

## RESEARCH ARTICLE

# Characterization of sequence variability hotspots in Cranichideae plastomes (Orchidaceae, Orchidoideae)

Eric de Camargo Smidt<sup>1†\*</sup>, Michelle Zavala Páez<sup>1</sup>, Leila do Nascimento Vieira<sup>1</sup>, Juan Viruel<sup>2</sup>, Valter Antônio de Baura<sup>3</sup>, Eduardo Balsanelli<sup>3</sup>, Emanuel Maltempi de Souza<sup>3</sup>, Mark W. Chase<sup>2,4</sup>

**1** Departamento de Botânica, Universidade Federal do Paraná, Curitiba, Paraná, Brazil, **2** Royal Botanic Gardens, Kew, Richmond, Surrey, England, United Kingdom, **3** Departamento de Bioquímica, Universidade Federal do Paraná, Núcleo de Fixação Biológica de Nitrogênio, Curitiba, Paraná, Brazil, **4** Department of Environment and Agriculture, Curtin University, Perth, Western Australia, Australia

† Current address: Royal Botanic Gardens, Kew, Richmond, Surrey, England, United Kingdom  
\* [ecsmidt@yahoo.com.br](mailto:ecsmidt@yahoo.com.br)



## OPEN ACCESS

**Citation:** Smidt EdC, Páez MZ, Vieira LdN, Viruel J, de Baura VA, Balsanelli E, et al. (2020) Characterization of sequence variability hotspots in Cranichideae plastomes (Orchidaceae, Orchidoideae). PLoS ONE 15(1): e0227991. <https://doi.org/10.1371/journal.pone.0227991>

**Editor:** Zhong-Jian Liu, The National Orchid Conservation Center of China; The Orchid Conservation & Research Center of Shenzhen, CHINA

**Received:** November 20, 2019

**Accepted:** January 3, 2020

**Published:** January 28, 2020

**Copyright:** © 2020 Smidt et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All sequence files are available from the NCBI-Genbank database (accession numbers MN597437, MN597436, MN597435, MN597434, MN597433, MN597432).

**Funding:** ECS would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for grant Bolsa de Produtividade em Pesquisa CNPq-Nível 2 (proc. 311001/2014-9) and Pos-doc grant (proc. 203304/

## Abstract

This study reports complete plastome sequences for six species of Neotropical Cranichideae and focuses on identification of the most variable regions (hotspots) in this group of orchids. These structure of these six plastomes is relatively conserved, exhibiting lengths ranging between 142,599 to 154,562 bp with 36.7% GC on average and exhibiting typical quadripartite arrangement (LSC, SSC and two IRs). Variation detected in the LSC/IR and SSC/IR junctions is explained by the loss of *ndhF* and *ycf1* length variation. For the two genera of epiphytic clade in Spiranthinae, almost whole sets of the *ndh*-gene family were missing. Eight mutation hotspots were identified based on nucleotide diversity, sequence variability and parsimony-informative sites. Three of them (*rps16-trnQ*, *trnT-trnL*, *rpl32-trnL*) seem to be universal hotspots in the family, and the other five (*trnG-trnR*, *trnR-atpA*, *trnP-psaJ*, *rpl32-infA*, and *rps15-ycf1*) are described for the first time as orchid molecular hotspots. These regions have much more variation than all those used previously in phylogenetics of the group and offer useful plastid markers for phylogenetic, barcoding and population genetic studies. The use of whole plastomes or exclusive no-gap matrices also positioned with high support the holomycotrophic *Rhizanthella* among Orchidoideae plastomes in model-based analyses, showing the utility of plastomes for phylogenetic placement of this unusual genus.

## Introduction

Orchidaceae are one of the largest families of flowering plants, including five subfamilies, 736 genera and more than 30,000 species [1, 2]. One of the five subfamilies is Orchidoideae, divided into four tribes, including Cranichideae. As currently recognized, this cosmopolitan subtribe has eight subtribes and includes 96 genera with ca. 1,400 spp [1, 3]. They are characterized by their fleshy, but not tuberose, fasciculate roots, clustered or distributed along rhizome [3].

2018-7) and Brazilian Program of National Institutes of Science and Technology-INCT/Brazilian Research Council-CNPq/MCT.

**Competing interests:** The authors have declared that no competing interests exist.

Many studies have dealt with phylogenetic issues in the tribe, some in a more comprehensive way focusing on revised limits of the subtribes or involving circumscription or new positions of genera [4–18], among others. In these studies, largely based on plastid markers, the nuclear ribosomal internal transcribed spacers (nrITS) have been the exclusive source of nuclear information in the tribe (exception being [18], who used *Xdh* in Goodyerinae). Most of these studies used only a few plastid regions, usually *rbcl* and *matK* exons, the intron/spacer of *trnL-F* and less frequently the *atpB* and *psaB* exons and the *rpoB-trnC* spacer. Use of these regions was purely based on their previous use in other studies without prior analysis of their variability. Moreover, the low resolution and support in several clades in these studies necessitates identification of new and more variable regions for obtaining well-supported results.

In recent years, high-throughput sequencing technologies have provided an opportunity for study of genomic evolution and interspecific relationships of organisms based on large-scale genomic resources, such as complete plastome sequences [19, 20]. Due to their abundance in the cells, generally maternal inheritance and conserved structural characteristics, plastomes can provide information for studying species divergence and interspecific relationships of plants [21]. In most angiosperms, plastomes have a quadripartite structure: two inverted regions (IRs) intercalated between large single-copy (LSC) and small single-copy (SSC) regions [21, 22]. The plastome is circular and usually ranges from 107 to 218 kbp in length with about 120 genes [20]. In the orchid family, the length of the plastome is usually within this range [23, 24], except for holomycotrophic species in which some plastomes have shrunk to 19 kbp [25, 26].

In this study, we aim to identify the most variable plastid regions in this orchid tribe for use in future phylogenetics and population genetics studies of Atlantic Rainforest endemics. To do that, complete plastomes of six Cranichidinae species native to the Neotropics were first assembled and annotated. We analysed the differences in genome size, content, and structure, characterizing sequence divergence in hotspot regions. We also included six previously published Paleotropical Cranichidinae plastomes. In addition, we evaluated the phylogenetic relationships in Orchidoideae based on eighteen plastome datasets to contribute to an understanding the position of these genera with the phylogenomic data available so far. We focus in particular on the Australian underground orchid, holomycotrophic *Rhizanthella gardneri* R.S. Rogers, which has been the subject of phylogenetic controversy due to having one of the most highly reduced plastomes of any orchid sequenced to date, with only 59,910 bp and 33 retained genes [3, 27, 28].

## Material and methods

### Taxon sampling, DNA extraction and sequencing

Fresh leaf material of *Aspidogyne longicornu* (Cogn.) Garay (Goodyerinae) *Prescottia stachyodes* (Sw.) Lindl. (Cranichidinae), *Cyclopogon argyriifolius* (Barb.Rodr.) Barb.Rodr., *Sauroglossum elatum* Lindl., *Eurystyles cotyledon* Wawra, *Lankesterella ceracifolia* (Barb.Rodr.) Mansf. (all Spiranthinae) was collected from southeastern Brazil in the Atlantic Rainforest (Table 1).

