BMC Ecology and Evolution

RESEARCH

Open Access

Plastome phylogenomics of the diverse neotropical orchid genus *Lepanthes* with emphasis on subgenus *Marsipanthes* (*Pleurothallidinae*: *Orchidaceae*)

Tatiana Arias[1,8*], Juan Sebastian Moreno[2], Sebastian Reyes[3], Martin Llano Almario[4], Alejandra Serna-Sánchez[5], Gabriel A. Iturralde[6], Janice Valencia[7], Luis Baquero[6] and Alejandro Zuluaga[4]

Abstract

Well-resolved phylogenetic relationships within the diverse Neotropical orchid genus *Lepanthes* are presented based on a genome skimming approach that yielded nine newly sequenced chloroplast genomes. We complemented this with 17–86 plastome coding genes for 26 species retrieved from GenBank, alongside amplified *matK* and *rITS* regions. The *Lepanthes* plastomes (157,185—158,260 bp, 37.15% GC content) contained 136 annotated genes, including 86 protein-coding, 42 tRNA, and 8 rRNA genes. We identified six hypervariable regions, including parts of the *ycf1* gene, as potential DNA barcodes. Phylogenetic analyses revealed that Carl Luer's subgeneric classifications are non-monophyletic, a finding confirmed by PCA of continuous morphological traits, reflecting significant morphological homoplasy. Six major clades were identified, though resolution for the phylogenetic backbone remains unresolved at two nodes. Subgenus *Marsipanthes* is not monophyletic as currently circumscribed, with two subclades recovered in distinct positions within the phylogeny. An early-diverging lineage, comprising species restricted to the eastern Andean slopes from southern Colombia to Peru, includes members of both *Marsipanthes* and *Lepanthes*. A derived clade, consisting of species from both subgenera, confined to the Chocó biogeographic region, forms an unresolved polytomy. Although only a subset of *Lepanthes* diversity was sampled, this study captures significant taxonomic, geographic, and morphological variation. It provides foundational insights into the genu's evolutionary history, along with tools and hypotheses that can be expanded upon in future research to further refine our understanding of its biogeographic history.

Keywords Andes, Chloroplast, Colombia, Eastern Andes, Chocó Biogeographic Region, Ecuador, *Ycf1*

*Correspondence:
Tatiana Arias
tatianaarias@orchidsforpeace.org
Full list of author information is available at the end of the article

Introduction

Lepanthes Sw. (Orchidaceae) stands out as one of the largest genera in the subtribe Pleurothallidinae [1], with around 1164 accepted species broadly distributed across the Neotropics, from Mexico to Bolivia and the Antilles, and from the Andes to the Amazon Basin [2–4] (Fig. 1). The interplay between their specialized habitat requirements and the complex environmental gradients of the Andean region and the Talamanca Mountain range in Central America has made *Lepanthes* one of the most species-rich and endemic-rich genera in the Neotropics [5].

The genus displays its highest diversity in Colombia and Ecuador, with 377 species recently reported for Colombia [10] and more than 300 species for Ecuador [23]. In this region, species span an impressive altitudinal gradient, from sea-level to nearly 4000 m in elevation near the Andean snow line [2, 6]. They thrive in diverse habitats, ranging from lowland areas along the Pacific Ocean and the Amazon foothills to the cloud forest of Sierra Nevada de Santa Marta and the Andes. *Lepanthes* predominantly inhabit ecosystems characterized by high humidity, constant air movement, and horizontal precipitation. They grow attached to litter on the forest floor or on various phorophytes, including shrubs and trees, from the main trunk to branches and the canopy. This remarkable ecological flexibility, combined with the intersection of three major biodiversity hotspots—Mesoamérica, Chocó/Darién/Western Ecuador, and the tropical Andes [11, 12]—and the narrow geographic distributions of many species [5], are proposed to be responsible for the extraordinary diversity observed in the genus.

Lepanthes are characterized by their diminutive size, and stems called ramicauls, which are enclosed in funnel-shaped sheaths known as lepanthiform sheaths. These sheaths often display microscopic cilia or pubescence, and variable apical dilation. The flowers typically feature transversely elongated, often bilobed petals and a compound lip comprised of a body, connectives, and two blades. The lip plays a crucial role in pseudocopulatory pollination [13]. Notably, these orchids possess some of the smallest flowers in the orchid family and within the genus, various morphological groups exhibit specific similarities [14–16]. *Lepanthes* are notoriously difficult to classify due to the subtle differences between species. Comprehensive molecular studies combined with thorough morphological documentation will provide effective species delimitation and a clearer species concept for this genus.

