

RESEARCH ARTICLE

New plastomes of eight *Ipomoea* species and four putative hybrids from Eastern Amazon

Marcele Laux¹, Renato R. M. Oliveira^{1,2}, Santelmo Vasconcelos^{1*}, Eder S. Pires¹, Talvâne G. L. Lima¹, Mayara Pastore³, Gisele L. Nunes¹, Ronnie Alves¹, Guilherme Oliveira¹

1 Instituto Tecnológico Vale, Belém, Pará, Brazil, **2** Programa Interunidades de Pós-Graduação em Bioinformática, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, **3** Programa de Pós-Graduação em Botânica Tropical, Museu Paraense Emílio Goeldi, Belém, Pará, Brazil

* santelmo.vasconcelos@itv.org



Abstract

Ipomoea is a large pantropical genus globally distributed, which importance goes beyond the economic value as food resources or ornamental crops. This highly diverse genus has been the focus of a great number of studies, enriching the plant genomics knowledge, and challenging the plant evolution models. In the Carajás mountain range, located in Eastern Amazon, the savannah-like ferruginous ecosystem known as *canga* harbors highly specialized plant and animal populations, and *Ipomoea* is substantially representative in such restrictive habitat. Thus, to provide genetic data and insights into whole plastome phylogenetic relationships among key *Ipomoea* species from Eastern Amazon with little to none previously available data, we present the complete plastome sequences of twelve lineages of the genus, including the *canga* microendemic *I. cavalcantei*, the closely related *I. marabaensis*, and their putative hybrids. The twelve plastomes presented similar gene content as most publicly available *Ipomoea* plastomes, although the putative hybrids were correctly placed as closely related to the two parental species. The *cavalcantei-marabaensis* group was consistently grouped between phylogenetic methods. The closer relationship of the *I. carnea* plastome with the *cavalcantei-marabaensis* group, as well as the branch formed by *I. quamoclit*, *I. asarifolia* and *I. maurandioides*, were probably a consequence of insufficient taxonomic representativity, instead of true genetic closeness, reinforcing the importance of new plastome assemblies to resolve inconsistencies and boost statistical confidence, especially the case for South American clades of *Ipomoea*. The search for k-mers presenting high dispersion among the frequency distributions pointed to highly variable coding and intergenic regions, which may potentially contribute to the genetic diversity observed at species level. Our results contribute to the resolution of uncertain clades within *Ipomoea* and future phylogenomic studies, bringing unprecedented results to *Ipomoea* species with restricted distribution, such as *I. cavalcantei*.

OPEN ACCESS

Citation: Laux M, Oliveira RRM, Vasconcelos S, Pires ES, Lima TGL, Pastore M, et al. (2022) New plastomes of eight *Ipomoea* species and four putative hybrids from Eastern Amazon. PLoS ONE 17(3): e0265449. <https://doi.org/10.1371/journal.pone.0265449>

Editor: Shilin Chen, Chinese Academy of Medical Sciences and Peking Union Medical College, CHINA

Received: October 14, 2021

Accepted: March 1, 2022

Published: March 17, 2022

Copyright: © 2022 Laux 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: The twelve plastomes described here are available in GenBank under the accession numbers MK086044 to MK086056.

Funding: This work was supported by Vale S.A. (RBRS000603.85) to SV, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES (88887.130628/2016-00) and Research Councils UK, GCRF-RCUK (BB/P027849/1) to GO. GO is a Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq fellow (307479/

2016-1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Introduction

Located in Eastern Amazon, the Carajás mountain range harbors the altitude ferruginous savannah-like ecosystem known as *canga*, characterized by shallow soils (0–10 cm) and potentially phytotoxic levels of metals [1]. The *canga* plateaus are isolated from each other by matrixes of rainforest and show high levels of both endemism and species turnover as the result of environmental heterogeneity [1–3]. Previous studies have found evidence that the *canga* soil properties are restrictive for the seedling establishment, working as primary drivers of vegetation composition and structure in the ecosystem [4–6]. Several studies have been performed with plant and animal populations sampled in the Carajás National Forest, or Floresta Nacional de Carajás, progressively improving the knowledge about the biodiversity and genetic diversity patterns from native and endemic populations, as well as to investigate the occurrence of hybridization and speciation events [2, 5–11].

Ipomoea species are broadly distributed in the world, especially in the tropics and subtropics [12–14], where about 600–700 species are known [12], or about 800–900 species in the broader generic concept based on recent phylogenetic analysis, considering the genera *Argyriaea*, *Astripomoea*, *Blinkworthia*, *Lepistemon*, *Lepistemonopsis*, *Mina*, *Paralepistemon*, *Rivea*, *Stictocardia* and *Turbina* nested within *Ipomoea* [11, 15, 16]. The genus *Ipomoea* is recognized within the tribe Ipomoeae, which has been subdivided into two main clades [11, 16, 17], with long history of taxonomic and nomenclatural problems [16, 18]. Widely known as “morning glories” or “bindweeds”, the species of the genus present a high commercial value, either as ornamental plants or food crops [13, 19], with the sweet potato (*I. batatas* (L.) Lam.) being one of the most widely cultivated species [20–22]. This genus has also served as a model for understanding many evolutionary questions and elucidate inter and intra-species relationships among populations [6, 9, 15, 18, 22–29].

The most recent inventory of the flora of the *cangas* of Carajás presented 116 seed plant families, encompassing 856 species [30]. Convolvulaceae is represented by eight genera and 20 species in Carajás, 12 of which belonging to the genus *Ipomoea* (Fig 1), including the flagship species *I. cavalcantei* D.F. Austin, known as “flor de Carajás” (flower of Carajás), endemic to the ferruginous fields from Carajás North ridge (N1 to N5) and considered an endangered species [30, 31]. *Ipomoea marabaensis* D.F. Austin & Secco is found in several ferruginous and granitic fields from Carajás, as well as in other rock outcrops in the states of Pará and Tocantins [2, 31]. *Ipomoea cavalcantei* and *I. marabaensis* are sister species, with similar genome sizes, shared chloroplast polymorphisms and overlaps in gene allele distributions [5]. The two species occur in sympatry in the *canga* sites N4 and N5, however, *I. cavalcantei* is common in N4, whereas *I. marabaensis* appears as small groups closer to the *canga*-forest boundaries [6]. Both species are clambering shrubs, perennials, and present elliptical, oblong to obovate leaves, being remarkably similar in terms of vegetative traits. They are mostly differentiated by the flower morphology, with a hypocrateriform deep red corolla adapted to hummingbird pollination in *I. cavalcantei*, and a campanulate to infundibuliform light pink to lilac corolla adapted to bee pollination in *I. marabaensis* [6, 12, 31, 32] (Fig 1C and 1D). Recently, a preliminary phylogenetic analysis was published, using a concatenated alignment of seven chloroplast genes and positioning *I. cavalcantei* and *I. marabaensis* within the Murucoides clade [5].

Considering that the modern plant classification systems [33] relies heavily on molecular data, which are often of chloroplast origin, the plastid genomes are efficient data source for building phylogenies on a broad scale and set species boundaries, inter-population variation, and gene flow at a local scale [34–38]. Also, recent studies have shown the higher variability in phylogenetic signal and different rates of evolution throughout the plastid regions and genes, as previously expected by single locus interpretations [39–44]. Therefore, the assembly of



Fig 1. Diversity of *Ipomoea* species and putative hybrids analyzed in this work. The canga environment (A); *Ipomoea cavalcantei* × *I. marabaensis* putative hybrids (B); *I. cavalcantei* (C); *I. marabaensis* (D); *I. carnea* (E); *I. triloba* (F); *I. quamoclit* (G); *I. goyazensis* (H); *I. asarifolia* (I and J); and *I. maurandioides* (K). Photos by Mayara Pastore (A, C-G, I-K), Pedro L. Viana (B) and Marcos E. L. Lima (H).

<https://doi.org/10.1371/journal.pone.0265449.g001>

complete plastomes have been extremely useful for providing an abundance of additional characters that can be used to resolve polytomies in phylogenetic trees [18, 45–48] and boost statistical confidence in deeply branching clades [49, 50].

As several individuals with intermediate phenotypes between *I. cavalcantei* and *I. marabaensis* have been observed at the N4 plateau (Table 1; Fig 1B), where the geographic ranges of

Table 1. List of *Ipomoea* lineages studied in the present work, including sampling information and plastome sequencing data for each analyzed specimen.