Plastid enrichment was performed according to the protocol developed by [29], adapted for a small amount of tissue [30]. Plastid-enriched DNA was extracted according to [31], scaled to 2 mL, and purified with DNA Clean and Concentrator kit (Zymo Research, Orange, CA). Approximately 1 ng of DNA was used for library preparation using the Nextera XT DNA Sample Prep Kit (Illumina Inc., San Diego, CA), following manufacturer's instructions. Sequencing was performed on an Illumina MiSeq (Illumina Inc., San Diego, CA). The number of paired reads ( $2 \times 250$  bp) per sample varied from 501,686 to 1,637,814 and were imported as FASTQ

**Table 1. Vouchers and general features of the complete plastomes of Cranichideae species analyzed in this study.**

Plastome	Voucher	Total cpDNA Size (bp)	Length of LSC region (bp)	Length of IR regions (bp)	Length of SSC region (bp)	% GC	GenBank accession
<i>Aspidogyne longicornu</i>	Engels, M.E. 1504 (UPCB)	153,436	82,437	26,492	18,015	36.9	MN597437
<i>Cyclopogon argyriifolius</i>	Carvalho, B.M. 153 (RB)	154,495	83,704	26,506	17,779	36.9	MN597436
<i>Eurystyles cotyledon</i>	Imig, D.C. 366 (UPCB)	139,811	79,842	24,396	11,177	36.5	MN597435
<i>Lankesterella ceracifolia</i>	Imig, D.C. 358 (UPCB)	142,599	78,992	26,892	9,823	36.3	MN597434
<i>Prescottia stachyoides</i>	Smidt, E.C. 1080 (UPCB)	154,044	83,373	26,493	17,685	36.7	MN597433
<i>Sauroglossum elatum</i>	Imig, D.C. 502 (UPCB)	154,562	83,822	26,503	17,734	37.0	MN597432

<https://doi.org/10.1371/journal.pone.0227991.t001>

format in CLC Genomics Workbench v.11.0 (<http://www.qiagenbioinformatics.com>). Reads were trimmed using a 0.05 quality threshold, resulting in 498,811 to 1,534,494 reads of 167 bp length on average (S1 Table). These reads were assembled into contigs in CLC Genomics Workbench (v. 8 and 12) and DNASTAR Lasergene v. 14. The correct size and positioning of the IRs were determined using REPuter [32], and the IR boundaries were analyzed and visually compared in Geneious Prime 2019.1.1 (<https://www.geneious.com>). Annotations were manually checked: protein-coding regions were validated with ExpASy [33] and tRNA with tRNAscan-SE 2.0 [34]. The graphic representation of the plastomes was then produced using OGDRAW [35].

Plastome structure, gene content, and general characteristics of the plastid genomes sequenced were compared with three published plastomes of the same subtribe from the Paleotropics available on the NCBI website: *Anoectochilus emeiensis* K.Y.Lang (NC033895), *Goodyera fumata* Thwaites (NC026773), and *Goodyera schlechtendaliana* Rchb.f. (NC029364). Sequences were aligned using the progressive Mauve algorithm [36] in Geneious Prime to verify structural differences between plastomes.

Comparison of the number of protein, tRNA and rRNA coding regions was performed by plastome alignment with MAFFT 7 software with the automatic algorithm plugin, 200PAM/ $k = 2$  [37] in Geneious Prime. In order to examine IR junctions, sequence and content of these regions were plotted using the online IRscope tool [38]. DnaSP v5.10.1. [39] was employed to assess sequence variability (SV) among each protein-coding gene (CDS), intron, and intergenic spacer (IGS) of more than 150 bp flanked by the same region. The number of mutations, insertion/deletion (indel) events, and conserved sites were counted for each sequence. SV was calculated according to the method of [40]:  $SV = (\text{number of nucleotide mutations} + \text{the number of indel events}) / (\text{the number of conserved sites} + \text{the number of nucleotide mutations} + \text{the number of indel events}) \times 100\%$ . The number of potentially parsimony informative sites (PIS) and nucleotide diversity (Pi) were obtained as calculated in DnaSP. The nrITS sequences of the same species were obtained from GenBank to compare the levels of variation in various plastid regions with this most widely used nuclear marker in Orchidaceae (except for the use of *Anoectochilus roxburghii* (Wall.) Lindl.—GQ32877 instead of *A. emeiensis* due the lack of an available nrITS).

## Phylogenetic analyses

For phylogenomic analyses, we used as an outgroup the plastome of *Cypripedium japonicum* Thunb. (NC027227), subfamily Cypripedioideae. We also included the available plastomes

from subfamily Orchidoideae, tribe Orchideae, subtribe Orchidinae: *Dactylorhiza majalis* (Rchb.) P.F.Hunt & Summerh. (MK984209), *Habenaria pantlingiana* Kraenzl. (NC026775), *Ophrys sphegodes* Mill. (AP018717), *Platanthera chlorantha* (Custer) Rchb. (MK937914), *P. japonica* (Thunb.) Lindl. (NC037440); tribe Diurideae, subtribe Rhizanthellinae: *Rhizanthella gardneri* R.S.Rogers (GQ413967); and tribe Cranichideae, Subtribe Goodyerinae: *Anoectochilus emeiensis* K.Y.Lang (NC033895), *Goodyera fumata* Thwaites (NC026773), *Goodyera procera* (Ker Gawl.) Hook. (NC029363), *Goodyera schlechtendaliana* Rchb. f. (NC029364), *Goodyera velutina* Maximowicz ex Regel (NC029365), and *Ludisia discolor* (Ker Gawl.) Blume (NC030540). Multiple sequence alignments were performed using MAFFT 7 [37] with settings described above, visually inspecting and manually adjusting them with Geneious Prime. Indels were treated as missing data. Phylogenetic analyses were performed using maximum likelihood (ML), maximum parsimony (MP), and Bayesian inference (BI) in order to explore the results with different assumptions. Due to the loss of many plastid genes in *R. gardneri*, we repeated the phylogenetic analyses with only non-gap positions to reduce the effect of missing data on the phylogenetic placement of this species. ML trees were calculated using IQ-tree 1.6.11 [41] with the concatenated unpartitioned scheme and tree search with 1,000 bootstrap replicates in a single run [42, 43]. The best-fit models of substitution inferred in the “super-gene” analyses by AIC were GTR+F+R5. For the support, we used the ultrafast bootstrap approximation (UFBoot) [44], with the strategy (-bnni) to reduce the risk of overestimating support. Also, we calculated the SH-like approximate likelihood ratio test (-alrt 1000) [45] as an alternative method of estimating support. We accepted only SH-aLRT  $\geq 80\%$  and UFboot  $\geq 95\%$  as a good support. MP analyses were performed with Fitch parsimony [46] using the software PAUP 4.0b10a [47]. Analyses included 1,000 random taxon-addition replicates, holding 10 trees per replicate, TBR swapping algorithm, followed by a second search to explore all topologies from the previous search, limited to 10,000 trees. Support was estimated with 1,000 bootstrap replicates [48], simple addition, TBR algorithm and holding 10 trees per replicate. BI was performed using MrBayes 3.2.6 [49] with GTR+I+G model. The analyses started from random trees and employed Markov chain Monte Carlo (MCMC) runs, over ten million generations, sampling trees every 1000 generations. We discarded 25% of the initial generations as the burn-in, after visual inspection for stabilization of tree log-likelihood, as measured by the standard deviation(s) and PSRF (potential scale reduction factor) values [50] and ESS  $> 300$  in the Tracer 1.7 [51]. The remaining trees were used to produce the BI tree with posterior probabilities. All trees were edited with FigTree 1.4.0 [52].

