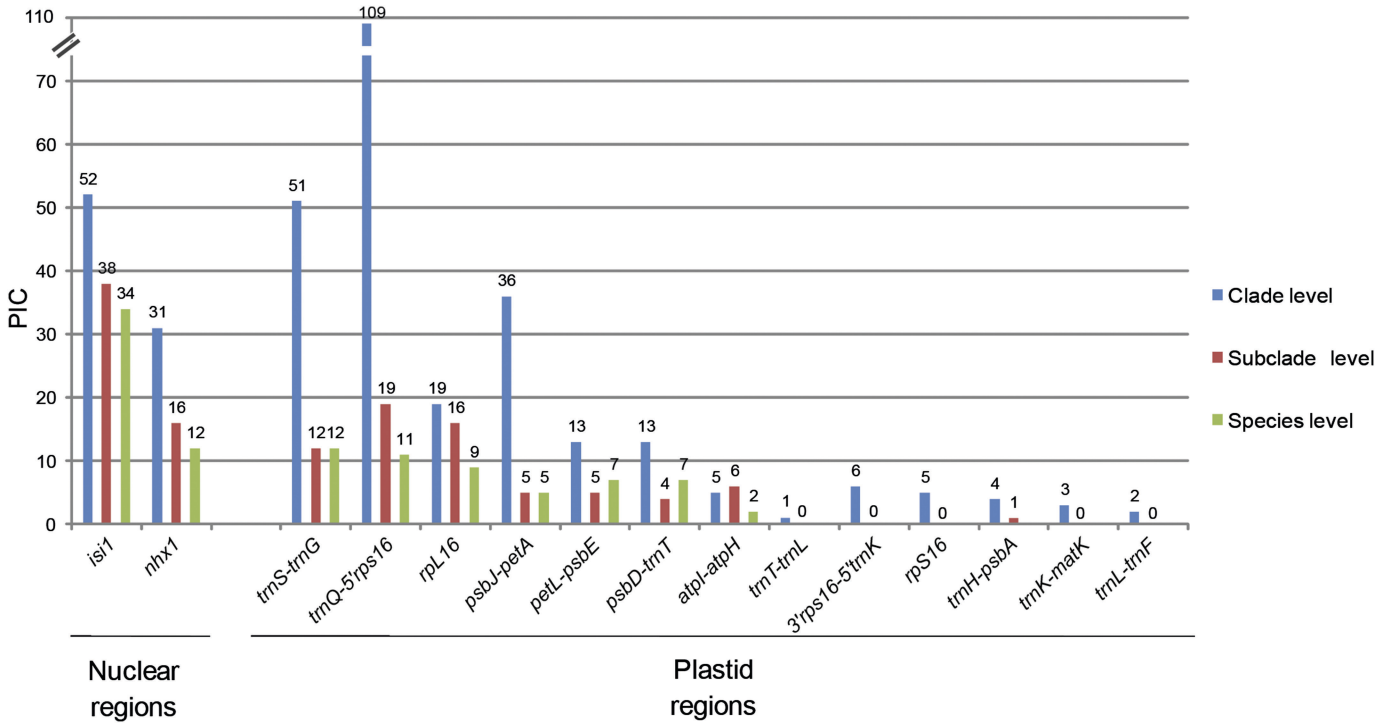

Fig. 1. PIC values of plastid and nuclear molecular markers successfully amplified in this study.