Taxonomy in *Lepanthes* was primarily developed by Carlyle Luer and collaborators using morphological

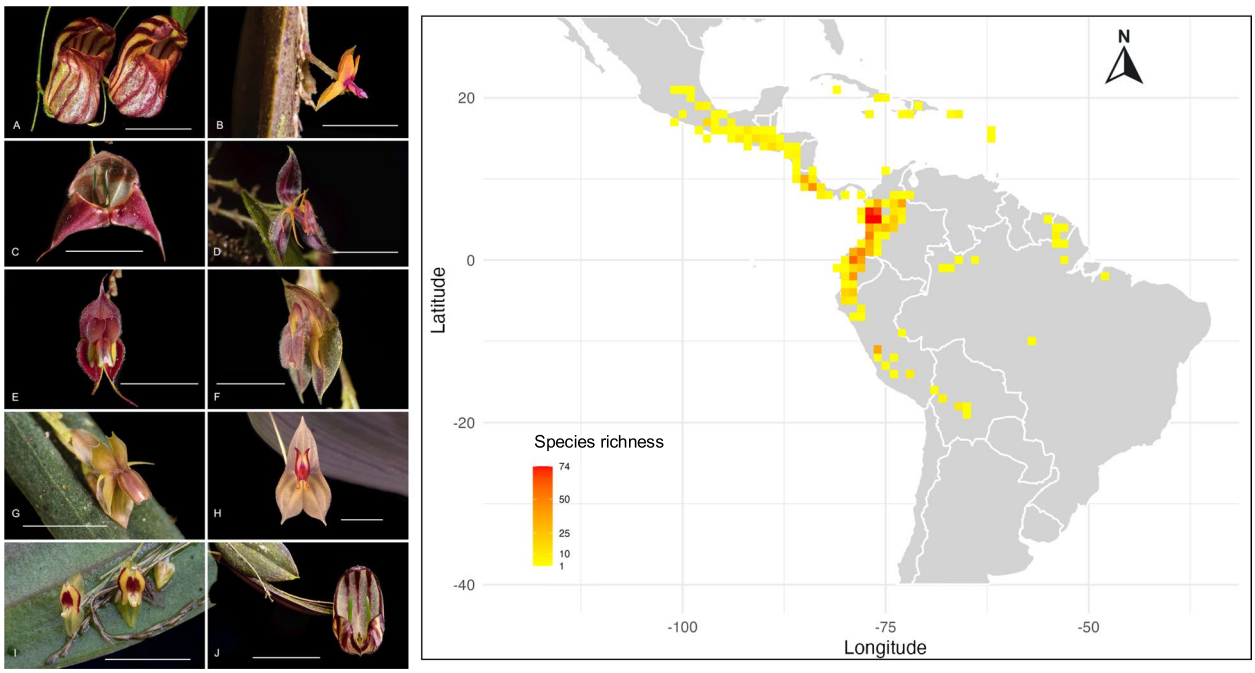

Fig. 1 Taxonomic representatives and distribution of *Lepanthes*. Left panel: Representatives of each taxonomic category proposed by Luer [6–8] and their morphological diversity. **A**. *L. caprimulgus* **B**. *L. eros* **C**. *L. felis* **D**. *L. hexapus* **E**. *L. manabina* **F**. *L. monoptera* **G**. *L. mucronata* **H**. *L. nicolasii* **I**. *L. narcissus* **J**. *L. ribes*. Photographs by Ron Parsons (A), J. S. Moreno (B, C, D, E, F and G), SCO (H), J. Riksen (I) and S. Vieira-Uribe (J). Scale bar = 5 mm. Right Panel: Spatial distribution of *Lepanthes* species richness across the Neotropics. Species richness was calculated within a 1-degree grid based on occurrence of 4582 records obtained from GBIF, iNaturalist, Species link and Moreno [9].

traits [7, 17–20]. Only in recent years, molecular studies have included *Lepanthes* species, mostly from Central American and the Caribbean [21, 22]. Despite hundreds of species distributed throughout diverse ecosystems of the Neotropics, Luer proposed the entire diversity of *Lepanthes* could be confined within two subgenera: subg. *Lepanthes* and subg. *Marsipanthes* [23] (Fig. 1). Subgenera *Lepanthes* encompasses the majority of species within the genus and is further divided into three subsections, as defined by Luer [7, 17–19]: subsect. *Breves* (including series *Breves* and *Filamentosa*), distinguished by lateral sepals with a single vein; subsect. *Lepanthes* (comprising series *Lepanthes*, *Elongatae*, and *Mucronatae*), characterized by lateral sepals with two or three veins; and subsect. *Bilabiatae*, in which the lateral sepals are fully fused into a two-veined synsepal. However, the concept of series within these subsections was later abandoned [2].