## Results

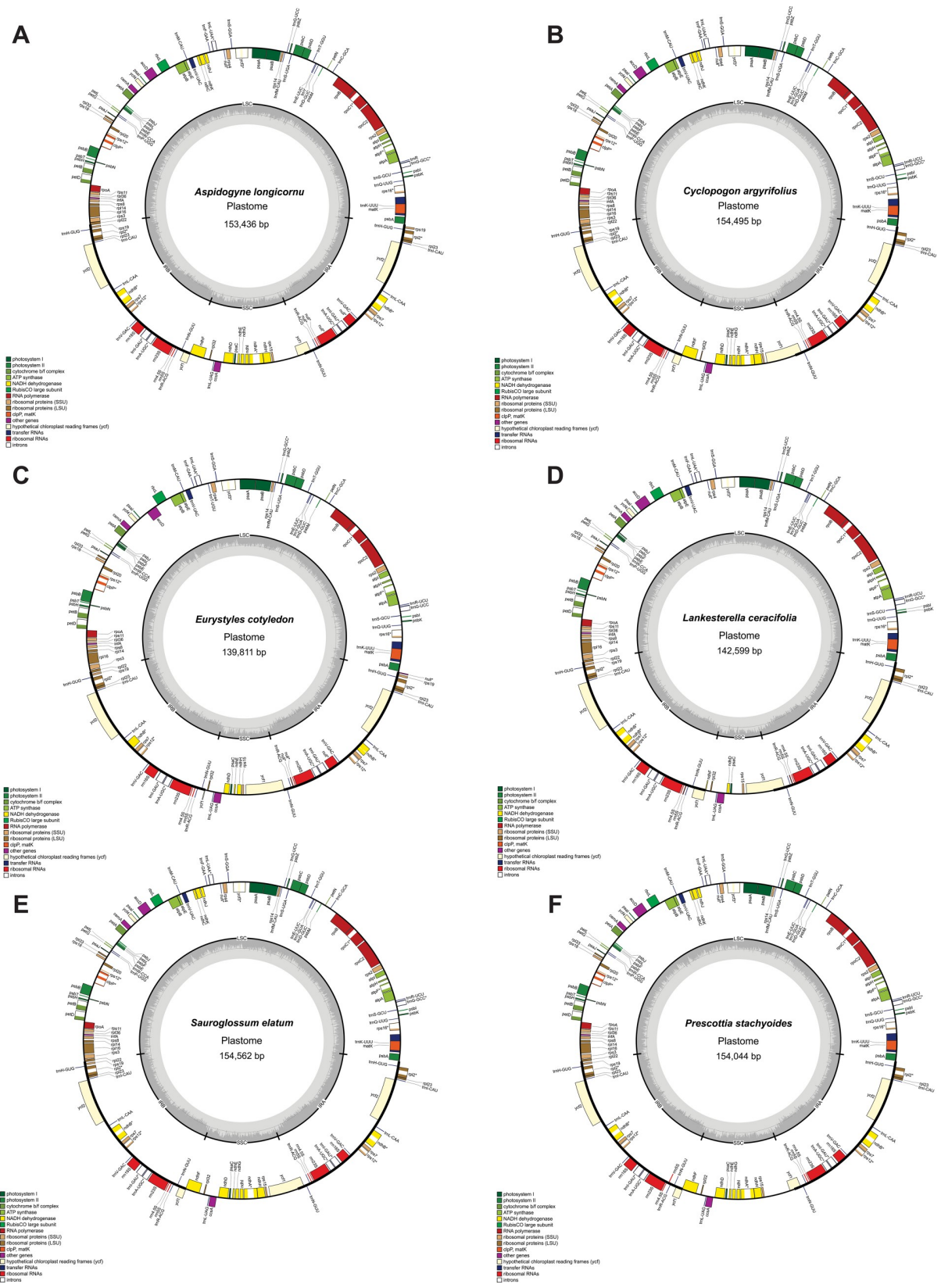
### Plastome description and hotspot characterization

The six newly sequenced plastomes exhibited lengths ranging from 139,811 to 154,562 bp, with 36.7% average GC content and the typical quadripartite structure (Table 1, Fig 1).

The Mauve alignment indicated that there are no major structural differences between the analyzed plastomes (S1 Fig). Variation in the length of some regions (S2 Table) and differences in the positioning of some genes in IR boundaries (Fig 2) was identified.

The list of plastid genes identified is presented in Table 2, and statistics for each CDS, intron, and intergenic spacer are in S2 Table.

A few length differences in one copy of the *rpl22* gene were detected in the LSC/IRb junction. The *ndhF* region was absent in *Eurystyles cotyledon*, and the length of *ycf1* was shorter than in the other studied species in the IRb/SSC junction. In the SSC junction with IRa, there are differences in the length of the second copy of the *ycf1* and a few differences in the length of *rpl22*.



**Fig 1. Cranichideae plastomes circular map.** A, *Aspidogyne longicornu*, B, *Cyclopogon argyriifolius*, C, *Eurystyles cotyledon*, D, *Lankesterella ceracifolia*, E, *Prescottia stachyoides*, F, *Sauroglossum elatum*. The genes represented inside the circle are transcribed clockwise and those outside are transcribed counter-clockwise. The dark gray rim inside the figure corresponds to the distribution of GC content across the plastome.

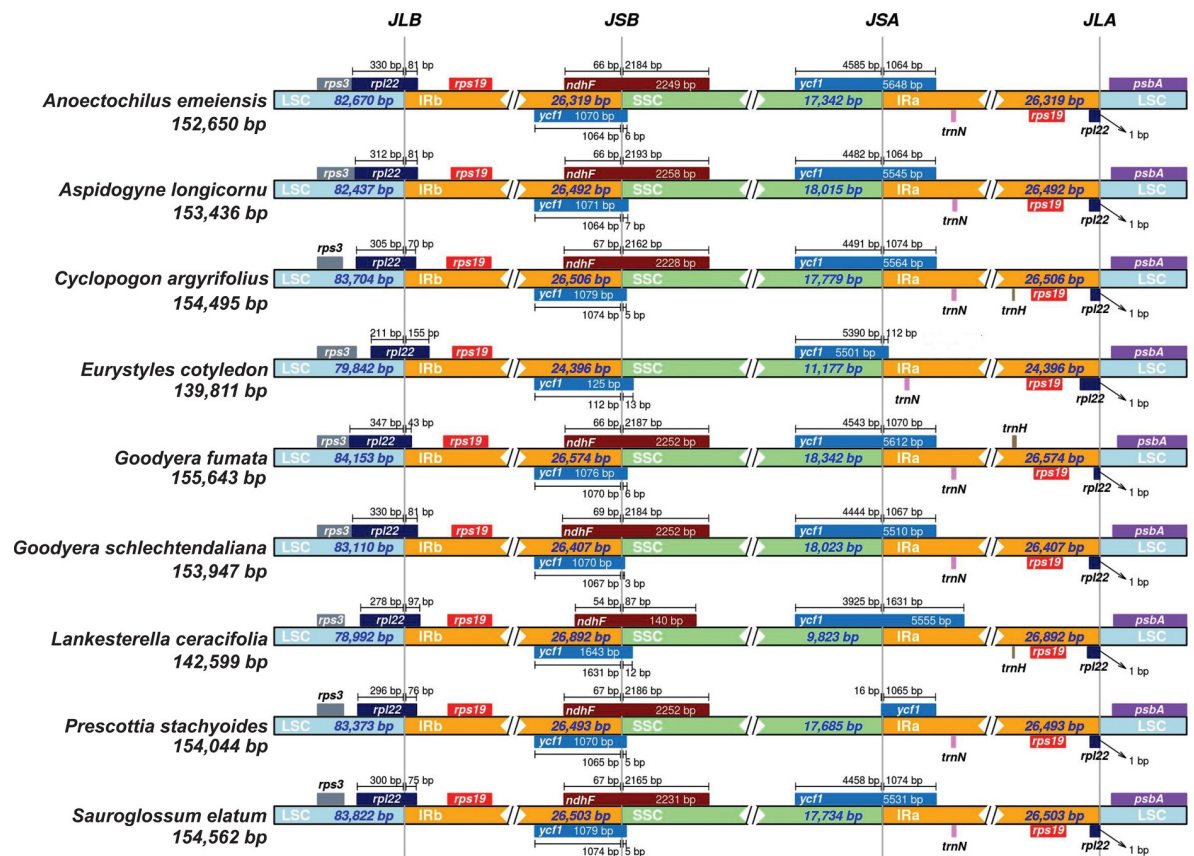
<https://doi.org/10.1371/journal.pone.0227991.g001>

All genes are present in all species, except for some subunits of NAD(P)H-quinone oxidoreductase, which are missing in *Lankesterella ceracifolia* (*ndhA*, *ndhC*, *ndhE*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*) and *Eurystyles cotyledon* (*ndhA*, *ndhC*, *ndhF*, *ndhG*, *ndhI*, *ndhJ*, *ndhK*).