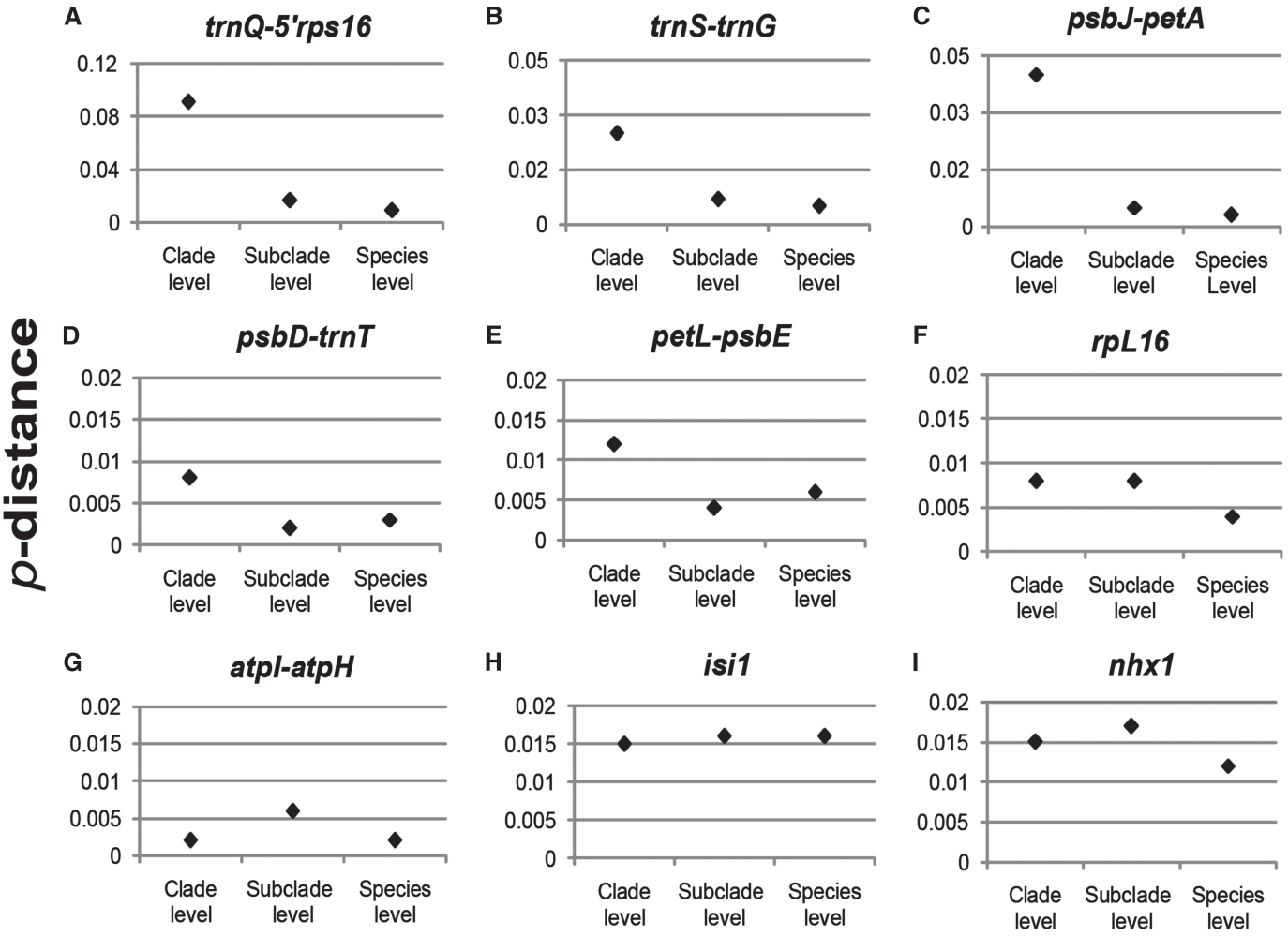

Fig. 2. The *p*-distance of the seven plastid segments (A–G) and two nuclear regions (H–I) were explored at three levels of analysis. The three markers (A–C) presented higher values of *p*-distance at the clade level (>0.05). Some regions, such as *petL-psbE*, *psbD-trnT*$^{(GGU)}$, and *atpI-atpH*, are examples of how heterogeneity occasionally varies along the levels of analysis. It is clear how abruptly the genetic distance can change through levels based on *trnS-trnG*, *psbJ-petA*, and *trnQ*$^{(UUG)}$*-5′rps16* patterns. Although nuclear regions also present heterogeneity variation among levels, they seem less prone to lineage-specific effects. Between-group genetic distance (*p*-distance) was calculated to obtain pairwise distance considering groups of taxa at each level of analysis (Appendix S5).

Non-clock events and punctuated molecular evolution occur widely in nature (Clegg et al., 1994; Pagel et al., 2006) and are extremely common in plants (Pagel et al., 2006), most likely due to patterns of hybridization, polyploidy, and gene duplication in these organisms (Pagel et al., 2006; Duchene and Bromham, 2013). Other explanations have been proposed for punctuated effects, including ecological traits such as life history (i.e., generation time associated with life forms such as shrubs, trees, and herbaceous plants; annual and perennial life cycles [Bromham, 2009]), environmental variables (i.e., temperature and UV radiation [Lancaster, 2010; Gaut et al., 2011]), and microevolutionary processes that occur during population differentiation (i.e., the founder effect and natural selection [Barraclough and Savolainen, 2001; Pagel et al., 2006; Duchene and Bromham, 2013; see also Lancaster, 2010; Pennell et al., 2014]).