Species	Sample number	Collection information	GenBank accession	Sequenced reads	Mapped reads	Plastome size (bp)	Mean coverage
<i>I. asarifolia</i>	ITV3285	E. Babiychuk, s.n.; Salinópolis, Pará, Brazil	MK086048	45,993,456	807,204	160,589	760.2
<i>I. carnea</i>	ITV2324	E. Babiychuk, s.n.; Salinópolis, Pará, Brazil	MK086049	71,273,986	2,241,650	160,819	2125.4
<i>I. cavalcantei</i>	ITV3206	E. Babiychuk, s.n.; N1, Serra Norte, Parauapebas, Pará, Brazil	MK086050	517,935,995	745,649	161,563	714.8
<i>I. goyazensis</i>	ITV4320	L.V. Vasconcelos, 1116; N4, Serra Norte, Parauapebas, Pará, Brazil	MK086051	55,417,682	243,786	160,414	229.8
<i>I. marabaensis</i>	ITV2328	E. Babiychuk, s.n.; Serra do Tarzan, Canaã dos Carajás, Pará, Brazil	MK086052	410,313,272	465,351	161,324	440.2
<i>I. maurandioides</i>	ITV4245	E. Babiychuk, s.n.; Serra Norte, Parauapebas, Pará, Brazil	MK086053	29,932,159	501,367	161,242	472.7
<i>I. quamoclit</i>	ITV4995	M.G.C. Nogueira, 665; Redenção, Pará, Brazil	MK086054	83,709,896	1,248,317	160,836	1175.4
<i>I. triloba</i>	ITV4963	M.G.C. Nogueira, 651; Redenção, Pará, Brazil	MK086056	37,827,247	830,975	161,835	774.4
<i>I. cavalcantei</i> × <i>I. marabaensis</i> H1	ITV2181	R.M. Harley, 57491; N4, Serra Norte, Parauapebas, Pará, Brazil	MK086044	55,410,366	798,743	161,495	754.6
<i>I. cavalcantei</i> × <i>I. marabaensis</i> H2	ITV280	F. Santos, E2FLO.05; N4, Serra Norte, Parauapebas, Pará, Brazil	MK086047	59,399,262	946,578	160,765	899.9
<i>I. cavalcantei</i> × <i>I. marabaensis</i> H3	ITV2294	E. Babiychuk, s.n.; N4, Serra Norte, Parauapebas, Pará, Brazil	MK086045	65,210,419	551,682	160,974	527.5
<i>I. cavalcantei</i> × <i>I. marabaensis</i> H4	ITV2295	E. Babiychuk, s.n.; N4, Serra Norte, Parauapebas, Pará, Brazil	MK086046	65,302,820	750,030	161,948	709.0

<https://doi.org/10.1371/journal.pone.0265449.t001>

both species overlap [5, 6, 31], we aimed to describe and analyze the complete plastome sequences from *I. cavalcantei*, *I. marabaensis* and their putative hybrids, plus six other *Ipomoea* species from Eastern Amazon, in order to better understand the phylogenetic relationships among the lineages of the genus in the region, besides providing genetic information to direct conservation planning in the unique *cangas* of the Serra de Carajás.

Materials and methods

DNA extraction and sequencing

DNA extractions were carried out following the automated protocol previously described [51], with approximately 20 mg of plant material collected in NaCl-saturated CTAB solution [52]. Afterwards, paired-end libraries were constructed from 50 ng of DNA. Samples were subjected to a step of enzymatic and random fragmentation in which the DNA were simultaneously fragmented and bound to adapters using the QXT SureSelect kit (Agilent Technologies) according to the manufacturer's instructions. The fragmented DNA was purified and subjected to an amplification reaction using primers complementary to the adapters. Next, the libraries were quantified using the Qubit 3.0 (Invitrogen) fluorimeter and checked for fragments size in the 2100 Bioanalyzer (Agilent Technologies). Then, the libraries were diluted in a solution of 0.1% Tris-HCl and Tween and pooled. The sequencing run was performed with a NextSeq 500 v2 kit high-output (300 cycles).

Plastome assembly and annotation

The quality of the generated dataset was checked using the FastQC 0.11.5 tool [53], and the adapters were removed using Trimmomatic 0.38 [54]. The assembly was performed using a

hybrid strategy with the de novo NOVOPlasty (NP) v2.6.3 assembler [55] and selected contigs from SPAdes v3.11 [56]. The NP config file was set as follows: insert size 300, read length 150, type chloro, genome range 120k-200k, K-mer 39, paired-end mode and original dataset with the full content of the DNA extracted as input. The seeds initially used for contig extension in NP assembler were genes from the reference genome *I. nil* (L.) Roth [57], and subsequently genes from the assembled plastomes, in a recursive strategy. The main seeds used for capture and contig extension were (in order of effectiveness) the complete sequences of the genes *psbK*, *psbB*, *psaC*, *ndhF*, *rbcL*, *psbC*, *rpoB*, *rrn23*, *trnH*, *ycf1*, *rps15*, *matK*, *rpl32*, and some partial sequences, as the junctions between the two inverted repeat (IR) regions and the small single copy (SSC) region, and the *ycf1-rps15* intergenic spacer. The resulting contigs were assembled in Geneious R11 (Biomatters) and the consensus sequences were annotated in the CpGAVAS web server [58].

SPAdes assembler is not designed to deal with the chloroplast (cp) genome architecture, especially because the inverted repeats, so the larger contigs generated were about 27–89 kbp long. To select chloroplast contigs in the SPAdes assembly, the contigs were chosen according to a two-step selection. The chloroplast DNA fragments are expected to be more abundant in the total DNA extracted, since such organelle is found in high numbers within each plant cell [59, 60]. A coverage cutoff was applied according to the overall contigs depth of coverage (DP) median. The selected high DP contigs were subsequently aligned to the plastid NCBI database (<ftp://ftp.ncbi.nlm.nih.gov/refseq/release/plastid/>) using MegaBLAST [61] with an e-value of 1e-5 and minimum percent identity of 85. The chloroplast contigs sequences were then extracted from the original SPAdes output.

Most plastomes were assembled only with NP, but for manual intervention and gap filling, we used the selected chloroplast contigs from SPAdes. For each sample, after the first draft genome was entirely assembled and annotated in CpGAVAS, the IRs were extracted and pair-wise aligned. The NP and SPAdes selected contigs were mapped against the draft genome, and the contigs which fell in such regions were used to guide the sequence edition. The final draft genome was then checked and edited using Artemis v18 [62] for curation. The twelve original datasets were re-mapped against the final plastomes using Bowtie2 [63] to check for coverage uniformity throughout the entire plastome. The duplicates were removed, and the mapped reads were used to calculate the average coverage for each assembly. The variable regions were located using the Geneious Find Variation tool, adopting *I. cavalcantei* as the reference genome for the population in study, since the four putative hybrids were closely related to it. RepeatMasker [64] was used to identify and locate the di- to pentameric and some hexameric simple repeat sequences with more than 20 bp, using default parameters. The circular map was generated using OGDRAW [65].

Phylogenetic analysis of concatenated genes

Thirteen complete genes from the 12 *Ipomoea* plastomes assembled, plus seven *Ipomoea* references (*I. nil*, *I. purpurea* (L.) Roth, *I. trifida* (Kunth) G. Don, *I. batatas*, *I. hederacea* Jacq, *I. lacunosa* L., *I. triloba* L.), and one outgroup (*Solanum dulcamara* (Moench) Dunal) were aligned using MAFFT v7 [66] multiple aligner using default settings, and concatenated using Geneious R11. The concatenated alignment of 13 genes selected according to interspecies variability, molecular function, chloroplast compartments, and conservation level (*atpA* 1524 bp, *ccsA* 978 bp, *matK* 1536 bp, *ndhB* 2222 bp, *ndhD* 1539 bp, *ndhG* 531 bp, *petA* 963 bp, *psaA* 2253 bp, *psaC* 246 bp, *psbD* 1062 bp, *rbcL* 1443 bp, *rpoB* 3213 bp and *rps15* 273 bp) was the input for the phylogenetic inference under the maximum likelihood (ML) criteria with RAxML

v8.2.8 [67], using the rapid bootstrapping (1,000 replications) and search for the best-scoring ML tree.

K-mer frequency distribution

The *Ipomoea* plastomes were also investigated according to the k-mer frequency distribution using the AAF (Alignment and Assembly-Free) software [68] and custom scripts available in [S1 Data](#). AAF reconstructs phylogeny from a distance matrix based on the proportion of shared k-mers between each sample. The k-mer lengths of 25, 31 and 35 were compared according to the support values and congruence with the ML tree generated with the concatenated genes matrix, and the k-mer length of 31 was adopted. The k-mers presenting frequencies with the highest dispersion values (sd function) among the samples were selected and extracted using a R custom script, aligned to consensus sequences and the plastome regions where the high dispersion k-mers were identified were referred as H-disp regions, potentially representing highly variable plastome sites, showing significantly distinct frequency patterns. Using the dispersion among the frequencies as a genetic diversity value parameter, we were able to progressively reduce the number of k-mers, selecting those with the highest contribution.

Results

Plastome assembly and comparative analysis

The plastomes were deposited in GenBank under the accession numbers MK086044-MK086056 ([Table 1](#)). All datasets presented high quality on average and high mean coverage, according to the respective assembled plastomes ([Table 1](#)). A total of 2,693 NP contigs and 11,442 SPAdes contigs were used to perform the assemblies of the 12 specimens. SPAdes generated more contigs, but a few consensus sequences by dataset.

The average length of the chloroplast genomes was 161 kbp, presenting a total of 123 genes, including 80 protein coding genes, one partial *infA* gene, four rRNA genes, and 33 tRNA genes, with nine tRNA genes presenting at least one duplicate ([Table 2](#)). The 12 plastomes assembled displayed the usual circular quadripartite structure ([Fig 2](#)), including one large single copy (LSC, 87 kbp in average), one SSC (12 kbp in average), and two IRs (IRa and IRb, 30 kbp in average). The average GC-content of the LSC region was 36%, and 40.8% in the IRs. Most plastomes presented a similar genome structure, with slight rearrangements in the SSC regions ([Fig 3](#)). *Ipomoea marabaensis* ([Fig 2B](#)) and *I. carnea* Jacq. presented the same SSC gene order as the seven *Ipomoea* references, displaying the complete *ndhA* gene (two exons) in SSC-IRb junction, while for remaining assembled plastomes the complete copy of the *ndhA* gene was located in the IRa-SSC junction ([Fig 2A](#)). All plastomes showed two complete copies of the *ycf1* located within the inverted repeats, close to each SSC junction.