For all regions longer than 150 bp, the statistics for Pi, SV and PIS/bp are relatively congruent. For the ten most variable regions in each statistic, all are intergenic spacers and eight most variable regions are the same for the three parameters (*rps16-trnQ*, *trnG-trnR*, *trnR-atpA*, *trnT-trnL*, *trnP-psaJ*, *rpl32-infA*, *rps15-ycf1* and *rpl32-trnL*). No individual region has more nucleotide diversity (Pi) than nrITS, but *rpl32-trnL* and *rps16-trnQ* contain more sequence variability (SV) and *rpL16-rpS6* contains more parsimony-informative sites (Fig 3, S2 Table).

### Phylogenetic relationships

The aligned dataset consisted of 195,841 characters, of which 153,336 (78%) are monomorphic, and 16,086 (8.2%) characters were potentially parsimony-informative. MP recovers only one tree (L = 64,472 steps) with a consistency index of 0.78 and retention index of 0.72. All phylogenetic analyses reconstructed the same topology with strong support for monophyly of



**Fig 2. Comparison of the boundaries of LSC, SSC, and IR regions of the complete plastomes of Cranichideae.** JLB (IRb/LSC), JSB (IRb/SSC), JSA (SSC/IRa) and JLA (IRa/LSC) denote the junctions between each corresponding region in the genome.

<https://doi.org/10.1371/journal.pone.0227991.g002>

**Table 2. List of genes identified in the plastomes of *Aspidogyne longicornu*, *Cyclopogon argyriifolius*, *Eurystyles cotyledon*, *Lankesterella ceracifolia*, *Prescottia stachyoides* and *Sauroglossum elatum*.**

Group of gene	Gene name
<b>Gene expression machinery</b>	
Ribosomal RNA genes	<i>rrn4.5<sup>b</sup></i> ; <i>rrn5<sup>b</sup></i> ; <i>rrn16<sup>b</sup></i> ; <i>rrn23<sup>b</sup></i>
Transfer RNA genes	<i>trnA-UGC<sup>ab</sup></i> ; <i>trnC-GCA</i> ; <i>trnD-GUC</i> ; <i>trnE-UUC</i> ; <i>trnF-GAA</i> ; <i>trnG-M-CAU</i> ; <i>trnG-GCC<sup>b</sup></i> ; <i>trnG-UCC</i> ; <i>trnH-GUG<sup>b</sup></i> ; <i>trnI-CAU<sup>b</sup></i> ; <i>trnI-GAU<sup>ab</sup></i> ; <i>trnK-UUU</i> ; <i>trnL-CAA<sup>b</sup></i> ; <i>trnL-UAA<sup>a</sup></i> ; <i>trnL-UAG</i> ; <i>trnM-CAU</i> ; <i>trnN-GUU<sup>b</sup></i> ; <i>trnP-UGG</i> ; <i>trnQ-UUG</i> ; <i>trnR-ACG<sup>b</sup></i> ; <i>trnR-UCU</i> ; <i>trnS-GCU</i> ; <i>trnS-GGA</i> ; <i>trnS-UGA</i> ; <i>trnT-GGU</i> ; <i>trnT-UGU</i> ; <i>trnV-GAC<sup>b</sup></i> ; <i>trnV-UAC<sup>a</sup></i> ; <i>trnW-CCA</i> ; <i>trnY-GUA</i>
Small subunit of ribosome	<i>rps2</i> ; <i>rps3</i> ; <i>rps4</i> ; <i>rps7<sup>b</sup></i> ; <i>rps8</i> ; <i>rps11</i> ; <i>rps12<sup>ab</sup></i> ; <i>rps14</i> ; <i>rps15</i> ; <i>rps16<sup>a</sup></i> ; <i>rps18</i> ; <i>rps19<sup>b</sup></i>
Large subunit of ribosome	<i>rpl2<sup>ab</sup></i> ; <i>rpl14</i> ; <i>rpl16<sup>a</sup></i> ; <i>rpl20</i> ; <i>rpl22<sup>b</sup></i> ; <i>rpl23<sup>b</sup></i> ; <i>rpl32</i> ; <i>rpl33</i> ; <i>rpl36</i>
DNA-dependent RNA polymerase	<i>rpoA</i> ; <i>rpo<sup>b</sup></i> ; <i>rpoC1<sup>a</sup></i> ; <i>rpoC2</i>
Translational initiation factor	<i>infA</i>
Maturase	<i>matK</i>
<b>Genes for photosynthesis</b>	
Subunits of photosystem I (PSI)	<i>psaA</i> ; <i>psaB</i> ; <i>psaC</i> ; <i>psaI</i> ; <i>psaJ</i> ; <i>ycf3<sup>a</sup></i> ; <i>ycf4</i>
Subunits of photosystem II (PSII)	<i>psbA</i> ; <i>psbB</i> ; <i>psbC</i> ; <i>psbD</i> ; <i>psbE</i> ; <i>psbF</i> ; <i>psbH</i> ; <i>psbI</i> ; <i>psbJ</i> ; <i>psbK</i> ; <i>psbL</i> ; <i>psbM</i> ; <i>psbN</i> ; <i>psbT</i> ; <i>psbZ</i>
Subunits of cytochrome <i>b<sub>6</sub>f</i>	<i>petA</i> ; <i>petB<sup>a</sup></i> ; <i>petD<sup>a</sup></i> ; <i>petG</i> ; <i>petL</i> ; <i>petN</i>
Subunits of ATP synthase	<i>atpA</i> ; <i>atpB</i> ; <i>atpE</i> ; <i>atpF<sup>a</sup></i> ; <i>atpH</i> ; <i>atpI</i>
Subunits of NADH dehydrogenase	<i>ndhA<sup>acd</sup></i> ; <i>ndhB<sup>abc</sup></i> ; <i>ndhC<sup>cd</sup></i> ; <i>ndhD<sup>cd</sup></i> ; <i>ndhE<sup>d</sup></i> ; <i>ndhF<sup>cd</sup></i> ; <i>ndhG<sup>cd</sup></i> ; <i>ndhH<sup>cd</sup></i> ; <i>ndhI<sup>cd</sup></i> ; <i>ndhJ<sup>cd</sup></i> ; <i>ndhK<sup>cd</sup></i>
Large subunit of Rubisco	<i>rbcL</i>
<b>Other functions</b>	
Envelope membrane protein	<i>cemA</i>
Subunit of acetyl-CoA carboxylase	<i>accD</i>
C-type cytochrome synthesis	<i>ccsA</i>
Subunit of protease Clp	<i>clpP<sup>a</sup></i>
Component of TIC complex	<i>ycf1<sup>b</sup></i>
Unknown function	<i>ycf2<sup>b</sup></i>