The subg. *Marsipanthes* comprises eleven species, distinguished by connate dorsal and lateral sepals with numerous veins forming an inflated, saccate calyx, and thick, fleshy, erect petals with one lobe [24]. Within *Marsipanthes*, three sections are recognized: species of sect. *Marsipanthes* have narrowly elongated petals; while those of sect. *Caprimulginae* have short, transverse petals, and deeply connate sepals; and species of sect. *Felinae* have non-inflated flowers and asymmetrical petals with the upper lobe much longer than the lower one [6]. Sect. *Caprimulginae* includes *Lepanthes caprimulgus* Luer and a recently described species *L. attenboroughii* Baquero & Monteros [25]. Sect. *Marsipanthes* includes *L. ribes* Luer and *L. portillae* Luer; and sect. *Felinae* includes *L. carunculigera* Rchb.f., *L. equicalceolata* Luer & Escobar, *L. felis* Luer & Escobar, *L. lucifer* Luer & Hirtz, *L. niessienae* Luer and *L. quadricornis* Luer & Escobar [6, 23]. *L. tulcanensis* Baquero & Yeager was described in recent years, and based on morphological traits was considered to belong to subg. *Marsipanthes* sect. *Felinae* [26].

The current taxonomy of *Lepanthes* requires reevaluation within an evolutionary framework. Previous studies using few DNA molecular markers have failed to resolve relationships [22], limiting our understanding of the genu's origin and evolution. Robust phylogenetic analyses based on tens of thousands of nucleotides, can greatly enhance our confidence in the resulting evolutionary hypotheses [27–29]. Plastomes have been widely used in orchid phylogenetics due to their high conservatism and relatively slow evolutionary rates [30]. However, it remains questionable whether large-scale plastome datasets, emcompassing both coding and non-coding regions, possess enough phylogenetic signal to resolve relationships in rapidly diversifying Andean orchids like *Lepanthes* [31].

This study employed high-throughput sequencing (HTS) and genome skimming to re-examine infrageneric taxonomic classifications and determine their consistency with morphological traits. By analyzing the plastome, we aimed for higher levels of sequence divergence to resolve phylogenetic relationships within the genus, beyond what traditional Sanger sequencing allows. We present complete plastome sequences of nine representative *Lepanthes* species to: 1) describe and interpret the plastome structure and evolution within main *Lepanthes* taxonomic groups; 2) reconstruct phylogenetic relationships among selected *Lepanthes* species using plastid genome sequences, with a focus on subgenus *Marsipanthes*; and 3) test preliminary hypothesis on the evolution of continuous morphological characters for *Lepanthes*. This paper initiates a broader project aimed at unraveling the genu's phylogenetic and phylogeographic history.

Methods

Taxon sampling, DNA extraction, amplification and sequencing

Fresh leaves from adult plants were collected in Colombia from Colomborquídeas in El Retiro, Antioquia, under a collection permit from the Corporación para Investigaciones Biológicas (resolution ANLA 1263) and CITES (permit number: 42664). We sequenced ten new plastomes representing valid *Lepanthes* species, using the circumscription and infrageneric classification of Luer [7, 8, 19] (Fig. 1). Additionally, 26 taxa from Ecuador were sampled for this study, 24 species of *Lepanthes* and two outgroups *Pseudolepanthes colombiae* Archila and *Brachionidium valerioi* Ames & C. Schweinf (Table S4) under a collection permit for Ecuador (MAATE-DBI-CM-2021-0187). Herbarium voucher was prepared including a dry specimen accompany by spirit flowers preserved in Copenhagen solution [78]. All vouchers were deposited at QCNE (Table S4). For all species, fresh leaves were stored in silica gel for later DNA extraction using the CTAB method for species sampled in Colombia [32]. Total DNA was purified with silica columns (Epoch Life Science Inc. Missouri, Texas, USA) and eluted in Tris-EDTA. DNA concentration was measured using a fluorometer (Qubit Fluorometer, Invitrogen, USA), and the integrity and purity of the samples were verified through agarose gel electrophoresis. DNA from species sampled in Ecuador was extracted using a rapid extraction procedure [33].

For plastome sequencing, a 300-bp DNA library was constructed using the TruSeq Illumina platform (Illumina Inc., San Diego, CA, USA). A total of 1.5 µg of DNA was fragmented using a Covaris ultrasonicator (Covaris Inc., Woburn, MA), and the fragmented DNA was assessed via gel electrophoresis (Bio-Rad Laboratories, Hercules, CA).