Thirty-five and 33 polymorphic simple repeats were identified in *I. cavalcantei* and *I. marabaensis*, respectively, mostly located in intergenic regions ([Fig 2](#)), seven located in the coding regions of the plastomes of *I. cavalcantei* (*accD*, *atpF*, *clpP*, *ndhA*, *trnR*, *trnY* and *ycf3*.) and five of *I. marabaensis* (*accD*, *atpF*, *clpP*, *trnY* and *ycf3*). All chloroplast genomes presented a homogeneous coverage throughout the entire plastome and presented mean coverage above 200× in the remapping analysis ([Table 1](#)). The genes and regions with higher interspecific variability among the plastomes analyzed were *accD*, *clpP*, *matK*, *ndhB-rps7*, *ndhG*, *psaA*, *rpoC1*, *rps16* and *ycf2*.

The H-disp regions with higher contribution (frequency distribution dispersion) were the *ycf1* repeat, also found in the reference genomes, at the same position, and the H-disp regions found in *rps15* gene, in the intergenic region between *ndhH* and *rps15*, *ycf2* and *trnL* and

Table 2. Gene content in the assembled plastome of *Ipomoea cavalcantei*.

Cell function	Gene codification	Observed genes
Self-replication	DNA-directed RNA polymerase	rpoA, rpoB, rpoC1 and rpoC2
Self-replication	Large subunit ribosomal protein	rpl2, rpl14, rpl16, rpl20, rpl22, rpl23, rpl32, rpl33 and rpl36
Self-replication	rRNA genes	rrn4.5 ^r , rrn5i ^r , rrn16 ^r and rrn23 ^r
Self-replication	Small subunit ribosomal protein	rps2, rps3, rps4, rps7 ^r , rps8, rps11, rps12 ^{r,s} , rps14, rps15 ^r , rps16, rps18 and rps19
Self-replication	tRNA genes	trnA ^{i,r} , trnC, trnD, trnE, trnF, trnGt, trnH, trnI ^{d,i,r} , trnL ^{i,r,t} , trnK ⁱ , trnM ^d , trnN ^r , trnP, trnQ, trnR ^t , trnS ^t , trnT ^d , trnV ^{d,i,r} , trnW and trnY ^d
Photosynthesis	ATP synthase	atpA, atpB, atpE, atpF, atpH and atpI
Photosynthesis	Cytochrome b6/f complex	petA, petB, petD, petG, petL and petN
Photosynthesis	NADH oxidoreductase	ndhA ^{p,r} , ndhB ^r , ndhC, ndhD, ndhE, ndhF, ndhG, ndhH ^R , ndhI, ndhJ and ndhK
Photosynthesis	Photosystem I	psaA, psaB, psaC, psaI and psaJ
Photosynthesis	Photosystem II	psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT and psbZ
Photosynthesis	Rubisco	rbcl
Other functions	Acetyl-CoA-carboxylase subunit	accD
Other functions	c-type cytochrome synthesis	ccsA
Other functions	Envelop membrane protein	cemA
Other functions	Maturase	matK
Other functions	Protease	clpP
Other functions	Translational initiation factor	infA ^p
Putative genes	Conserved ORF	ycf1 ^t , ycf2 ^t , ycf3, ycf4 and ycf15 ^r

^dduplicated;ⁱintron;^ppartial;^rinverted repeat;^strans-splicing; and^ttriplicated

<https://doi.org/10.1371/journal.pone.0265449.t002>

between rps7 and ndhB genes. The H-disp between rps7 and ndhB was not present in *I. goyazensis*. The H-disp found in the ycf1 gene consisted in a 48 bp region which occurred in two copies, except for *I. carnea*, which contained only a single copy (Table 3). We found the ycf1 H-disp repeat in the reference genomes, even for *I. batatas*, which lacks the ycf1 gene [21] but presents the two copies of the repeat in the same position, between trnN and rps15.

Phylogenetic reconstruction

Both tree reconstruction methods showed high support for the *cavalcantei-marabaensis* group and *I. carnea* was the closest related species among the analyzed specimens, also grouping *I. asarifolia* and *I. maurandioides*, with *I. quamoclit* being placed in the same branch. Both

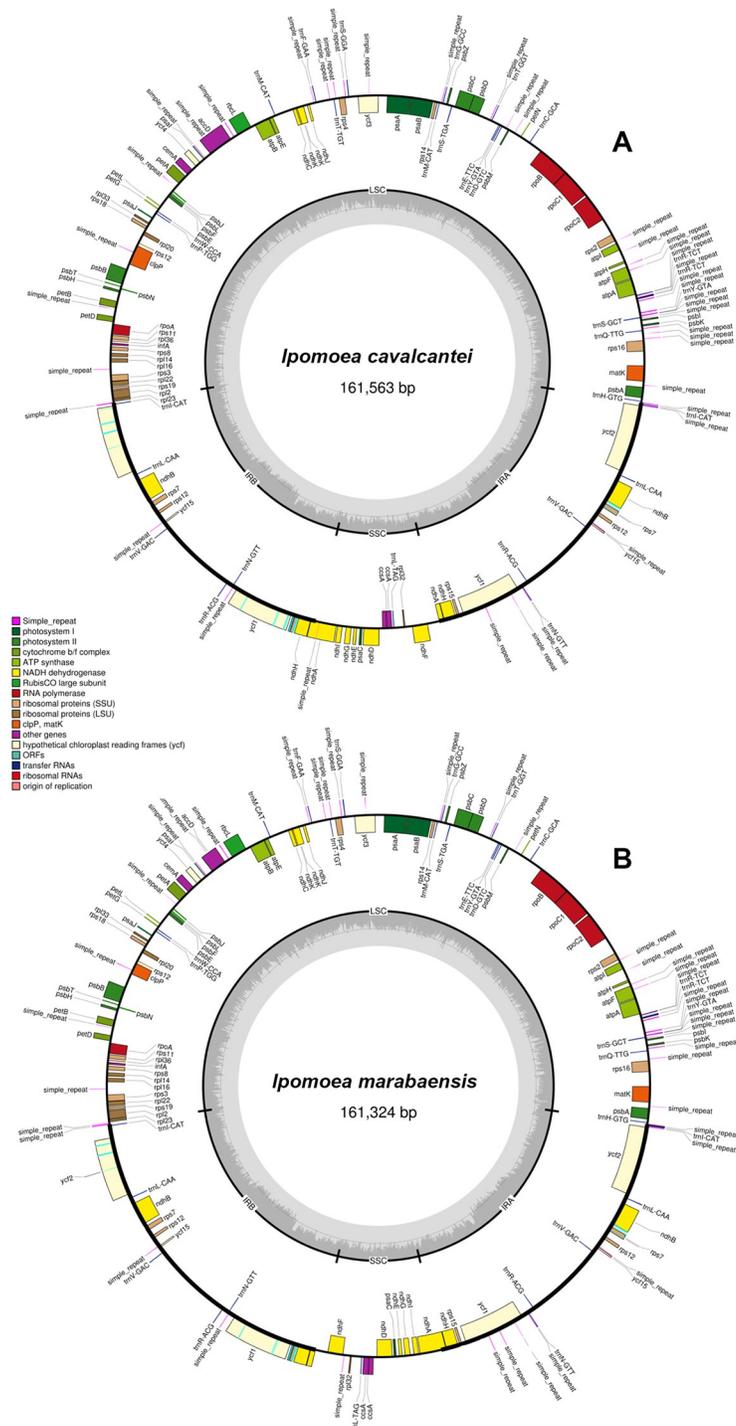


Fig 2. Circular map of the plastomes of *Ipomoea cavalcantei* (A) and *I. marabaensis* (B), showing the typical quadripartite structure indicated in the inner circle (IRB, LSC, IRA and SSC). The gene categories are shown in colored boxes and simple repeats sites are indicated in pink.