Recently, Duchene and Bromham (2013) and Bromham et al. (2015) have reported faster rates of molecular evolution in plastid genes in species-rich lineages. Such observations suggest that cytonuclear interactions among plastid and nuclear genes promote hybrid incompatibility and sterility and possibly accelerate the number of substitutions in a single lineage, thus promoting hybrid incompatibility (Bromham et al., 2015). These interactions may indirectly drive the divergence of neutral chloroplast regions, resulting in heterogeneity among lineages (e.g., Shaw et al., 2014).

Although the causes of lineage-specific rates are under discussion, the possibility that rate heterogeneity in molecular markers in plants, especially those from cpDNA, may initiate early during diversification should be considered in the experimental design used to perform empirical screening for variation. In general, most screenings first compare allied species and then suggest the use of the best markers for the intraspecific level (e.g., Miller et al., 2009; Dong et al., 2012). For example, this is the design used in the publications by Shaw et al. (2005, 2007, 2014). Undoubtedly, this strategy may work, but our data agree that such interspecific comparisons do not ensure variability at the intraspecific level. For this reason, we recommend that the

experimental design of the initial screening should consider the biological unit(s) (i.e., species, subspecies, and populations) that is as close as possible to the target unit(s) to be studied.

In summary, our data indicate that intronic regions of nuclear genes *isi1* and *nhx1* are candidate markers for comparative studies in cacti, in concordance with previous data (Franck et al., 2012, 2013). The plastid segments *trnQ*$^{(UUG)}$*-5′rps16*, *rpL16*, and *trnS-trnG* showed higher levels of variability among the cpDNA markers tested at the population level for *Cereus*, becoming candidate markers for further phylogeographic studies in this genus, which is broadly distributed in South America. It is worth noting that despite the lack of universality in cpDNA variability, these three markers were considered potentially variable regions as reviewed by Shaw et al. (2005, 2007, 2014), and when these three regions were screened together, at least one of them showed useful information (e.g., Byrne and Hankinson, 2012; Martinez-Nieto et al., 2013). Thus, we suggest that these three segments are candidate regions to be included in initial variation screening for plant phylogeographic studies.

LITERATURE CITED

Aguirre-Liguori, J. A., E. Scheinvar, and L. E. Eguiarte. 2014. Gypsum soil restriction drives genetic differentiation in *Fouquieria shrevei* (Fouquieriaceae). *American Journal of Botany* 101: 730–736.

Arakaki, M., P. Christin, R. Nyffeler, A. Lendel, U. Eggli, R. M. Ogburn, E. Spriggs, et al. 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. *Proceedings of the National Academy of Sciences, USA* 108: 8379–8384.

Avise, J. C. 2009. Phylogeography: Retrospect and prospect. *Journal of Biogeography* 36: 3–15.

Barraclough, T. G., and V. Savolainen. 2001. Evolutionary rates and species diversity in flowering plants. *Evolution* 55: 677–683.

Bonatelli, I. A. S., D. C. Zappi, N. P. Taylor, and E. M. Moraes. 2013. Usefulness of cpDNA markers for phylogenetic and phylogeographic analyses of closely related cactus species. *Genetics and Molecular Research* 12: 4579–4585.

Bonatelli, I. A. S., M. F. Perez, A. T. Peterson, N. P. Taylor, D. C. Zappi, M. C. Machado, I. Koch, et al. 2014. Interglacial microrefugia and diversification of a cactus species complex: Phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology* 23: 3044–3063.

Borsch, T., and D. Quandt. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolution* 282: 169–199.

Bromham, L. 2009. Why do species vary in their rate of molecular evolution? *Biology Letters* 5: 401–404.

Bromham, L., X. Hua, R. Lanfear, and P. F. Cowman. 2015. Exploring the relationships between mutation rates, life history, genome size, environment, and species richness in flowering plants. *American Naturalist* 185: 507–524.

Byrne, M., and M. Hankinson. 2012. Testing the variability of chloroplast sequences for plant phylogeography. *Australian Journal of Botany* 60: 569–574.

Calvente, A., D. C. Zappi, F. Forest, and L. G. Lohmann. 2011. Molecular phylogeny of tribe Rhipsalidae (Cactaceae) and taxonomic implications for *Schlumbergera* and *Hartiora. Molecular Phylogenetics and Evolution* 58: 456–468.