The fragmented DNA was treated with End Repair Mix (New England BioLabs, Ipswich, MA) and incubated at 20 °C for 30 min. The end-repaired DNA was purified with the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany), then treated with A-Tailing Mix (Thermo Fisher Scientific, Waltham, MA) and incubated at 37 °C for 30 min. The adenylated DNA was ligated with adapters using the Adapter and Ligation Mix, with the ligation reaction incubated at 20 °C for 15 min. Adapter-ligated DNA was selected by running a 2% agarose gel to recover the target fragments, followed by purification with the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). Multiple rounds of PCR amplification were performed using a PCR Primer Cocktail and PCR Master Mix to enrich the adapter-ligated DNA fragments. The PCR products were again selected via 2% agarose gel and purified with the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). The final library was quantified by determining the average molecule length with an Agilent 2100 Bioanalyzer (Agilent DNA 1000 Reagents, Santa Clara, CA) and by real-time quantitative PCR (qPCR) using TaqMan Probes (Thermo Fisher Scientific, Waltham, MA). Libraries were multiplexed and sequenced on an Illumina HiSeq 4000 platform (San Diego, CA), producing paired end reads of 150 base pairs in length at the Beijing Genomics Institute (BGI), Hong Kong.

Polymerase Chain Reaction (PCR) was used for amplification of two regions, nuclear ribosomal internal transcribed spacer (*rITS*) and plastid maturase K (*matK*), with 7.5 µL GoTaq Green Master Mix 2X (Promega), 3 µL of extracted DNA, 7.5 µL ultra-pure water, and 1 µL of each primer (Table S1). PCR conditions were as follows: 2 min–95 °C, 1 min–95 °C, 1 min–55 °C, 1 min–72 °C, 35 cycles, and 5 min–72 °C for final extension. PCR products were purified, and Sanger sequenced (ABI 3500xL Genetic Analyzer, Applied Biosystem).

Plastome assembly, annotation, and comparative analyses

On average 4.5 GB of high-quality data was obtained for each sample. To evaluate the quality of the reads and check for the presence of adapters or other contaminants, we used FastQC [34]. Based on this analysis, the reads were trimmed using Trimmomatic v0.39 [35], removing reads with a Phred score below 33 (parameter "-phred33") or a length shorter than 100 bp (MINLEN:100). Adapters within the reads (parameter "-ILLUMINACLIP") and low-quality bases at the start (LEADING) and end (TRAILING) of the reads were also removed. The remaining paired-end sequences were merged with a maximum overlap of 100 bp in Geneious Prime 2021 v2.2.

The quality-filtered reads were then assembled de novo using GetOrganelle v1.7.5 [36], specifying chloroplasts as the target organelle (parameter "-F embplant_pt"). Gaps between scaffolds were closed through iterative mapping in Geneious Prime 2021 v2.2 using default settings. Out of the ten sequenced plastomes, only nine were successfully assembled. The sequencing reads for *Lepanthes eros* Luer & R. Escobar were contaminated, and we were unable to assemble its plastome with confidence. The plastome sequences were deposited in GenBank (Bioproject ID PRJNA1209138) (Table S4).

We annotated the assembled plastomes using the Geneious annotation tool with default parameters and a 95% similarity threshold. The genomes of *Anathallis obovata* (Lindl.) Pridgeon and M.W. Chase [37] and *Lepanthes caprimulgus* Luer [30] were used as reference. Annotations were manually verified to ensure no stop codons within genes. Circular plastome maps were created using OGDRAW v2 [38]. Several open reading frames (ORFs) that did not match the initial annotations were identified using BLASTn [39] followed by BLASTx [40] v2.15.0 searches in NCBI.

Plastome structure and sequence divergence analyses

To assess potential expansions and contractions of the IR boundaries, the genes in the boundary regions of LSC/IRb/SSC/IRa were visualized using IRscope v3.1 [41]. Additionally, gene arrangement was analyzed through a collinearity analysis using the Mauve v1.1.3 plugin in Geneious v9.1.4 with default parameters. Sequence divergence of *Lepanthes* plastomes was examined using mVISTA [79], with the *Anathallis obovata* plastome serving as a reference. Nucleotide diversity (Pi) values were calculated using DnaSP v6 with a sliding window of 1000 sites and a step size of 25 sites [42].

Plastome repetitive sequence and codon usage analyses

Three types of repeat sequences were analyzed in the plastomes, simple sequence repeats (SSRs), tandem repeats, and long repeats. SSRs were detected using MISA v2.1 [43] and visualized with the R package ggplot2. Tandem repeats were identified using Tandem Repeats Finder v4.09 [43]. Forward, reverse, complement, and palindromic long repeats were detected in REPuter [44] using default parameters. The Relative Synonymous Codon Usage (RSCU) ratio was estimated using CodonW v1.4.2 [45]. A RSCU value greater than 1 indicates positive codon usage bias, while a value less than 1 indicates less frequent usage. The R package "pheatmap" was used to generate a heatmap for the RSCU analysis.