<https://doi.org/10.1371/journal.pone.0265449.g002>

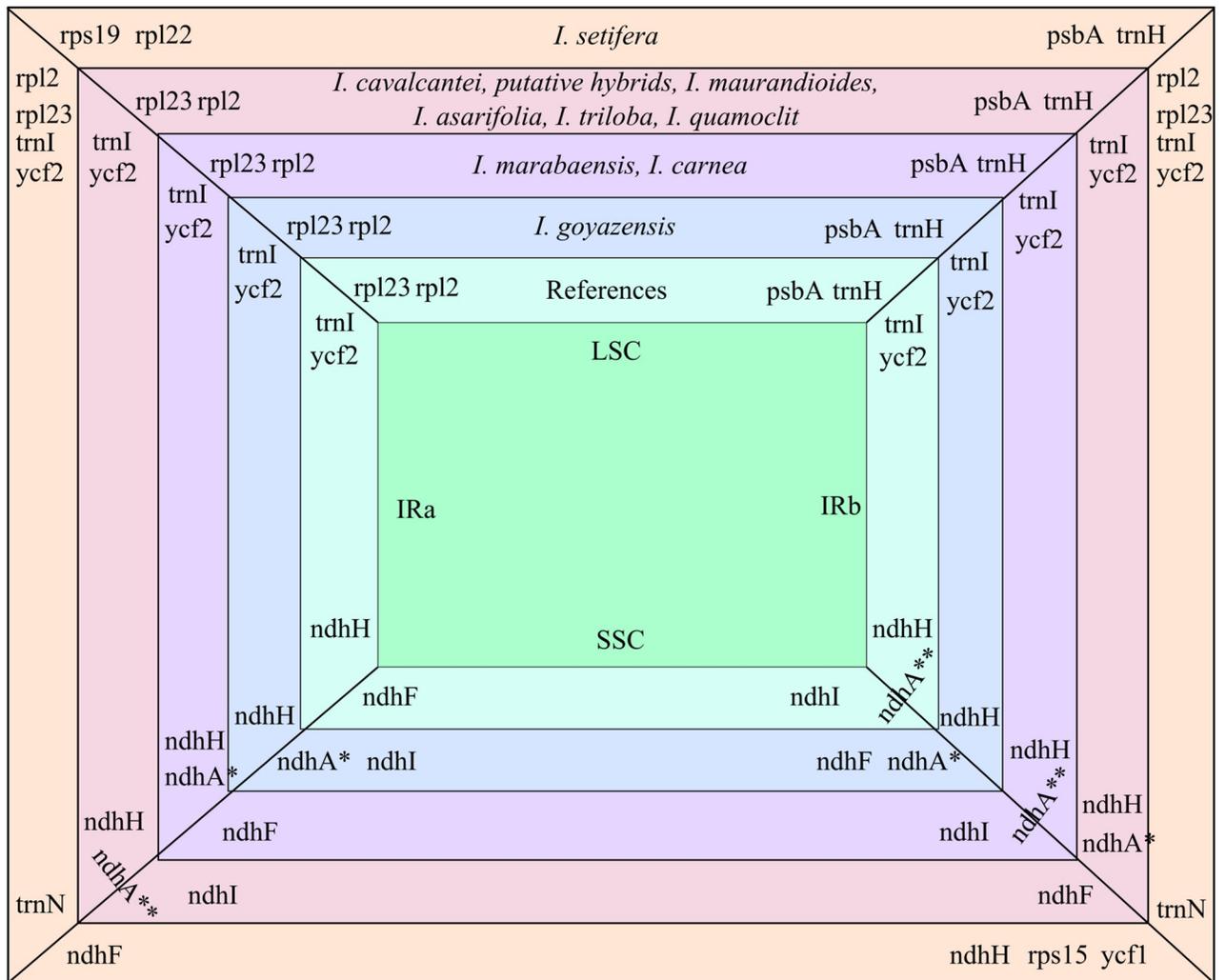


Fig 3. Comparison of the four junctions of the chloroplast quadripartite structure, grouping species with similar junctions, based on [21]. *ndhA with one exon; **ndhA with two exons.

<https://doi.org/10.1371/journal.pone.0265449.g003>

Table 3. H-disp regions captured through the k-mer frequency dispersion among the samples.

Plastome region	H-disp
5'-ndhB	TTCATTCTGTACATGCCAGCTCATGAATTAGT
ndhH-rps15	CTATTTGTTTCGTTTCAATTTGA
rps7-ndhB	TCCAATTTCAAAAAAAAAATCCAATTGTGT
rps15	GTACGTTATAAGAATTAATTGAGAAATGGATATTCGAGAGACAAA
trnI-ycf2	GTTTCAAGTAATGTTTGATCAATTACGTATTTATACACGTATTCGTATTAATCAATTTTGGATGAATTA
ycf1 repeat	GAAGAGATCAATCCCAGCAGTAATCAAAGACTCCAATTGGGACTAAT

<https://doi.org/10.1371/journal.pone.0265449.t003>

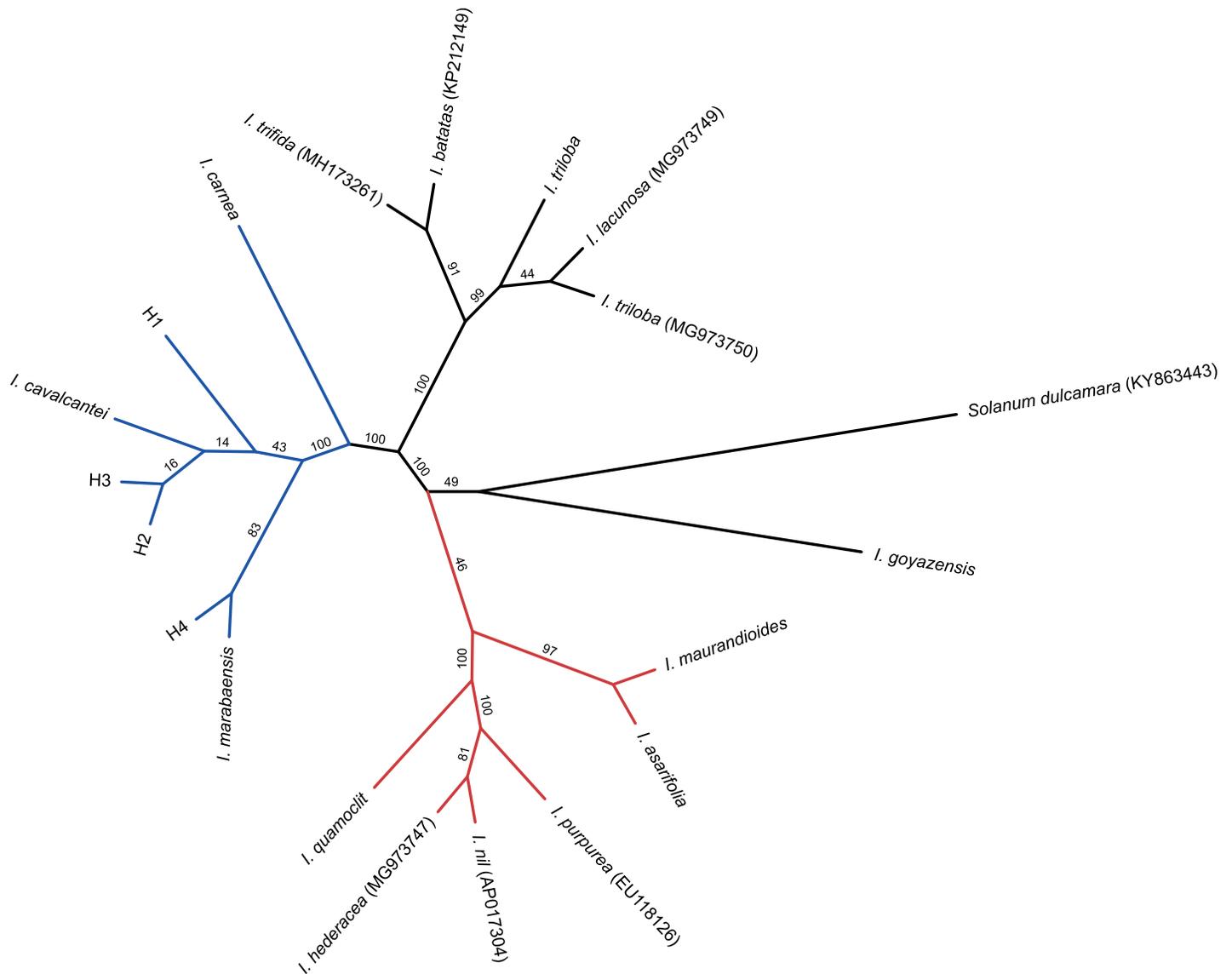


Fig 4. Phylogenetic relationships among *Ipomoea* species. Unrooted ML phylogenetic tree of the 12 *Ipomoea* sequenced plastomes, plus five *Ipomoea* plastomes previously available in public databases and one outgroup from Solanales (*Solanum dulcamara*) using RAxML with the GTR+G model and rapid bootstrap with 1,000 replicates, based on a concatenated matrix of 13 genes. The *cavalcantei-marabaensis* group is highlighted in blue and the *quamoclit-asarifolia-maurandioides* in brown.

<https://doi.org/10.1371/journal.pone.0265449.g004>

methods also showed the distinct relationship between the putative hybrids and the two sister species.

The putative hybrids H1, H2 and H3 were closely related to *I. cavalcantei*, while H4 showed a higher similarity to the *I. marabaensis* plastome, according to the ML tree (Fig 4). On the other hand, the two parental plastomes were closely related to each other according to the AAF tree, and the H4 appeared as sister to the remaining lineages of the *cavalcantei-marabaensis* group (Fig 5). Three larger groups were formed according to the ML trees, being the *cavalcantei-marabaensis* group (BS = 100), the weakly supported *asarifolia-maurandioides-quamoclit* group (BS = 49) and the *triloba-batatas* group (BS = 100) (Fig 4). However, among those, only the *cavalcantei-marabaensis* group was recovered, with the other lineages forming