Clegg, M. T., B. S. Gaut, G. H. Learn, and B. R. Morton. 1994. Rates and patterns of chloroplast DNA evolution. *Proceedings of the National Academy of Sciences, USA* 91: 6795–6801.

Dong, W., J. Liu, J. Yu, L. Wang, and S. Zhou. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One* 7: e35071.

Duchene, D., and L. Bromham. 2013. Rates of molecular evolution and diversification in plants: Chloroplast substitution rates correlated with species-richness in the Proteaceae. *BMC Evolutionary Biology* 13: 65.

Franck, A. R., B. J. Cochrane, and J. R. Garey. 2012. Low-copy nuclear primers and *ycf1* primers in Cactaceae. *American Journal of Botany* 99: 405–407.

Franck, A. R., B. J. Cochrane, and J. R. Garey. 2013. Phylogeny, biogeography, and infrageneric classification of *Harrisia* (Cactaceae). *Systematic Botany* 38: 210–223.

Gaut, B., L. Yang, S. Takuno, and L. E. Eguiarte. 2011. The patterns and causes of variation in plant nucleotide substitution rates. *Annual Review of Ecology Evolution and Systematics* 42: 245–266.

Gaxiola, R. A., R. Rao, A. Sherman, P. Grisafi, S. L. Alper, and G. R. Fink. 1999. The *Arabidopsis thaliana* proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast. *Proceedings of the National Academy of Sciences, USA* 96: 1480–1485.

Hall, T. A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/ NT. *Nucleic Acids Symposium Series* 41: 95–98.

Hamilton, M. B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology Notes* 8: 521–523.

Hernández-Hernández, T., H. M. Hernández, J. A. De-Nova, R. Puente, L. E. Eguiarte, and S. Magallón. 2011. Phylogenetic relationships and evolution of growth form in Cactaceae (Caryophyllales, Eudicotyledoneae). *American Journal of Botany* 98: 44–61.

Hunt, D., N. Taylor, and G. Charles. 2006. The new cactus lexicon. DH Books, Milborne Port, United Kingdom.

Jiménez-Mejías, P., M. Luceño, K. A. Lye, C. Brochmann, and G. Gussarova. 2012. Genetically diverse but with surprisingly little geographical structure: The complex history of the widespread herb *Carex nigra* (Cyperaceae). *Journal of Biogeography* 39: 2279–2291.

Korotkova, N., T. Borsch, D. Quandt, N. P. Taylor, K. F. Müller, and W. Barthlott. 2011. What does it take to resolve relationships and to identify species with molecular markers? An example from the epiphytic Rhipsalideae (Cactaceae). *American Journal of Botany* 98: 1549–1572.

Korotkova, N., L. Nauheimer, H. Ter-Voskanyan, M. Allgaier, and T. Borsch. 2014. Variability among the most rapidly evolving plastid genomic regions is lineage-specific: Implications of pairwise genome comparisons in *Pyrus* (Rosaceae) and other angiosperms for marker choice. *PLoS One* 9: e112998.

Lancaster, L. T. 2010. Molecular evolutionary rates predict both extinction and speciation in temperate angiosperm lineages. *BMC Evolutionary Biology* 10: 162.

Larridon, I., H. E. Walter, P. C. Guerrero, M. Duarte, M. A. Cisternas, C. P. Hernandez, K. Bauters, et al. 2015. An integrative approach to understanding the evolution and diversity of *Copiapoa* (Cactaceae), a threatened endemic Chilean genus from the Atacama Desert. *American Journal of Botany* 102: 1506–1520.

Librado, P., and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics (Oxford, England)* 25: 1451–1452.

Majure, L. C., R. Puente, M. P. Griffith, W. S. Judd, P. S. Soltis, and D. E. Soltis. 2012. Phylogeny of *Opuntia* s.s. (Cactaceae): Clade delineation, geographic origins, and reticulate evolution. *American Journal of Botany* 99: 847–864.

Martínez-Nieto, M. I., J. G. Segarra-Moragues, E. Merlo, F. Martínez-Hernández, and J. F. Mota. 2013. Genetic diversity, genetic structure and phylogeography of the Iberian endemic *Gypsophila struthium* (Caryophyllaceae) as revealed by AFLP and plastid DNA sequences: Connecting habitat fragmentation and diversification. *Botanical Journal of the Linnean Society* 173: 654–675.