Phylogenetic analyses

Nine *Lepanthes* species along with *Anathallis obovata*, *Draconanthes aberrans* (Schltr.) Luer, *Trichosalpinx dirhamphis* (Luer) Luer and *Zootrophion lappaceum* Luer

and R. Escobar as outgroups, were included in the phylogenetic analyses. Plastomes were aligned using MAFFT v7.526 [46]. To reconstruct the phylogenetic relationships within *Lepanthes*, Maximum Likelihood (ML) analyses were conducted using RAxML v7.4.2 [47] under the general time-reversible (GTR) substitution model. Bootstrap support was assessed with 1000 bootstrap replicates. Additionally, Bayesian Inference (BI) analysis were performed using MrBayes v3.2.7a with 20 million generations sampled every 1000 generations. The majority rule (> 75%) consensus tree was obtained after removing the first 25% of the sampled trees as "burn-in" [48].

Whole plastid phylogenies were reconstructed based on the following datasets: (1) complete plastid genomes, (2) different outgroup sets, (3) one inverted repeat (IR), (4) coding sequences, and (5) non-coding sequences. The *Filamentosae* series (subg. *Lepanthes*, sect. *Lepanthes*, subsect. *Breves*) and the *Elongatae* series (subg. *Lepanthes*, sect. *Lepanthes*, subsect. *Lepanthes*) were not represented in this phylogeny by any species.

In addition to constructing complete plastid phylogenies, we downloaded 26 SRA files for *Lepanthes* from GenBank (Table S4) and mined all plastid sequences. Chloroplast assemblies were generated using GetOrganelle v1.7.7.1 [36] and Captu v1.0.0 [49]. To perform alignments and phylogenetic reconstructions, we employed two approaches: (1) coding sequences concatenation with an incomplete data matrix for a subset of species, and (2) a *matK* phylogeny for all available species.

For the phylogenetic reconstruction of a combined dataset (*rITS* + *matK*) focusing on subgenus *Marsipanthes*, analyses were conducted using *Brachionidium valerioi* Ames & C. Schweinf. and *Pseudolepanthes colombiae* Archila as outgroup species. ML analyses were performed using the PhyML plugin [50] in Geneious Prime, employing the GTR + G substitution model with 1000 bootstrap replicates. BI analyses were conducted in BEAST v1.10.4 [51], with the GTR + G substitution model and the Yule speciation model. The BI analysis ran for 10 million generations, with tree sampling occurring every 10,000 generations. A single consensus tree was generated using TreeAnnotator v1.10.4.

Morphological analysis

Patterns of morphological variation and potential correlations between traits and phylogenetic clades were assessed using Principal Component Analysis (PCA) and correlation matrices. Protologues of species included here were consulted to build a database including continuous measurements of vegetative and floral structures such as the ramicauls, leaves, peduncles, sepals, petals, lip, and column. Both minimum and maximum dimensions were considered for each continuous morphological character. The analysis involved several R packages, including *ggplot2* v3.4.4 for data visualization, and *ggcorrplot* v0.1.4 for generating a correlation matrix plot. We pre-processed the data to address missing values using the *na.omit* function and converted categorical variables to numeric format. After filtering species with missing data, only 20 species were used for the analysis. Normalization of the data was performed using z-score scaling to ensure comparability among variables.

Results

Plastomes molecular descriptions

We obtained 7,150,658 (*L. eros*) to 11,373,812 bp (*L. narcissus*) paired reads after cleaning, with an average read length of 145 bp and a recovery rate of approximately 82% of the original reads (Table S2). The nine *Lepanthes* plastomes assembled had an average length of 157,710 bp, with *L. narcissus* Luer & Hirtz having the smallest plastome (157,185 bp) and *L. caprimulgus* Luer the largest (158,260 bp) (Table 1). The G/C content of all plastomes was approximately 37%. Each plastome exhibited the typical quadripartite structure seen in angiosperm plastomes, including a Large Single Copy (LSC), a Small Single Copy (SSC), and two Inverted Repeat (IR) regions (Fig. 2). *L. ribes* Luer had the smallest LSC (89,500 bp) while *L. narcissus* had the largest (90,156 bp). In terms of SSC, *L. nicolasii* Luer & R.Escobar had the smallest at 19,095 bp, while *L. ribes* had the largest (23,401 bp). The IR regions ranged from 24,495 bp in *L. nicolasii* to 21,346 bp in *L. ribes*. All plastomes contained 136 unique genes, including 42 tRNA genes, 8 rRNA genes, and 86 protein-coding genes. Additionally, 25 genes were duplicated in the IR regions (Table 1, Fig. 2).