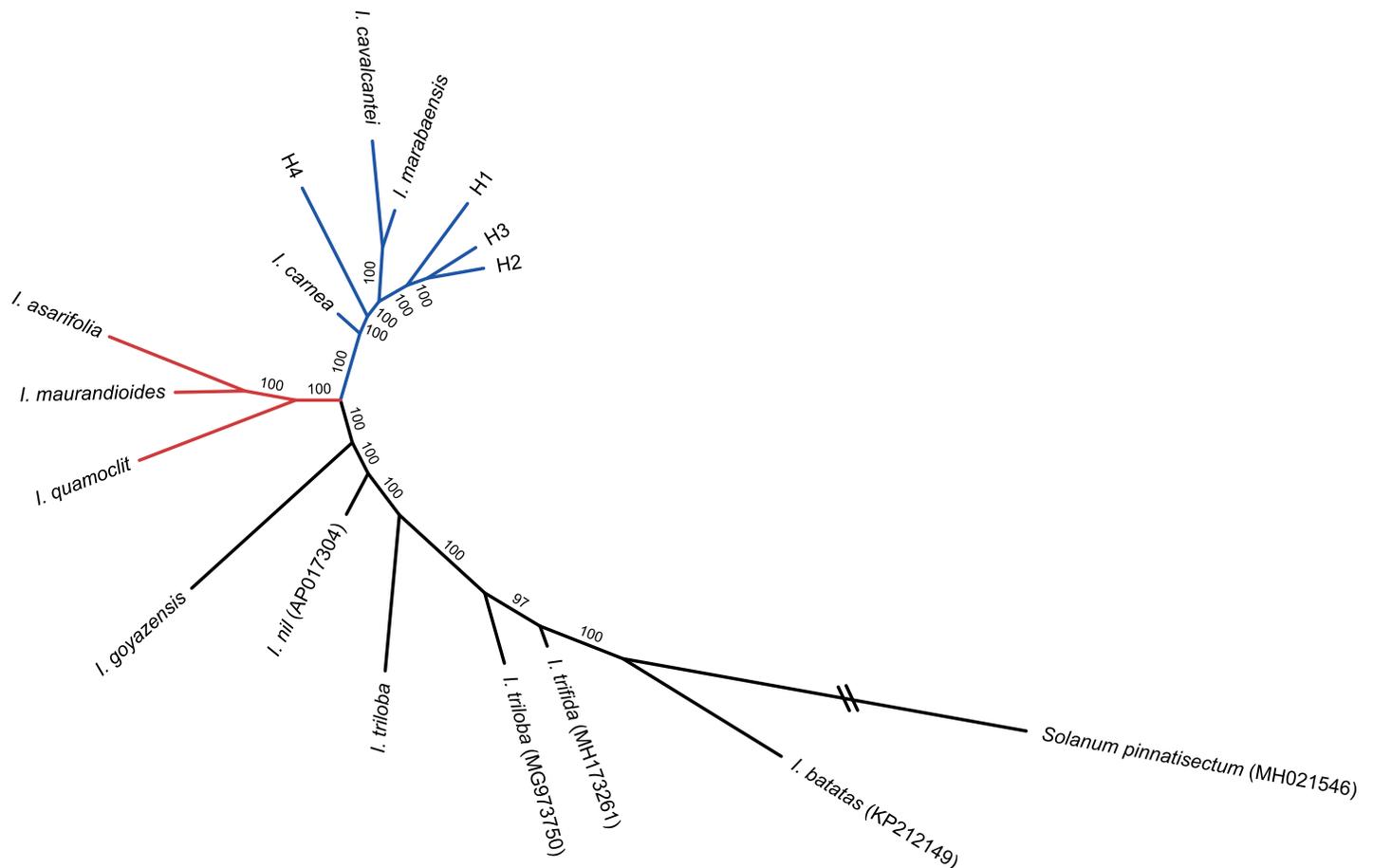


Fig 5. Dendrogram representing the genetic distances among *Ipomoea* species. Unrooted genetic distance tree based on k-mer frequency distribution using the AAF approach, using the 12 *Ipomoea* sequenced plastomes, four *Ipomoea* plastomes previously available in public databases and one outgroup from Solanales (*Solanum pinnatisectum*). The *cavalcantei-marabaensis* group is highlighted in blue and the *quamoclit-asarifolia-maurandioides* in brown.

<https://doi.org/10.1371/journal.pone.0265449.g005>

different clusters (Fig 5). *Ipomoea asarifolia* and *I. maurandioides* were closer between each other in both methods (BS > 97), but the grouping with *I. quamoclit* was inconsistent (BS = 46) in the ML analysis, showing *I. quamoclit* closer (BS > 81) to three *Ipomoea* references from the Pharbitis clade (*I. hederacea*, *I. nil* and *I. purpurea*). *Ipomoea triloba* was grouped in a highly supported branch with four references, all belonging to the Batatas clade, but showing an inconsistent polyphyletic pattern with a reference of the same species, *I. triloba* MG973750 (Figs 4 and 5). Most of the differences between the branches were found in intergenic regions. Among the coding regions, we highlight the *accD* and *clpP* genes, which were the most variable genes.

The phylogenetic reconstruction based on the assembly-free shared k-mers (AAF) generated a total of 1,187,671 k-mers for all 12 datasets and two major groups could be observed (Fig 5). Similar to the ML tree, the *cavalcantei-marabaensis* group was highly supported (BS = 100) and *I. carnea* was, again, the closest relative. The *asarifolia-maurandioides-quamoclit* branch was not topologically consistent and the long branches reflect the large genetic distance between *I. asarifolia* and *I. quamoclit* (Fig 5).

Discussion

Interspecific genetic diversity

With the advent of the next-generation sequencing, a substantial increase in number of organelle genomes newly assembled has been observed, along with the remarkable expansion of studies investigating new evolutionary and structural diversification patterns, which reinforce the importance of the genome-wide genetic diversity exploration and gene variability, especially below the order and family levels [39, 42, 43]. We identified potentially variable sites according to single nucleotide variants and the H-dips regions, which pointed to genes harboring potentially informative genetic diversity sites. The *accD* gene was among the most variable genes, as previously observed for *Ipomoea* plastomes [26, 40, 69]. The genes *clpP*, *ndhA*, *rps16*, *ycf1* and *ycf2* also showed high interspecific variability, especially *clpP*, which is highly variable and presents a remarkable acceleration in plastome evolutionary rate [44, 69, 70]. Shared chloroplast polymorphisms between *I. cavalcantei* and *I. marabaensis* in *rpoC1* was previously observed [5]. We could not find any reference for the *ycf1* H-disp repeat in the repeats databases, but the *ycf1* gene is considered a genomic marker [27, 71] and variable sites were also identified by different studies [26, 27, 29, 72]. The H-disp region found between *ndhH* and *rps15* was also observed in all plastomes, and already adopted as marker region for species discrimination in some studies [73, 74]. The *accD* and *ycf3* simple repeats were also observed in recently published *Ipomoea* plastomes [21, 27].

Phylogenetic considerations

The definition of infrageneric clades in *Ipomoea* is hampered by its extreme evolutionary lability in morphology and the widespread homoplasmy among the species, especially in highly diverse tropical ecosystems [13, 15, 18, 75, 76]. *Ipomoea cavalcantei* and *I. marabaensis* are sister species, sharing plastome types and alleles distribution, but a greater diversity was already documented in ITS2 alleles in *I. cavalcantei*, suggesting different diversification rates between the two species [5]. Despite recognized as two species belonging to the Murucoides clade, *I. cavalcantei* and *I. marabaensis* were not sampled in previous phylogenetic studies of *Ipomoea*, which included mainly better-known species for understanding the infrageneric classification. The Murucoides clade was further investigated, including *I. murucoides* Roem. & Schult. (Southern Mexico to Guatemala) and *I. polpha* R.W. Johnson (Australia) [18]. According to the most recent and comprehensive study of *Ipomoea* in Americas, both species were arbitrarily placed within the New World clade called clade A1, related morphologically to species of a smaller clade designated Arborescens [11].

According to recent studies concerning the organization, composition, and inheritance patterns of the organelle genomes, concatenated datasets may present strong variation of phylogenetic signal across the matrix, resulting in distinct or conflicting topologies [39, 42, 43]. Among the 13 genes composing the concatenated matrix, two were chosen for carrying H-disp regions (*ndhB* and *rps15*), two variable genes (*ndhG* and *psaA*), traditional markers (*matK* and *rbcL*), and seven more conserved genes related to self-replication, photosynthesis, and transmembrane molecular function (*atpA*, *ccsA*, *ndhD*, *petA*, *psaC*, *psbD* and *rpoB*). Both ML and AAF tree reconstruction methods applied in this work showed high support for the *cavalcantei-marabaensis* group, but the relationships among the internal nodes are still unclear. According to the ML method, the putative hybrid H4 was closer to *I. marabaensis* (BS = 83), while H1, H2 and H3 were closer to *I. cavalcantei*, but with low support values (Fig 4). The occurrence of hybridization may be one of the reasons for the low support internal branches observed [24, 77, 78], especially considering that an interspecific hybridization

between the sister species has already been confirmed, producing fertile offspring [6]. Both methods also placed *I. carnea* as the closest related species to the *cavalcantei-marabaensis* group among the analyzed plastomes, as well as *I. quamoclit* as the closest species to *I. asarifolia* and *I. maurandioides* plastomes, an incongruent relationship according to previous phylogenetic studies. Despite the AAF highly supported branch, *asarifolia-maurandioides-quamoclit* was artificially grouped probably due to an insufficient sampling representativity, and the long branches demonstrate the substantial genetic distance between the three species (Fig 5). In the ML trees, the *asarifolia-maurandioides-quamoclit* branch support was low (BS = 46) and *I. quamoclit* grouped with three references from Pharbitis clade in a high support internal branch. *Ipomoea asarifolia*, *I. maurandioides* and *I. quamoclit* are not closely related, belonging to different clades within the genus and were probably grouped due the lack of tropical and South American *Ipomoea* references [15], instead of true genetic closeness or low phylogenetic resolution [24, 77, 79, 80].

Morphological considerations

The vegetative characters combined with the hypocrateriform red corolla make *I. cavalcantei* very distinct, tackling problems about its clustering in the identification key with all species of the genus in the Americas [11]. The members of the *cavalcantei-marabaensis* group showed flowers with a sharp color gradient from lilac, considered an ancestral characteristic, to red, which evolved independently within the tribe Ipomoeae [18, 23, 81] (Fig 1). Both *I. cavalcantei* and *I. quamoclit* (Quamoclit clade) present a red hypocrateriform corolla, but the two species are not closely related, illustrating the independent diversification process among our sampled *Ipomoea* species plastomes [11, 12, 23]. Moreover, the same SSC gene order observed in *I. marabaensis*, *I. carnea* and the references, and the higher similarity of H4 putative hybrid with *I. marabaensis*, could be related to the heterogeneity in diversification rates [28], the relaxed radiation of Central-South America *Ipomoea* clades [15] or even to the color flower and pollination preferences [23].