<sup>a</sup> Genes containing introns

<sup>b</sup> Duplicated genes

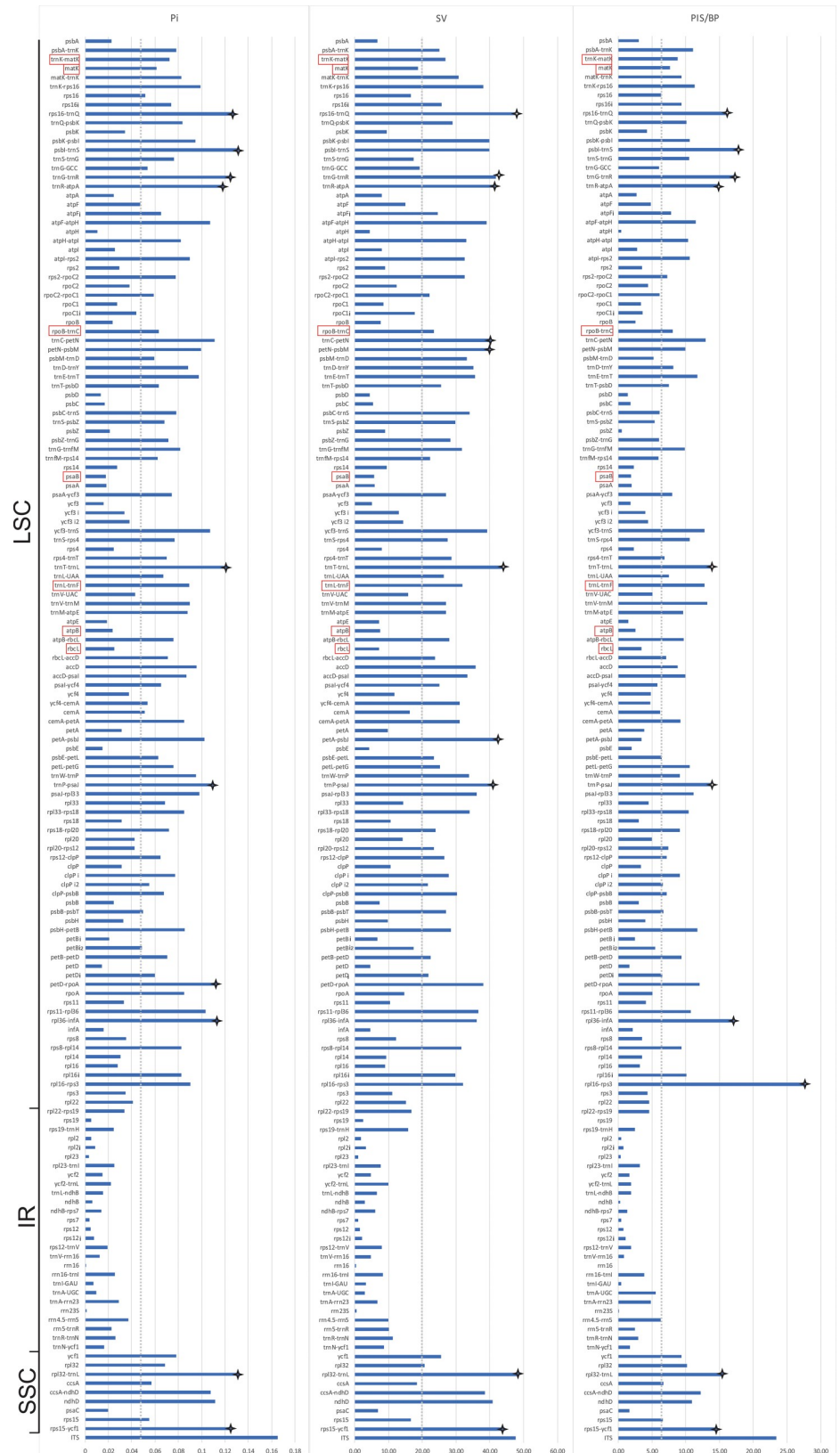
<sup>c</sup> Deleted in *Eurystyles*

<sup>d</sup> Deleted in *Lankesterella*

<https://doi.org/10.1371/journal.pone.0227991.t002>

tribes Cranichideae and Orchideae (Fig 4). In Cranichideae, Cranichidinae are monophyletic with strong support, *Prescottia stachyoides* is sister of Spiranthinae and both sister to Goodyerinae. In Spiranthinae, *Eurystyles cotyledon* plus *Lankesterella ceracifolia* are sister to *Cyclopogon argyriifolius* plus *Sauroglossum elatum*. In subtribe Goodyerinae, *Goodyera* is biphyletic, and *Aspidogyne* is sister to *G. fumata*. All these relationships receive strong support. *Rhizanthella* is sister to Orchideae with high support in the BI and ML trees, but with MP this species is sister to both Cranichideae and Orchideae with low support (not shown).

The matrix excluding gaps consisted of 51,193 characters, of which 36,614 (73%) are constant, and 4,090 (8.1%) characters were potentially parsimony-informative. MP recovered one

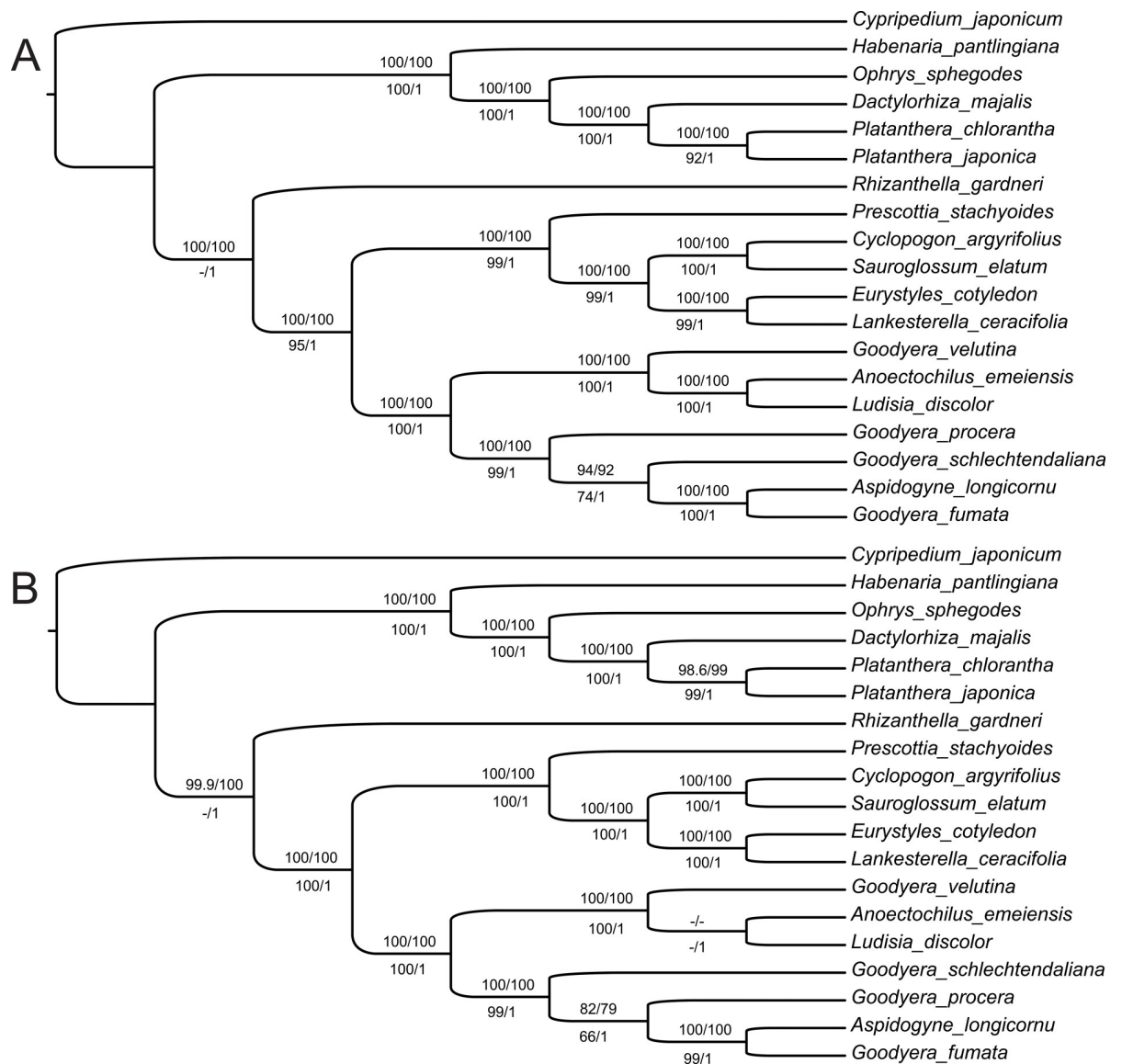




**Fig 3. Expected number of genetic differences within each protein-coding sequence (CDS), intron, and intergenic spacer (IGS) of Cranichideae plastomes.** Gene order is preserved beginning at the top with the large single-copy (LSC), intersection of the inverted repeat (IR), and the small single-copy (SSC) region at the bottom, ending with *rps15-ycf1*. For comparison, nrITS is placed at the bottom. The vertical dotted lines highlight the average value across regions (Pi = 0.056, SV = 20.14 and PIS/bp = 6.55), and the stars highlight the ten regions that on average contain the greatest amount of nucleotide diversity (Pi), highest sequence variability (SV), and greatest number of potentially informative characters (PIS/bp). Regions in red were used in previous phylogenetic studies of the group.