Miller, J. S., A. Kamath, and R. Levin. 2009. Do multiple tortoises equal a hare? The utility of nine noncoding plastid regions for species-level phylogenetics in Tribe Lycieae (Solanaceae). *Systematic Botany* 34: 796–804.

Müller, K., and T. Borsch. 2005. Phylogenetics of Amaranthaceae based on *matk/trnK* sequence data: Evidence for parsimony, likelihood and Bayesian analyses. *Annals of the Missouri Botanical Garden* 92: 66–102.

Naumann, J., L. Symmank, M. Samain, K. F. Müller, C. Neinhuis, C. W. de Pamphilis, and S. Wanke. 2011. Chasing the hare: Evaluating the phylogenetic utility of a nuclear single copy gene region at and below species level within the species rich group *Peperomia* (Piperaceae). *Evolutionary Biology* 11: 357.

Nyffeler, R. 2002. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from *trnK-matK* and *trnL-trnF* sequences. *American Journal of Botany* 89: 312–326.

Ornelas, J. F., and F. Rodríguez-Gómez. 2015. Influence of Pleistocene glacial/interglacial cycles on the genetic structure of the mistletoe cactus *Rhipsalis baccifera* (Cactaceae) in Mesoamerica. *Journal of Heredity* 106: 196–210.

Pagel, M., C. Venditti, and A. Meade. 2006. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science* 314: 119–121.

Pennell, M. W., L. J. Harmon, and J. C. Uyeda. 2014. Speciation is unlikely to drive divergence rates. *Trends in Ecology & Evolution* 29: 72–73.

Pleines, T., S. S. Jakob, and F. R. Blattner. 2008. Application of non-coding DNA regions in intraspecific analyses. *Plant Systematics and Evolution* 282: 281–294.

Prince, L. 2015. Plastid primers for angiosperm phylogenetics and phylogeography. *Applications in Plant Sciences* 3: 1400085.

Quandt, D., K. Müller, and S. Huttunen. 2003. Characterisation of the chloroplast DNA *psb*T-H region and the influence of dyad symmetrical elements on phylogenetic reconstructions. *Plant Biology* 5: 400–410.

Ravi, V., J. P. Khurana, A. K. Tyagi, and P. Khurana. 2007. An update on chloroplast genomes. *Plant Systematics and Evolution* 271: 101–122.

Rook, F., F. Corke, M. Baier, R. Holman, A. G. May, and M. W. Bevan. 2006. *Impaired sucrose induction1* encodes a conserved plant-specific protein that couples carbohydrate availability to gene expression and plant growth. *Plant Journal* 46: 1045–1058.

Sang, T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology* 37: 121–147.

Savolainen, V., J. F. Manen, E. Douzery, and R. Spichiger. 1994. Molecular phylogeny of families related to Celastrales based on *rbcL* 5′ flanking sequences. *Molecular Phylogenetics and Evolution* 3: 27–37.

Schaal, B. A., D. A. Hayworth, K. M. Olsen, J. T. Rauscher, and W. A. Smith. 1998. Phylogeography studies in plants: Problems and prospects. *Molecular Ecology* 7: 465–474.

Schaal, B., and K. Olsen. 2000. Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences, USA* 97: 7024–7029.

Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Siripun, et al. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142–166.

Shaw, J., E. B. Lickey, E. E. Schilling, and R. L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* 94: 275–288.

Shaw, J., H. L. Shafer, O. R. Leonard, M. J. Kovach, M. Schorr, and A. B. Morris. 2014. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: The tortoise and the hare IV. *American Journal of Botany* 101: 1987–2004.

Shneyer, V. S. 2009. DNA barcoding is a new approach in comparative genomics of plants. *Russian Journal of Genetics* 45: 1267–1278.

Small, R. L., R. C. Cronn, and J. F. Wendel. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17: 145–170.

Soltis, D. E., M. A. Gitzendanner, D. D. Strenge, and P. S. Soltis. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution* 206: 353–373.

Sramkó, G., A. Molnar, J. A. Hawkins, and R. M. Bateman. 2014. Molecular phylogeny and evolutionary history of the Eurasiatic orchid genus *Himantoglossum* s.l. (Orchidaceae). *Annals of Botany* 114: 1609–1626.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731–2739.

Taylor, N. P., and D. Zappi. 2004. Cacti of eastern Brazil. Royal Botanic Gardens, Kew, Richmond, Surrey, United Kingdom.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.

Wolfe, K. H., W. H. Li, and P. M. Sharp. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences, USA* 84: 9054–9058.