The closeness of *I. carnea* and the sister species agrees with morphological studies because these species have the woody branches, subequal sericeous sepals, and seeds with long side trichomes. According to [82], *I. carnea* belongs to the series Jalapae, while [11] classified it within the sister clade Jalapa. *Ipomoea goyazensis*, the most genetically distant among the plastomes covered in this study (Figs 4 and 5), belongs to a different monophyletic group with about 30 species characterized by subequal coriaceous sepals, glabrous, usually convex and glabrous corolla [11]. The polyphyletic pattern observed for *I. triloba* and the same species reference plastome (Figs 4 and 5) is probably a result of low phylogenetic resolution captured by the plastomes or even due to biogeographic genetic variability. On the other hand, the whole branch grouped the described plastome of *I. triloba* with the references *I. batatas*, *I. trifida*, *I. triloba* MG973750 and *I. lacunosa*, all belonging to the section *Batatas*, confirmed as monophyletic group [11, 15, 18].

Furthermore, the complex plastome relationships observed within the *cavalcantei-marabaensis* group and among the eight sampled species may suggest that our newly assembled plastomes could be from species of a highly diverse South American *Ipomoea* clade, as already proposed, which phylogenetic relationships are currently poorly resolved [11, 15, 22, 25]. In S1 Table we present results of the allele distribution among five chloroplast genes, which show the distinct allele distribution among *cavalcantei-marabaensis* group and the remaining species. In the attempt to define natural clades, phylogenetic inferences were initially proposed and phylogenetic analyses have been testing the monophyly of traditional groups, providing advances towards a biogeographically and taxonomically representative phylogenetic

classification of the tribe Ipomoeae [11, 13–15, 17, 18, 76, 81–85]. However, it is important to emphasize our sampling here was focused on species within a definite geographic scale for conservation purposes. Therefore, the phylogenetic relationships among *I. cavalcantei*, *I. marabaensis* and the South American *Ipomoea* clade could only be robustly investigated by the inclusion of additional morphologically similar species in future phylogenetic reconstructions, therefore improving the taxonomic resolution with a more comprehensive genetic diversity within the systematic context.

Conclusions

As research efforts are starting to address the genetic diversity of plant species with restricted geographic distribution in the highly explored ironstone outcrops of the Serra dos Carajás [e.g. 5, 6, 9, 86, 87], the main focus has been directed to understand population dynamics of a few lineages employing NGS-derived SNP analyses. However, there is still the need for a better coverage of genetic studies on the several known endemics of the *cangas* [2]. Yet, phylogenetic diversity analyses using proper cladistic approaches and robust molecular data with high genomic coverage are not being developed in the same pace for the species of the region, with just a couple of previous studies available, including two rare *canga* quillworts [88, 89]. Thus, the results we presented here for *Ipomoea* species are important to provide a better view of the phylogenetic context of rare morning glories, especially considering the occurrence of hybridization and introgression between *I. cavalcantei* and *I. marabaensis* [6], two important components of the flora of the *cangas*, besides populating the genetic databases with information on one of the most diverse and economically important angiosperm genera.

Supporting information

S1 Table. Visual representation of the allele distributions detected for the genes *matK*, *ndhG*, *psaA*, *rpoC1* and *rps16*.

(XLSX)

S1 Data. H-disp regions script and analysis pipeline.

(PDF)

Acknowledgments

We thank the MG herbarium at Museu Paraense Emílio Goeldi for specimens consulted and incorporated.

Author Contributions

Conceptualization: Marcele Laux, Guilherme Oliveira.

Data curation: Marcele Laux, Renato R. M. Oliveira, Santelmo Vasconcelos, Gisele L. Nunes, Ronnie Alves, Guilherme Oliveira.

Formal analysis: Marcele Laux, Renato R. M. Oliveira, Eder S. Pires, Talvâne G. L. Lima, Mayara Pastore.

Funding acquisition: Santelmo Vasconcelos, Guilherme Oliveira.

Investigation: Marcele Laux, Ronnie Alves.

Methodology: Marcele Laux, Renato R. M. Oliveira, Eder S. Pires, Talvâne G. L. Lima, Gisele L. Nunes, Ronnie Alves.

Project administration: Santelmo Vasconcelos, Guilherme Oliveira.

Resources: Santelmo Vasconcelos, Gisele L. Nunes, Ronnie Alves, Guilherme Oliveira.

Software: Renato R. M. Oliveira.

Supervision: Santelmo Vasconcelos, Guilherme Oliveira.

Validation: Marcele Laux, Renato R. M. Oliveira, Santelmo Vasconcelos, Mayara Pastore, Gisele L. Nunes, Ronnie Alves, Guilherme Oliveira.

Visualization: Marcele Laux.

Writing – original draft: Marcele Laux.

Writing – review & editing: Renato R. M. Oliveira, Santelmo Vasconcelos, Mayara Pastore, Gisele L. Nunes, Ronnie Alves, Guilherme Oliveira.

References

1. Souza-Filho PWM, Giannini TC, Jaffé R, Giulietti AM, Santos DC, Nascimento WR, et al. Mapping and quantification of ferruginous outcrop savannas in the Brazilian Amazon: a challenge for biodiversity conservation. *PLoS One*. 2019; 14: e0211095. <https://doi.org/10.1371/journal.pone.0211095> PMID: 30653607
2. Giulietti AM, Giannini TC, Mota NFO, Watanabe MTC, Viana PL, Pastore M, et al. Edaphic endemism in the Amazon: vascular plants of the canga of Carajás, Brazil. *Bot Rev*. 2019; 85: 357–383.
3. Rodrigues M, Souza AIAF, Goulart SL, Kohler SV, Lima GCP, Anjos LJS, et al. Geostatistical modeling and conservation implications for an endemic *Ipomoea* species in the Eastern Brazilian Amazon. *J Nat Conserv*. 2020; 57: 125893.
4. Nunes JA, Schaefer CEGR, Ferreira Júnior WG, Neri AV, Correa GR, Enright NJ. Soil-vegetation relationships on a banded ironstone 'island', Carajás Plateau, Brazilian Eastern Amazonia. *An Acad Bras Cienc*. 2015; 87: 2097–2110. <https://doi.org/10.1590/0001-376520152014-0106> PMID: 26648541
5. Babiychuk E, Kushnir S, Vasconcelos S, Dias MC, Carvalho-Filho N, Nunes GL, et al. Natural history of the narrow endemics *Ipomoea cavalcantei* and *I. marabaensis* from Amazon canga savannahs. *Sci Rep*. 2017; 7: 7493. <https://doi.org/10.1038/s41598-017-07398-z> PMID: 28790327
6. Babiychuk E, Teixeira JG, Tyski L, Guimaraes JTF, Romeiro LA, Silva EF, et al. Geography is essential for reproductive isolation between florally diversified morning glory species from Amazon canga savannahs. *Sci Rep*. 2019; 9: 18052. <https://doi.org/10.1038/s41598-019-53853-4> PMID: 31792228
7. Salas RM, Viana PL, Cabral EL, Dessein S, Janssens S. *Carajasia* (Rubiaceae), a new and endangered genus from Carajás mountain range, Pará, Brazil. *Phytotaxa*. 2015; 206: 14–29.
8. Almeda F, Michelangeli FA, Viana PL. *Brasilianthus* (Melastomataceae), a new monotypic genus endemic to ironstone outcrops in the Brazilian Amazon. *Phytotaxa*. 2016; 273: 269–282.
9. Lanes ÉC, Pope NS, Alves R, Carvalho Filho NM, Giannini TC, Giulietti AM, et al. Landscape genomic conservation assessment of a narrow-endemic and a widespread morning glory from Amazonian savannas. *Front Plant Sci*. 2018; 9: 532. <https://doi.org/10.3389/fpls.2018.00532> PMID: 29868042
10. Isobe S, Shirasawa K, Hirakawa H. Current status in whole genome sequencing and analysis of *Ipomoea* spp. *Plant Cell Rep*. 2019; 38: 1365–1371. <https://doi.org/10.1007/s00299-019-02464-4> PMID: 31468128
11. Wood JRI, Muñoz-Rodríguez P, Williams BRM, Scotland RW. A foundation monograph of *Ipomoea* (Convolvulaceae) in the new world. *PhytoKeys*. 2020; 143: 1–823. <https://doi.org/10.3897/phytokeys.143.32821> PMID: 32577084
12. Austin DF. Novidades nas Convolvulaceae da flora amazônica. *Acta Amaz*. 1981; 11: 291–295.
13. Wilkin P. A morphological cladistic analysis of the Ipomoeae (Convolvulaceae). *Kew Bull*. 1999; 54: 853–876.
14. Miller RE, McDonald JA, Manos PS. Systematics of *Ipomoea* subgenus *Quamoclit* (Convolvulaceae) based on its sequence data and a Bayesian phylogenetic analysis. *Am J Bot*. 2004; 91: 1208–1218. <https://doi.org/10.3732/ajb.91.8.1208> PMID: 21653478
15. Muñoz-Rodríguez P, Carruthers T, Wood JRI, Williams BRM, Weitemier K, Kronmiller B, et al. A taxonomic monograph of *Ipomoea* integrated across phylogenetic scales. *Nat Plants*. 2019; 5: 1136–1144. <https://doi.org/10.1038/s41477-019-0535-4> PMID: 31712754