Several Open Reading Frames (ORFs) were identified across the *Lepanthe*s plastomes. In the LSC region, two hypothetical proteins were detected via BLASTx. One was found within *trnC-GCA* in all species except *L. nicolasii* and *L. caprimulgus* (408 bp, 136 aa). The second one partially overlapped *clpP1* and was exclusively found in the *L. manabina* Dodson plastome, showing 97.96% identity to a *Clp* protease subunit in *Epidendrum ciliare* L. (e-value: 1e-21).

In both inverted repeat regions (IRa and IRb), a hypothetical protein, *M5 K25*, previously identified in *Dendrobium thyrsiflorum* B.S.Williams, was detected within *trnI-GAU* in *Lepanthes ribes* Luer *L. narcissus*, and *L. caprimulgus*. Furthermore, another hypothetical protein, conserved across all *Lepanthes* plastomes, was located within the *rrn23* rRNA gene (408 bp, 136 amino acids). This protein exhibited high similarity to chloroplast

Table 1 List of the *Lepanthes* species used in this study and their infrageneric taxonomic classification following Luer [6–8]**.** The series *Filamentosae* (subg. *Lepanthes*, sect. *Lepanthes*, subsect. *Breves*) is not represented in this phylogeny by any species (NA not included here). **Lepanthes eros* was not included in further analysis due to sequence contamination

Subgenera	Section	Subsection	Series abandoned by Luer and Thoerle [2]	Total species number	Species sequenced here	Species sequenced elsewhere
Marsipanthes	*Caprimulginae*	-		2	*L. caprimulgus* *L. attenboroughii*	
	Felinae	-		6	*L. felis* *L. lucifer* *L. niesseniae* *L. tulcanensis*	*L. lucifer* *L. niesseniae*
	Marsipanthes	-		3	*L. ribes*	
Lepanthes	*Lepanthes*	*Bilabiatae*		7	*L. narsissus*	
		Breves	*Breves*	85	*L. monoptera* *L. eros**	*L. elata* *L. tachirensis*
			Filamentosae	13	NA	NA
		Lepanthes	*Mucronatae*	60	*L. mucronata* *L. hexapus*	*L. ankistra* *L. heteroloba* *L. pelix*
			Elongatae	154	NA	*L. acarina* *L. clowesii* *L. katleri* *L. martineae* *L. nycteris* *L. terborchii*
			Lepanthes	795	*L. barbelifera* *L. calodyction* *L. manabina* *L. nicolasii*	*L. adrianae* *L. auriculata* *L. ballatrix* *L. calodyction* *L. ciliisepala* *L. clareae* *L. confusa* *L. crucitasensis* *L. dactylopetala* *L. dodsonii* *L. gargantua* *L. grandiflora* *L. imitator* *L. juninensis* *L. matamorosii* *L. meniscophora* *L. mystax* *L. orion* *L. saltatrix* *L. tentaculata* *L. terpsichore* *L. turialvae* *L. urbaniana* *L. volador*

proteins from *Acorus calamus* L. (95.56%, e-value: 2e-86), *Dendrobium chrysanthum* Rchb.f. (93.04%), and *D. thyrsiflorum* (95.6%).

The *ycf1* gene, located around the IRa-SSC boundary, exhibits variable copy numbers and is found in the SSC in most species, except for *L. mucronata* and *L. nicolasii.* In these species, one copy of the gene is located in the IRa and the second in the SSC (Fig. S1). *L. felis* Luer & R.Escobar, *L. narcissus* and *L. manabina* contained two functional copies, where as *L. caprimulgus* has three. In *L. manabina*, an additional ORFs was been found embedded within the overlap of the two *ycf1* copies. This ORF showed similarity to a *Dracula erythrocochaete* Luer *NADH dehydrogenase subunit F* (94.51%, e-value: 1e-90) and a *Masdevallia picturata* Rchb.f. *NADH plastoquinone oxidoreductase subunit 5* (94.51%, e-value:

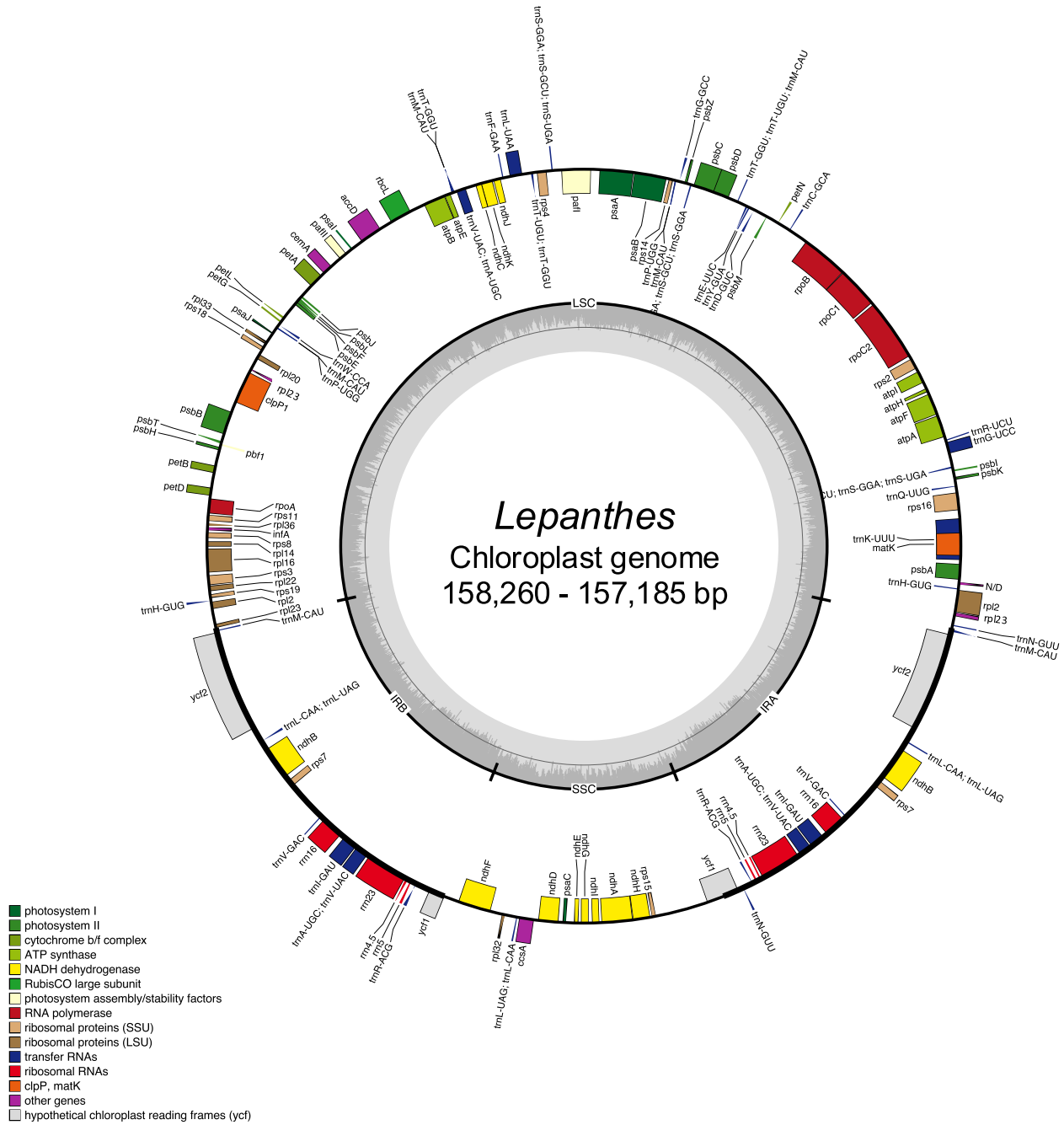

Fig. 2 Plastome structure of nine *Lepanthes* species. The darker gray in the inner circle corresponds to the GC content. Bars of different colors indicate different functional groups. Genes on the inside circle are transcribed clockwise, while genes on the outside circle are transcribed counterclockwise. LSC: large single copy region; SSC: small single copy region IR: inverted repeat regions. *Lepanthes nicolasii* chloroplast

2e-90). *L. narcissus* had two copies of the *ndhF* gene, one of them overlapping with *ycf1* and *ndhF*. Only two species have multiple copies of *ycf1* in the boundary of SSC-IRb, *L. felis* Luer & R.Escobar (3 copies) and *L. nicolasii* (2 copies) (Fig. S1).

Plastome structural variations

An IR boundary map was generated by comparing the *Lepanthes* plastomes using IRscope (Fig. 3). The *rpl23* gene was consistently positioned at the end of LSC in the JLB junction. At the IRb-SSC junction (JSB), *ycf1* was