<https://doi.org/10.1371/journal.pone.0227991.g003>

tree (L = 19,860 steps) with a consistency index of 0.79 and retention index of 0.69. All analyses reconstructed the same topology for complete matrix, except for lack of support for a clade of Goodyerinae that contains *Ludisia discolor*, *Anoectochilus emeiensis* and *Goodyera* (Fig 4).



**Fig 4. Phylogenetic relationships inferred with ML, MP, and BI.** A. Full-length plastome sequences of representative Orchidoideae. B. With only non-gap sites of the plastomes. Node support is derived from IQ-tree bootstrap, and SH-like approximate likelihood ratio test above the branches and maximum parsimony bootstraps and Bayesian posterior probabilities below the branches. *Cypripedium japonicum* was specified as the outgroup.

<https://doi.org/10.1371/journal.pone.0227991.g004>

*Rhizanthella* is sister to Orchideae with high support in BI and ML analyses, and with MP this species is again sister to Cranichideae and Orchideae with low support (not shown).

## Discussion

### Plastome description and hotspot characterization

This study sequenced for the first time six whole plastomes for six orchid species in subtribe Cranichideae. The genomic structure and gene content are conserved in general. Plastome length in *Eurystyles cotyledon* and *Lankesterella ceracifolia* were shorter than those of the other species, due to loss of most *ndh* genes encoding subunits of the NAD(P)H-quinone oxidoreductase proteins. The *ndh* gene complex comprises 11 genes that act in electron transport in photosystem I [53]. However, in Orchidaceae plastomes, there is a great variation in retention/deletion of the *ndh* genes among different clades, suggesting independent losses [24, 54–57]. Because they are important genes for photosynthesis, there may be functional copies in the nuclear or mitochondrial genomes [21, 26, 55, 58, 59], but it has also been suggested that the *ndh* genes are dispensable [60]. In several photosynthetic groups that are neither mycoheterotrophic nor parasitic (summarized in [61]), *ndh* genes have been lost and often are the first plastid genes to be lost [26, 62]. *Eurystyles* and *Lankesterella* are the only genera of the epiphytic clade in Spiranthinae (*sensu* [17]), and their *ndh* gene loss should probably be considered an apomorphic character in the relation to the other two Spiranthinae plastomes sequenced. [54, 55] suggested that there is a clear relationship between the loss of *ndh* genes and instability in IR/SSC boundaries, and [24] also reported length variation in *ycf1*. In our study, both conditions are related to differences in these boundaries (Fig 2).

Niu et al. [63] reported mutational hotspots as genus-specific in orchid plastomes and proposed three markers for Epidendroideae (*trnK-rps16*, *trnS-trnG*, and *rps16-trnQ*) and two for Cypripedioideae (*clpP-psbB* and *rps16-trnQ*). In a study of plastome variation in *Dendrobium*, [64] proposed the combination of five markers (*trnT-trnL*, *rpl32-trnL*, *clpP-psbB*, *trnL intron*, and *rps16-trnQ*) for phylogenetic and DNA barcoding studies of the genus. Here, we suggest eight hotspot markers for Cranichideae. Due to previously being regarded in other studies as hot spots, three of them (*rps16-trnQ*, *trnT-trnL*, *rpl32-trnL*) seem to be universal in the family, but the other five (*trnG-trnR*, *trnR-atpA*, *trnP-psa*, *rpl32-infA*, and *rps15-ycf1*) are identified for the first time.

### Phylogenetic relationships

In all analyses, the currently recognized taxa (tribes and subtribes) with more than one plastome included in the analysis were recovered with high support and, although with limited taxon sampling, our results agreed with the classification proposed in [1]. In Cranichidinae, as expected, *Prescottia* is sister to the Spiranthinae plastomes sampled, and in Spiranthinae, the epiphytic clade (*Eurystyles* plus *Lankesterella*) is also recovered with high support as sister to *Cyclopogon* plus *Sauroglossum*, in agreement with the results of [17]. By contrast, for the Goodyerinae, although the subtribe is recovered with high support, *Goodyera* is biphyletic. Chen et al. (2019) demonstrated the non-monophyly of *Goodyera* as currently circumscribed using two nuclear and five plastid markers. Our results place *Goodyera fumata* sister to *Aspidogyne*, which corresponds to the *Microchilus* subclade *sensu* [18].

Many non-photosynthetic taxa (both mycotrophic and parasitic) have a history of being phylogenetically difficult to place because many of the plastid regions utilized in early phylogenies are absent or degraded, and there have been such problems as well with holomycotrophic orchids [25, 26, 65, 66]. The position of holomycotrophic *Rhizanthella* in our analyses agrees with the classification scheme of [1], sister to Cranichideae with high ML and BI support and

low support with MP, which placed *Rhizanthella* as a sister to both tribes. The full plastome analysis shows better support than the use only non-gapped sites. In addition, absence of some regions in the plastomes of non-photosynthetic taxa species and rate heterogeneity could lead to phylogenetic artifacts [65].

## Conclusions

In this study, six plastid genomes of Cranichideae species were sequenced and annotated. They are relatively conserved, but for the epiphytic clade of Spiranthinae almost the whole set of the *ndh*-gene complex was truncated or pseudogenized. Eight mutation hotspot regions were identified in these plastid genomes and have much more variation than any those used for the group before, giving them potential as markers for phylogenetic, barcoding and population genetic studies. Although different criteria and methods could affect hotspot identification, our findings in this study compared with previous studies in other orchid groups suggest many plastid hotspots in are clade-specific, whereas other hotspots are shared by all orchid genera. The use of complete plastomes supported placement of holomycotrophic *Rhizanthella* as sister to Cranichideae, demonstrating the utility of plastomes in orchid phylogenetics.

## Supporting information

**S1 Fig. Mauve alignment with the progressive Mauve algorithm of the plastomes.** Monochromatic alignment indicates that there is no structural variation among the plastomes. (TIFF)

**S1 Table. General features of reads in sequencing of plastomes for Cranichideae.** (DOCX)

**S2 Table. Statistics of the expected number of genetic differences within each protein-coding gene (CDS), introns, and intergenic spacer (IGS) of Cranichideae plastomes.** For comparison, nrITS is placed at bottom. The columns indicate: A, position in the plastome, B, Name of the region, C, Region of the plastome (SSC, IR, LSC), D, Type of the region (CDS, IGS, INTRON), E, Nucleotide diversity (Pi), F, Number of conserved sites (CS), G, Number of Mutations (NS), H, Indels Events (ID), I, number of inversions (IV), J, Sequence variability (SV), K, Number of parsimony informative sites (PIS), L, PIS/base pairs, M, Total sequence size (L), N, L less gaps sites. (XLSX)

## Acknowledgments

ECS would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Brazilian Program of National Institutes of Science and Technology-INCT/Brazilian Research Council-CNPq/MCT.

## Author Contributions

**Conceptualization:** Eric de Camargo Smidt, Mark W. Chase.

**Data curation:** Eric de Camargo Smidt.

**Formal analysis:** Eric de Camargo Smidt, Michelle Zavala Páez, Leila do Nascimento Vieira, Juan Viruel, Mark W. Chase.

**Methodology:** Michelle Zavala Páez, Leila do Nascimento Vieira, Juan Viruel, Valter Antônio de Baura, Eduardo Balsanelli, Emanuel Maltempi de Souza.

**Project administration:** Eric de Camargo Smidt.

**Writing – original draft:** Eric de Camargo Smidt, Michelle Zavala Páez, Leila do Nascimento Vieira, Juan Viruel, Valter Antônio de Baura, Eduardo Balsanelli, Emanuel Maltempi de Souza, Mark W. Chase.