16. Eserman LA, Sosef MSM, Simão-Bianchini R, Utteridge TMA, Barbosa JCJ, Buriel MT, et al. Proposal to change the conserved type of *Ipomoea*, nom. cons. (Convolvulaceae). *Taxon*. 2020; 69: 1369–1371.
17. Stefanović S, Krueger L, Olmstead RG. Monophyly of the Convolvulaceae and circumscription of their major lineages based on DNA sequences of multiple chloroplast loci. *Am J Bot*. 2002; 89: 1510–1522. <https://doi.org/10.3732/ajb.89.9.1510> PMID: 21665753
18. Eserman LA, Tiley GP, Jarret RL, Leebens-Mack JH, Miller RE. Phylogenetics and diversification of morning glories (tribe Ipomoeae, Convolvulaceae) based on whole plastome sequences. *Am J Bot*. 2014; 101: 92–103. <https://doi.org/10.3732/ajb.1300207> PMID: 24375828
19. Bovell-Benjamin AC. Sweet potato: a review of its past, present, and future role in human nutrition. *Adv Food Nutr Res*. 2007; 52, 1–59. [https://doi.org/10.1016/S1043-4526\(06\)52001-7](https://doi.org/10.1016/S1043-4526(06)52001-7) PMID: 17425943
20. Khoury CK, Heider B, Castañeda-Álvarez NP, Achicanoy HA, Sosa CC, Miller RE, et al. Distributions, ex situ conservation priorities, and genetic resource potential of crop wild relatives of sweetpotato [*Ipomoea batatas* (L.) Lam., *I. series Batatas*]. *Front Plant Sci*. 2015; 6, 251. <https://doi.org/10.3389/fpls.2015.00251> PMID: 25954286
21. Yan L, Lai X, Li X, Wei C, Tan X, Zhang Y. Analyses of the complete genome and gene expression of chloroplast of sweet potato [*Ipomoea batata*]. *PLoS One*. 2015; 10: e0124083. <https://doi.org/10.1371/journal.pone.0124083> PMID: 25874767
22. Wadl PA, Olukolu BA, Branham SE, Jarret RL, Yencho GC, Jackson DM. Genetic diversity and population structure of the USDA sweetpotato (*Ipomoea batatas*) germplasm collections using GBSpoly. *Front Plant Sci*. 2018; 9: 1166. <https://doi.org/10.3389/fpls.2018.01166> PMID: 30186293
23. Streisfeld MA, Rausher MD. Genetic changes contributing to the parallel evolution of red floral pigmentation among *Ipomoea* species. *New Phytol*. 2009; 183: 751–763. <https://doi.org/10.1111/j.1469-8137.2009.02929.x> PMID: 19594698
24. Rosas-Guerrero V, Quesada M, Armbruster WS, Pérez-Barrales R, Smith SDW. Influence of pollination specialization and breeding system on floral integration and phenotypic variation in *Ipomoea*. *Evolution*. 2011; 65: 350–364. <https://doi.org/10.1111/j.1558-5646.2010.01140.x> PMID: 20874738
25. Roullier C, Duputié A, Wennekes P, Benoit L, Bringas VMF, Rossel G, et al. Disentangling the origins of cultivated sweet potato (*Ipomoea batatas* (L.) Lam.). *PLoS One*. 2013; 8: e62707. <https://doi.org/10.1371/journal.pone.0062707> PMID: 23723970
26. Park I, Yang S, Kim WJ, Noh P, Lee HO, Moon BC. The complete chloroplast genomes of six *Ipomoea* species and indel marker development for the discrimination of authentic pharbitidis semen (seeds of *I. nil* or *I. purpurea*). *Front Plant Sci*. 2018; 9: 965. <https://doi.org/10.3389/fpls.2018.00965> PMID: 30026751
27. Sun J, Dong X, Cao Q, Xu T, Zhu M, Sun J, et al. A systematic comparison of eight new plastome sequences from *Ipomoea* L. *PeerJ*. 2019; 7: e6563. <https://doi.org/10.7717/peerj.6563> PMID: 30881765
28. Carruthers T, Muñoz-Rodríguez P, Wood JRI, Scotland RW. The temporal dynamics of evolutionary diversification in *Ipomoea*. *Mol Phylogenet Evol*. 2020; 146: 106768. <https://doi.org/10.1016/j.ympev.2020.106768> PMID: 32081764
29. Xiao S, Xu P, Deng Y, Dai X, Zhao L, Heider B, et al. Comparative analysis of chloroplast genomes of cultivars and wild species of sweetpotato (*Ipomoea batatas* [L.] Lam). *BMC Genomics*. 2021; 22: 262. <https://doi.org/10.1186/s12864-021-07544-y> PMID: 33849443
30. Mota NFO, Watanabe MTC, Zappi DC, Hiura AL, Pallos J, Viveros RS, et al. Amazon canga: the unique vegetation of Carajás revealed by the list of seed plants. *Rodriguésia*. 2018; 69: 1435–1488.
31. Simão-Bianchini R, Vasconcelos LV, Pastore M. Flora das cangas da Serra dos Carajás, Pará, Brasil: Convolvulaceae. *Rodriguésia*. 2016; 67: 1301–1318.
32. Austin DF, Secco RS. *Ipomoea marabaensis*, nova Convolvulaceae da Serra dos Carajás (PA). *Bol. Mus. Para. Emílio Goeldi. Série Botânica*. 1988; 4: 187–194.
33. APG IV. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc*. 2016; 181: 1–20.
34. Ravi V, Khurana JP, Tyagi AK, Khurana P. An update on chloroplast genomes. *Plant Syst Evol*. 2008; 271: 101–122.
35. Tonti-Filippini J, Nevill PG, Dixon K, Small I. What can we do with 1000 plastid genomes? *Plant J*. 2017; 90: 808–818. <https://doi.org/10.1111/tj.13491> PMID: 28112435
36. Foster CSP, Henwood MJ, Ho SYW. Plastome sequences and exploration of tree-space help to resolve the phylogeny of riceflowers (Thymelaeaceae: *Pimelea*). *Mol Phylogenet Evol*. 2018; 127: 156–167. <https://doi.org/10.1016/j.ympev.2018.05.018> PMID: 29803950

37. Jiang Y, Yang Y, Lu Z, Wan D, Ren G. Interspecific delimitation and relationships among four *Ostrya* species based on plastomes. *BMC Genet.* 2019; 20: 33. <https://doi.org/10.1186/s12863-019-0733-0> PMID: 30866795
38. Cho M-S, Kim JH, Yamada T, Maki M, Kim SC. Plastome characterization and comparative analyses of wild crabapples (*Malus baccata* and *M. toringo*): insights into intraspecific plastome variation and phylogenetic relationships. *Tree Genet Genomes.* 2021; 175: 41.
39. Schwarz EN, Ruhlman TA, Weng ML, Khiyami MA, Sabir JSM, Hajarrah NH, et al. Plastome-wide nucleotide substitution rates reveal accelerated rates in Papilionoideae and correlations with genome features across legume subfamilies. *J Mol Evol.* 2017; 84: 187–203. <https://doi.org/10.1007/s00239-017-9792-x> PMID: 28397003
40. D'Agostino N, Tamburino R, Cantarella C, De Carluccio V, Sannino L, Cozzolino S, et al. The complete plastome sequences of eleven *Capsicum* genotypes: insights into DNA variation and molecular evolution. *Genes.* 2018; 9: 503. <https://doi.org/10.3390/genes9100503> PMID: 30336638
41. Fan W-B, Wu Y, Yang J, Shahzad K, Li Z-H. Comparative chloroplast genomics of Dipsacales species: insights into sequence variation, adaptive evolution, and phylogenetic relationships. *Front Plant Sci.* 2018; 9: 689. <https://doi.org/10.3389/fpls.2018.00689> PMID: 29875791
42. Gonçalves DJP, Simpson BB, Ortiz EM, Shimizu GH, Jansen RK. Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. *Mol Phylogenet Evol.* 2019; 138: 219–232. <https://doi.org/10.1016/j.ympev.2019.05.022> PMID: 31146023
43. Gonçalves DJP, Jansen RK, Ruhlman TA, Mandel JR. Under the rug: abandoning persistent misconceptions that obfuscate organelle evolution. *Mol Phylogenet Evol.* 2020; 151: 106903. <https://doi.org/10.1016/j.ympev.2020.106903> PMID: 32628998
44. Williams AM, Friso G, van Wijk KJ, Sloan DB. Extreme variation in rates of evolution in the plastid Clp protease complex. *Plant J.* 2019; 98: 243–259. <https://doi.org/10.1111/tpj.14208> PMID: 30570818
45. Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A. Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *Am J Bot.* 2012; 99: 349–364. <https://doi.org/10.3732/ajb.1100335> PMID: 22174336
46. Rousseau-Gueutin M, Bellot S, Martin GE, Boutte J, Chelaifa H, Lima O, et al. The chloroplast genome of the hexaploid *Spartina maritima* (Poaceae, Chloridoideae): comparative analyses and molecular dating. *Mol Phylogenet Evol.* 2015; 93: 5–16. <https://doi.org/10.1016/j.ympev.2015.06.013> PMID: 26182838
47. Williams AV, Miller JT, Small I, Nevill PG, Boykin LM. Integration of complete chloroplast genome sequences with small amplicon datasets improves phylogenetic resolution in *Acacia*. *Mol Phylogenet Evol.* 2016; 96: 1–8. <https://doi.org/10.1016/j.ympev.2015.11.021> PMID: 26702955
48. Xu J, Shen X, Liao B, Xu J, Hou D. Comparing and phylogenetic analysis chloroplast genome of three *Achyranthes* species. *Sci Rep.* 2020; 10: 10818. <https://doi.org/10.1038/s41598-020-67679-y> PMID: 32616875
49. Sun Y, Moore MJ, Zhang S, Soltis PS, Soltis DE, Zhao T, et al. Phylogenomic and structural analyses of 18 complete plastomes across nearly all families of early-diverging eudicots, including an angiosperm-wide analysis of IR gene content evolution. *Mol Phylogenet Evol.* 2016; 96: 93–101. <https://doi.org/10.1016/j.ympev.2015.12.006> PMID: 26724406
50. Zhang R, Wang YH, Jin J-J, Stull GW, Bruneau A, Cardoso D, et al. Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. *Syst Biol.* 2020; 69: 613–622. <https://doi.org/10.1093/sysbio/syaa013> PMID: 32065640
51. Vasconcelos S, Nunes GL, Dias MC, Lorena J, Oliveira RRM, Lima TGL, et al. Unraveling the plant diversity of the Amazonian canga through DNA barcoding. *Ecol Evol.* 2021; 11: 13348–13362. <https://doi.org/10.1002/ece3.8057> PMID: 34646474
52. Rogstad SH. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. *Taxon.* 1992; 41: 701–708.
53. Andrews S. FastQC: A quality control tool for high throughput sequence data. 2010 [cited 2021 Jul 5]. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
54. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014; 30: 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
55. Dierckxsens N, Mardulyn P, Smits G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 2017; 45: e18. <https://doi.org/10.1093/nar/gkw955> PMID: 28204566
56. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012; 19: 455–477. <https://doi.org/10.1089/cmb.2012.0021> PMID: 22506599