Zimmer, E. A., and J. Wen. 2012. Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molecular Phylogenetics and Evolution* 65: 774–785.

Appendix 1. Origin of the material used in this analysis and GenBank accession numbers. NS = not sampled. Taxon; population code; GenBank accessions: *trnS-trnG*, *rpL16*, *psbJ-petA*, *atpI-atpH*, *psbD-trnT*$^{(GGU)}$, *petL-psbE*, *trnQ*$^{(UUG)}$*-5′rps16*, *trnH-psbA*, *rpS16*, *3′rps16-5′trnK*$^{(UUU)}$, *trnT-trnL*, *trnL-trnF*, *3′trnK-matK*, *nhx1*, *isi1*.

C. hildmannianus* K. Schum.**; S90F; KP017425, KR998134, KR998096, KR998187, KR998083, KR998114, NS, KR998090, KR998157, KR998169, KR998125, NS, NS, KT984782, NS. ***C. insularis* Hemsl.**; S77A29; KP017432, KR998135, KR998105, KR998179, KR998082, NS, NS, NS, KR998158, NS, NS, NS, NS, NS, NS. ***C. insularis* Hemsl.**; S115A; NS, NS, NS, NS, NS, KR998113, KR998147, KR998091, NS, KR998170, NS, NS, NS, KT984781, KT984772. ***C. fernambucensis* Lem. subsp. *sericifer* (Ritter) N. P. Taylor & Zappi**; S76F; KP017431, KR998133, KR998102, KR998178, KR998081, KR998115, KR998143, KR998089, KR998156, KR998171, KR998127, NS, KR998173, KT984775, KT984771. ***C. fernambucensis* Lem. subsp. *sericifer* (Ritter) N. P. Taylor & Zappi**; S88F; KR998162, KR998136, KR998097, KR998183, KR998084, KR998116, KR998142, NS, NS, NS, NS, NS, NS, KT984776, NS. ***C. fernambucensis* Lem. subsp. *fernambucensis; S72F; KP017430, KR998128, KR998101, KR998177, KR998076, KR998108, KR998140, KR998088, NS, KR998168, KR998126, KR998119, KR998172, KT984777, KT984767. ***C. fernambucensis* Lem. subsp. *fernambucensis***; S89F; KR998163, KR998137, KR998098, KR998182, KR998085, KR998117, NS, NS, NS, NS, NS, NS, NS, KT984778, NS. ***C. fernambucensis* Lem. subsp. *fernambucensis***; S104F; KR998160, KR998138, KR998099, KR998181, KR998086, KR998118, KR998141, NS, NS, NS, NS, NS, NS, KT984780, KT984773. ***C. fernambucensis* Lem. subsp. *fernambucensis***; S114F; KR998161, KR998139, KR998100, KR998188, KR998087, NS, NS, NS, NS, NS, NS, NS, NS, KT984779, KT984774. ***C. hankeanus* F. A. C. Weber ex K. Schum.**; S77A31; KR998159, KR998132, KR998104, KR998186, KR998080, KR998111, KR998146, KR998095, KR998095, KR998167, NS, KR998122, KR998175, KT984785, KT984770. ***C. saddianus* (Rizzini & A. Mattos) P. J. Braun**; S103D; KP017445, KR998131, KR998103, KR998180, KR998079, KR998112, KR998145, NS, KR998148 and KR998149, KR998166, NS, KR998120, KR998176. ***C. saddianus* (Rizzini & A. Mattos) P. J. Braun**; S99; NS, NS, NS, NS, NS, NS, NS, NS, KR998150 and KR998153, NS, NS, NS, NS, KT984783, NS. ***C. mirabella* N. P. Taylor**; S77A19; KP017444, KR998130, NS, NS, KR998078, KR998110, KR998144, NS, NS, KR998164, KR998124, NS, KR998174, NS, NS. ***C. mirabella* N. P. Taylor**; S77A18; NS, NS, KR998107, KR998185, NS, NS, NS, KR998093, NS, NS, NS, NS, NS, KT984786, KT984769. ***C. albicaulis* (Britton & Rose) Luetzelb.**; S77A17; KP017442, KR998129, NS, KR998184, KR998077, KR998109, NS, NS, KR998154, KR998165, NS, KR998121, NS, NS, NS. ***C. albicaulis* (Britton & Rose) Luetzelb.**; S38V6B; NS, NS, KR998106, NS, NS, NS, NS, KR998094, NS, NS, KR998123, NS, NS, KT984784, KT984768.