**Writing – review & editing:** Eric de Camargo Smidt.

## References

1. Chase MW, Cameron KM, Freudenstein JV, Pridgeon AM, Salazar G, van den Berg C, et al. An updated classification of Orchidaceae. *Bot J Linn Soc.* 2015; 177:151–174. <https://doi.org/10.1111/boj.12234>
2. POWO. 2019. [cited 15 Nov 2019] Plants of the world online. Facilitated by the Royal Botanic Gardens, Kew. Available from: <http://www.plantsoftheworldonline.org/>
3. Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. *Genera orchidacearum*, vol. 3 Orchidoideae (part 2) Vanilloideae. London: Oxford University Press.; 2003.
4. Álvarez-Molina A, Cameron KM. Molecular phylogenetics of Prescottiinae s.l. and their close allies (Orchidaceae, Cranichideae) inferred from plastid and nuclear ribosomal DNA sequences. *Am J Bot.* 2009; 96:1020–1040. <https://doi.org/10.3732/ajb.0800219> PMID: 21628253
5. Batista JAN, Meneguzzo TEC, Salazar GA, Ramalho AJ, de Bem Bianchetti L. Phylogenetic placement, taxonomic revision and a new species of *Nothostele* (Orchidaceae), an enigmatic genus endemic to the cerrado of central Brazil. *Bot J Linn Soc.* 2011; 165:348–363. <https://doi.org/10.1111/j.1095-8339.2011.01113.x>
6. Borba EL, Salazar GA, Mazzoni-Viveiros S, Batista JAN. Phylogenetic position and floral morphology of the Brazilian endemic, monospecific genus *Cotylolabium*: a sister group for the remaining Spiranthinae (Orchidaceae). *Bot J Linn Soc.* 2014; 175:29–46. <https://doi.org/10.1111/boj.12136>
7. Cameron KM. Utility of plastid *psaB* gene sequences for investigating intrafamilial relationships within Orchidaceae. *Mol Phylogenet Evol.* 2004; 31:1157–1180. <https://doi.org/10.1016/j.ympev.2003.10.010> PMID: 15120407
8. Cameron KM, Chase MW, Whitten WM, Kores PJ, Jarrell DC, Albert VA, et al. A phylogenetic analysis of the Orchidaceae: evidence from *rbcL* nucleotide sequences. *Am J Bot.* 1999; 86:208–224. <https://doi.org/10.2307/2656938> PMID: 21680360
9. Cisternas Ma, Salazar Ga, Verdugo G, Novoa P, Calderón X, Negritto Ma. Phylogenetic analysis of Chloraeinae (Orchidaceae) based on plastid and nuclear DNA sequences. *Bot J Linn Soc.* 2012; 168:258–277. <https://doi.org/10.1111/j.1095-8339.2011.01200.x>
10. Clements MA, Tupac Otero CJ, Miller JT. Phylogenetic relationships in Pterostylidinae (Cranichideae: Orchidaceae): combined evidence from nuclear ribosomal and plastid DNA sequences. *Aust J Bot.* 2011; 59: 99–117. <https://doi.org/10.1071/BT10190>
11. Figueroa C, Salazar GA, Zavaleta A, Engleman M. Root character evolution and systematics in Crani-chidinae, Prescottiinae and Spiranthinae (Orchidaceae, Cranichideae). *Ann Bot.* 2008; 101: 509–520. <https://doi.org/10.1093/aob/mcm328> PMID: 18263628
12. Janes JK, Steane DA, Vaillancourt RE, Duretto MF. A molecular phylogeny of the subtribe Pterostylidinae (Orchidaceae): resolving the taxonomic confusion. *Aust. Syst. Bot.* 2010; 23, 248–259. <https://doi.org/10.1071/SB10006>
13. Hu C, Tian H, Li H, Hu A, Xing F, Bhattacharjee A, et al. Phylogenetic analysis of a 'jewel orchid' genus *Goodyera* (Orchidaceae) based on DNA sequence data from nuclear and plastid regions. *PLoS ONE* 2016; 11(2): e0150366. <https://doi.org/10.1371/journal.pone.0150366> PMID: 26927946
14. Salazar GA, Chase MW, Soto Arenas MA, Ingrouille M. Phylogenetics of Cranichideae with emphasis on Spiranthinae (Orchidaceae, Orchidoideae): evidence from plastid and nuclear DNA sequences. *Am. J. Bot.* 2003; 90: 777–795. <https://doi.org/10.3732/ajb.90.5.777> PMID: 21659175
15. Salazar GA, Cabrera LI, Madriñán S, Chase MW. Phylogenetic relationships of Crani-chidinae and Prescottiinae (Orchidaceae, Cranichideae) inferred from plastid and nuclear DNA sequences. *Ann. Bot.* 2009; 104: 403–416. <https://doi.org/10.1093/aob/mcn257> PMID: 19136493
16. Salazar GA, van den Berg C, Popovkin A. Phylogenetic relationships of *Discyphus scopulariae* (Orchidaceae, Cranichideae) inferred from plastid and nuclear DNA sequences: evidence supporting recognition of a new subtribe, Discyphinae. *Phytotaxa* 2014; 173: 127–139. <https://doi.org/10.11646/phytotaxa.173.2.3>

17. Salazar GA, Batista JAN, Cabrera LI, van den Berg C, Whitten WM, Smidt EC, et al. Phylogenetic systematics of subtribe Spiranthinae (Orchidaceae: Orchidoideae: Cranichideae) based on nuclear and plastid DNA sequences of a nearly complete generic sample. *Bot. J. Linn. Soc.* 2018; 186: 273–303. <https://doi.org/10.1093/botlinnean/box096>
18. Chen S, Tian H, Guan Q, Zhai J, Zhang G, Chen L, et al. Molecular systematics of Goodyerinae (Cranichideae, Orchidoideae, Orchidaceae) based on multiple nuclear and plastid regions. *Mol Phylogenet Evol.* 2019; 139: 106542. <https://doi.org/10.1016/j.ympev.2019.106542> PMID: 31229601
19. Olson PD, Hughes J, Cotton JA. Next generation systematics. Systematics Association Special Volume, vol. 85. Cambridge: Cambridge University Press; 2016.
20. Daniell H, Lin CS, Yu M, Chang WJ. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol Evol.* 2016; 17:134. <https://doi.org/10.1186/s13059-016-1004-2> PMID: 27339192
21. Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol. Biol.* 2011; 76: 273. <https://doi.org/10.1007/s11103-011-9762-4> PMID: 21424877
22. Zhu A, Guo W, Gupta S, Fan W, Mower JP. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytol.* 2016; 209: 1747–1756. <https://doi.org/10.1111/nph.13743> PMID: 26574731
23. Yang J, Tang M, Li H, Zhang Z, Li D. Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evol. Biol.* 2013; 13:84. <https://doi.org/10.1186/1471-2148-13-84> PMID: 23597078
24. Mauad AVSR, do Nascimento Vieira L, Bolson M, Baura V, Balsanelli E, Souza E, et al. Complete chloroplast genome of *Anathallis obovata* (Orchidaceae: Pleurothallidinae). *Braz. J. Bot.* 2019; 42: 345. <https://doi.org/10.1007/s40415-019-00524-3>
25. Schelkunov MI, Shtratnikova VY, Nuraliev MS, Selosse M, Penin AA, Logacheva MD. Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. *Genome Biol Evol.* 2015; 7:1179–1191. <https://doi.org/10.1093/gbe/evv019> PMID: 25635040
26. Graham SW, Lam VK, Merckx VS. Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes. *New Phytol.* 2017; 214: 48–55. <https://doi.org/10.1111/nph.14398> PMID: 28067952
27. Delannoy E, Fujii S, Colas des Francs-Small C, Brundrett M, Small I. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. *Mol Biol Evol.* 2011; 28:2077–86. <https://doi.org/10.1093/molbev/msr028> PMID: 21289370
28. Givnish TJ, Spalink D, Ames M, Lyon SP, Hunter SJ, Zuluaga A, et al. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc Royal Sci. Biol.* 2015; 282(1814). <https://doi.org/10.1098/rspb.2015.1553> PMID: 26311671
29. Vieira LN, Faoro H, Fraga HPF, Rogalski M, Souza EM, Pedrosa FO, et al. An improved protocol for intact chloroplasts and cpDNA isolation in conifers. *PLoS ONE.* 2014; 9: e84792. <https://doi.org/10.1371/journal.pone.0084792> PMID: 24392157
30. Sakaguchi S, Ueno S, Tsumura Y, Setoguchi H, Ito M, Hattori C, et al. Application of a simplified method of chloroplast enrichment to small amounts of tissue for chloroplast genome sequencing. *Appl. Plant Sci.* 2017; 5: 1700002. <https://doi.org/10.3732/apps.1700002> PMID: 28529832
31. Doyle J, Doyle J. A rapid DNA isolation procedure for small amounts of leaf tissue. *Phytochemical Bulletin* 1987; 19: 810–815.
32. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 2001; 29: 4633–4642. <https://doi.org/10.1093/nar/29.22.4633> PMID: 11713313
33. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003; 31(13):3784–3788. <https://doi.org/10.1093/nar/gkg563> PMID: 12824418
34. Lowe TM, Chan PP. tRNAscan-SE On-line: search and contextual analysis of transfer RNA genes. *Nucleic Acids Res.* 2016; 44: W54–57. <https://doi.org/10.1093/nar/gkw413> PMID: 27174935
35. Lohse M, Drechsel O, Kahlau S, Bock R. OrganellarGenome-DRAW: a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 2013; 41: W575–W581. <https://doi.org/10.1093/nar/gkt289> PMID: 23609545
36. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 2004; 14: 1394–1403. <https://doi.org/10.1101/gr.2289704> PMID: 15231754