57. Hoshino A, Jayakumar V, Nitasaka E, Toyoda A, Noguchi H, Itoh T, et al. Genome sequence and analysis of the Japanese morning glory *Ipomoea nil*. *Nat Commun*. 2016; 7: 13295. <https://doi.org/10.1038/ncomms13295> PMID: 27824041
58. Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, et al. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics*. 2012; 13: 715. <https://doi.org/10.1186/1471-2164-13-715> PMID: 23256920
59. McKain MR, Johnson MG, Uribe-Convers S, Eaton D, Yang Y. Practical considerations for plant phylogenomics. *Appl Plant Sci*. 2018; 6: e1038. <https://doi.org/10.1002/aps3.1038> PMID: 29732268
60. Sakamoto W, Takami T. Chloroplast DNA dynamics: copy number, quality control and degradation. *Plant Cell Physiol*. 2018; 59: 1120–1127. <https://doi.org/10.1093/pcp/pcy084> PMID: 29860378
61. Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA. Database indexing for production MegaBLAST searches. *Bioinformatics*. 2008; 24: 1757–1764. <https://doi.org/10.1093/bioinformatics/btn322> PMID: 18567917
62. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinformatics*. 2000; 16: 944–945. <https://doi.org/10.1093/bioinformatics/16.10.944> PMID: 11120685
63. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9: 357–359. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286
64. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. Version 4.0.9 [software]. 2020 Jan 20 [cited 2021 Jul 5]. <http://www.repeatmasker.org>.
65. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res*. 2019; 47: W59–W64. <https://doi.org/10.1093/nar/gkz238> PMID: 30949694
66. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002; 30: 3059–3066. <https://doi.org/10.1093/nar/gkf436> PMID: 12136088
67. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014; 30: 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033> PMID: 24451623
68. Fan H, Ives AR, Surget-Groba Y, Cannon CH. An assembly and alignment-free method of phylogeny reconstruction from next-generation sequencing data. *BMC Genomics*. 2015; 16: 522. <https://doi.org/10.1186/s12864-015-1647-5> PMID: 26169061
69. Logacheva MD, Schelkunov MI, Shtratnikova VY, Matveeva MV, Penin AA. Comparative analysis of plastid genomes of non-photosynthetic Ericaceae and their photosynthetic relatives. *Sci Rep*. 2016; 6: 30042. <https://doi.org/10.1038/srep30042> PMID: 27452401
70. Machado LO, Vieira LN, Stefenon VM, Faoro H, Pedrosa FO, Guerra MP, et al. Molecular relationships of *Campomanesia xanthocarpa* within Myrtaceae based on the complete plastome sequence and on the plastid *ycf2* gene. *Genet Mol Biol*. 2020; 43: e20180377. <https://doi.org/10.1590/1678-4685-GMB-2018-0377> PMID: 32555941
71. Yan M, Xiong Y, Liu R, Deng M, Song J. The application and limitation of universal chloroplast markers in discriminating East Asian evergreen oaks. *Front Plant Sci*. 2018; 9: 569. <https://doi.org/10.3389/fpls.2018.00569> PMID: 29868047
72. Dong W, Liu J, Yu J, Wang L, Zhou S. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One*. 2012; 7: e35071. <https://doi.org/10.1371/journal.pone.0035071> PMID: 22511980
73. Sebastiani F, Carnevale S, Vendramin GG. A new set of mono- and dinucleotide chloroplast microsatellites in Fagaceae. *Mol Ecol Notes*. 2004; 4: 259–261.
74. Shepherd LD, Lange PJ, Cox S, McLenachan PA, Roskrige NR, Lockhart PJ. Evidence of a strong domestication bottleneck in the recently cultivated New Zealand endemic root crop, *Arthropodium cirratum* (Asparagaceae). *PLoS One*. 2016; 11: e0152455. <https://doi.org/10.1371/journal.pone.0152455> PMID: 27011209
75. McDonald A. Revision of *Ipomoea* section *Exogonium* (Choisy) Griseb (Convolvulaceae). *Brenesia*. 1987; 8: 41–87.
76. Manos PS, Miller RE, Wilkin P. Phylogenetic analysis of *Ipomoea*, *Argyrea*, *Stictocardia*, and *Turbina* suggests a generalized model of morphological evolution in morning glories. *Syst Bot*. 2001; 26: 585–602.
77. Linder CR, Rieseberg LH. Reconstructing patterns of reticulate evolution in plants. *Am J Bot*. 2004; 91: 1700–1708.

78. Bernhardt N, Brassac J, Dong X, Willing EM, Poskar CH, Kilian B, et al. Genome-wide sequence information reveals recurrent hybridization among diploid wheat wild relatives. *Plant J.* 2020; 102, 493–506. <https://doi.org/10.1111/tpj.14641> PMID: 31821649
79. Heath TA, Hedtke SM, Hillis DM. Taxon sampling and the accuracy of phylogenetic analyses. *Artic J Syst Evol.* 2008; 46: 239–257.
80. Qin Y, Li M, Cao Y, Gao Y, Zhang W. Molecular thresholds of ITS2 and their implications for molecular evolution and species identification in seed plants. *Sci Rep.* 2017; 7: 17316. <https://doi.org/10.1038/s41598-017-17695-2> PMID: 29229945
81. Miller RE, Rausher MD, Manos PS. Phylogenetic systematics of *Ipomoea* (Convolvulaceae) based on ITS and Waxy sequences. *Syst Bot.* 1999; 24: 209–227.
82. Austin DF, Huaman Z. A synopsis of *Ipomoea* (Convolvulaceae) in the Americas. *Taxon.* 1996; 45: 3–38.
83. McDonald JA, Mabry TJ. Phylogenetic systematics of New World *Ipomoea* (Convolvulaceae) based on chloroplast DNA restriction site variation. *Plant Syst Evol.* 1992; 180: 243–259.
84. Miller RE, Buckley TR, Manos PS. An examination of the monophyly of morning glory taxa using Bayesian phylogenetic inference. *Syst Biol.* 2002; 51: 740–753. <https://doi.org/10.1080/10635150290102401> PMID: 12396588
85. Stefanović S, Austin DF, Olmstead RG. Classification of Convolvulaceae: a phylogenetic approach. *Syst Bot.* 2009; 28: 791–806.
86. Silva AR, Moreira LCR, Carvalho S, Lanes ECM, Ortiz-Vera MP, Viana PL, et al. Range-wide neutral and adaptive genetic structure of an endemic herb from Amazonian Savannas. *AoB Plants.* 2020; 12: plaa003. <https://doi.org/10.1093/aobpla/plaa003> PMID: 32128104
87. Dalapicolla J, Alves R, Jaffé R, Vasconcelos S, Pires ES, Nunes GL, et al. Conservation implications of genetic structure in the narrowest endemic quillwort from the Eastern Amazon. *Ecol Evol.* 2021; 11: 10119–10132. <https://doi.org/10.1002/ece3.7812> PMID: 34367563
88. Pereira JBS, Giulietti AM, Pires ES, Laux M, Watanabe MTC, Oliveira RRM, et al. Chloroplast genomes of key species shed light on the evolution of the ancient genus *Isoetes*. *J Syst Evol.* 2021; 59: 429–441.
89. Pereira JBS, Giulietti AM, Prado J, Vasconcelos S, Watanabe MTC, Pinangé DSB, et al. Plastome-based phylogenomics elucidate relationships in rare *Isoetes* species groups from the Neotropics. *Mol Phylogenet Evol.* 2021; 161: 107177. <https://doi.org/10.1016/j.ympev.2021.107177> PMID: 33866010