37. Katoh K, Standley DM. 2013. MAFFT Multiple sequence alignment software version 7: improvements in performance and ssability. *Molecular Biology and Evolution* 30: 772–780. <https://doi.org/10.1093/molbev/mst010> PMID: 23329690
38. Amiryousefi A, Hyvönen JT, Poczai P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics*. 2018; 34:3030–3031. <https://doi.org/10.1093/bioinformatics/bty220> PMID: 29659705
39. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Mol. Biol. Evol.* 2017; 34: 3299–3302. <https://doi.org/10.1093/molbev/msx248> PMID: 29029172
40. Shaw J, Shafer H, Leonard R, Kovach M, Schorr M, Morris A. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. *Am J Bot.* 2014; 101:1987–2004. <https://doi.org/10.3732/ajb.1400398> PMID: 25366863
41. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2015; 32: 268–274. <https://doi.org/10.1093/molbev/msu300> PMID: 25371430
42. Chernomor O, von Haeseler A, Minh BQ. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst Biol.* 2016; 65: 997–1008. <https://doi.org/10.1093/sysbio/syw037> PMID: 27121966
43. Kalyaanamoorthy S, Minh B, Wong T, von Haeseler A, Jermini L. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017; 14. <https://doi.org/10.1038/nmeth.4285> PMID: 28481363
44. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Phylogenet Evol.* 2013; 30: 1188–1195. <https://doi.org/10.1093/molbev/mst024> PMID: 23418397
45. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 2010; 59: 307–321. <https://doi.org/10.1093/sysbio/syq010> PMID: 20525638
46. Fitch WM. Towards defining the course of evolution: minimum change for a specific tree topology. *Syst Zool.* 1971; 20: 406–416. <https://doi.org/10.2307/2412116>
47. Swofford DL. PAUP\*. Phylogenetic analysis using parsimony (\* and other methods), version 4.0b10. Sunderland: Sinauer; 2003.
48. Felsenstein J. Confidence limits on phylogenies: an approach to using bootstrap. *Evolution.* 1985; 39: 783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x> PMID: 28561359
49. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 2012; 61: 539–542. <https://doi.org/10.1093/sysbio/sys029> PMID: 22357727
50. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statist Sci.* 1992; 7: 457–472. <https://doi.org/10.1214/ss/1177011136>
51. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst Biol.* 2018; syy032. <https://doi.org/10.1093/sysbio/syy032> PMID: 29718447
52. Rambaut A. FigTree v1.4: tree figure drawing tool. [Internet]; 2009. Available from: <http://tree.bio.ed.ac.uk/software/figtree>
53. Martín M, Sabater B. Plastid *ndh* genes in plant evolution. *Plant Physiol Biochem.* 2010; 48: 636–645. <https://doi.org/10.1016/j.plaphy.2010.04.009> PMID: 20493721
54. Kim HT, Kim JS, Moore MJ, Neubig KM, Williams NH, Whitten WM, et al. Seven new complete plastome sequences reveal rampant independent loss of the *ndh* gene family across orchids and associated instability of the inverted repeat/small single copy region boundaries. *PLoS ONE* 2015; 10: e0142215. <https://doi.org/10.1371/journal.pone.0142215> PMID: 26558895
55. Lin C, Chen JJW, Huang Y, Chan M, Daniell H, Chang W, et al. The location and translocation of *ndh* genes of chloroplast origin in the Orchidaceae family. *Sci. Rep.* 2015; 5:9040. <https://doi.org/10.1038/srep09040> PMID: 25761566
56. Lin C, Chen JJW, Chiu C, Hsiao HCW, Yang C, Jin X, et al. Concomitant loss of *ndh* complex-related genes with chloroplast and nuclear genomes in some orchids. *Plant J.* 2017; 90: 994–1006. <https://doi.org/10.1111/tj.13525> PMID: 28258650
57. Kim HT, Chase MW. Independent degradation in genes of the plastid *ndh* gene family in species of the orchid genus *Cymbidium* (Orchidaceae; Epidendroideae). *PLoS ONE* 2015; 12: e0187318. <https://doi.org/10.1371/journal.pone.0187318> PMID: 29140976
58. Chang CC, Lin HC, Lin IP, Chow TY, Chen HH, Chen WH, et al. The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and its

- phylogenetic implications. *Mol Phylogenet Evol.* 2006; 23:279–291. <https://doi.org/10.1093/molbev/msj029>
59. Peltier G, Aro E-M, Shinakai T. NDH-1 and NDH-2 plastoquinone reductases in oxygenic photosynthesis. *Annu. Rev. Plant. Biol.* 2016; 67: 55–80. <https://doi.org/10.1146/annurev-arplant-043014-114752> PMID: 26735062
  60. Ruhlman TA, Chang WJ, Chen JJ, Huang YT, Chan MT, Jin Z, et al. NDH expression marks major transitions in plant evolution and reveals coordinate intracellular gene loss. *BMC Plant Biol.* 2015; 15:100. <https://doi.org/10.1186/s12870-015-0484-7> PMID: 25886915
  61. Ross TG, Barrett CF, Soto Gomez M, Lam VKY, Henriquez CL, Les DH, et al. Plastid phylogenomics and molecular evolution of Alismatales. *Cladistics.* 2016; 32: 160–178. <https://doi.org/10.1111/cla.12133>
  62. Barrett CF, Davis JI. The plastid genome of the mycoheterotrophic *Corallorhiza striata* (Orchidaceae) is in the relatively early stages of degradation. *Am. J. Bot.* 2012; 99: 1513–1523. <https://doi.org/10.3732/ajb.1200256> PMID: 22935364
  63. Niu Z, Xue Q, Zhu S, Sun J, Liu W, Ding X. The complete plastome sequences of four orchid species: insights into the evolution of Orchidaceae and the utility of plastomic mutational hotspots. *Front Plant Sci.* 2017; 8: 715. <https://doi.org/10.3389/fpls.2017.00715> PMID: 28515737
  64. Zhitao N, Shuying Z, Jiajia P, Ludan L, Jing S, Xiaoyu D. Comparative analysis of *Dendrobium* plastomes and utility of plastomic mutational hotspots. *Sci Rep.* 2017; 7: 2073. <https://doi.org/10.1038/s41598-017-02252-8> PMID: 28522861
  65. Merckx V, Bakker F, Huysmans S, Smets E. Bias and conflict in phylogenetic inference of myco-heterotrophic plants. *Cladistics.* 2009; 25: 64–77. <https://doi.org/10.1111/j.1096-0031.2008.00241.x>
  66. Merckx V, Freudenstein JV. Evolution of mycoheterotrophy in plants: a phylogenetic perspective. *New Phytologist.* 2010; 185: 605–609. <https://doi.org/10.1111/j.1469-8137.2009.03155.x> PMID: 20356335