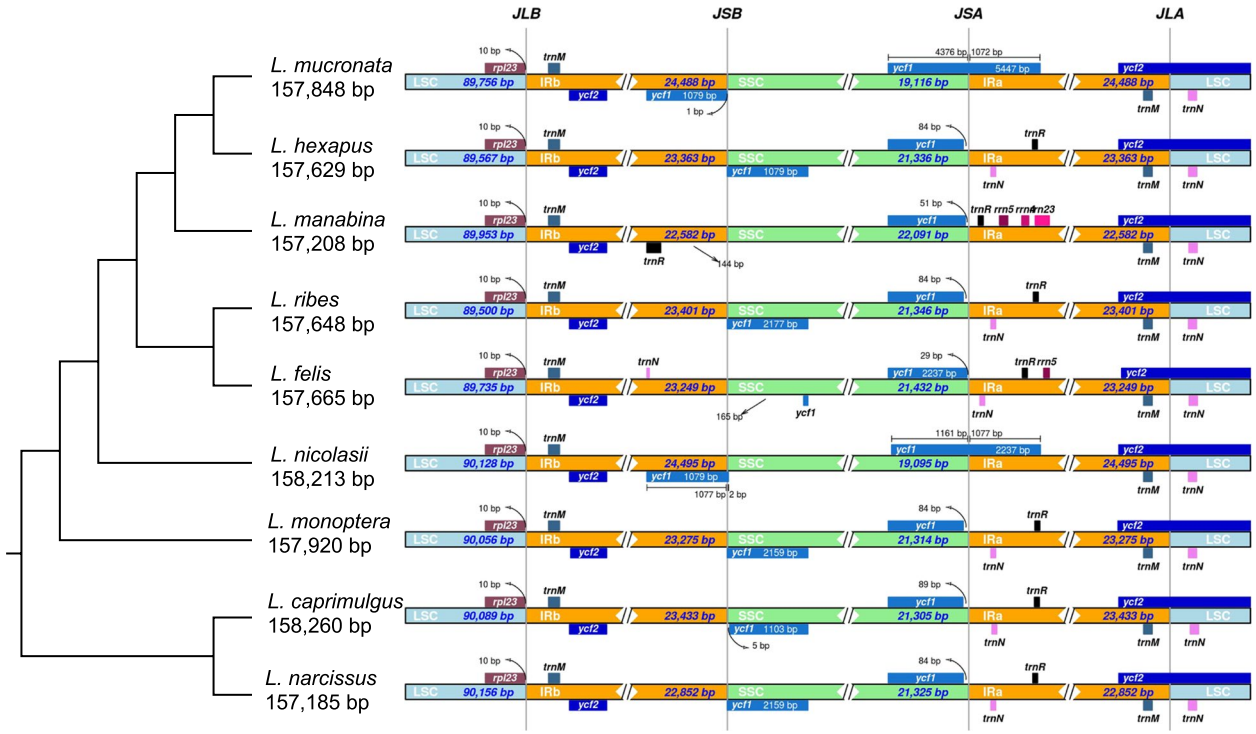

Fig. 3 Comparison of the boundaries between the LSC, SSC and IR regions in the nine *Lepanthes* plastomes. Tree topology from the ML phylogeny of complete plastomes. JLB: LSC/IRb junctions; JSB: SSC/IRb junctions; JSA: SSC/IRa junctions; JLA: LSC/IRa junctions

located at the end of the IRb in *L. mucronata* and *L. nicolasii.* However, in the remaining species, it was typically located at the start of the SSC.

At the SSC-IRa junction (JSA), *ycf1* was entirely located within the SSC, positioned from 29-89 bp before the boundary in most species. However, in *L. mucronata* and *L. nicolasii, ycf1* spanned 1072 and 1077 bp, respectively, extending from the SSC into IRa. At the IRa-LSC junction (JLA), *ycf2* was found spanning both regions.

Collinearity analysis showed no gene rearrangements or inversions in the *Lepanthes* plastomes (Fig. S2).

Identification of hypervariable regions

Whole-plastomes alignment revealed that sequence variation in conserved non-coding regions (highlighted in pink bars) was greater than in protein-coding regions (highlighted in purple bars) (Fig. S3). The variation rates in both coding and non-coding regions within the two IR regions were lower than those in the LSC and SSC regions. Highly divergent non-coding regions included *psbk-psbI*, *atpB-rbcL*, *petA-psbJ*, *psbB-psbN*, and *rpl32-trnL*. In contrast, the rRNA genes were highly conserved compared to other genes.

Nucleotide diversity (Pi) values were calculated using DnaSP (Fig. 4). The average Pi value across the nine plastomes was 0.00906, with a total of 4054 mutations identified. The IR region averaged a Pi value of 0.00229, the LSC region averaged 0.01050, and the SSC region averaged 0.01286. Based on Pi values (ranging from 0.02 to 0.03), seven hypervariable regions were identified: *trnKuuu–rps16*, *psbK–psbI*, *psbM–trnEucc*, *trnTUGU–trnLUAA*, *accD–psaI*, *clpP1–psbB*, and *ycf1*.

Repeated sequence analysis

The number of identified SSRs ranged from 58 in *L. hexapus* Luer & R.Escobar to 47 in *L. monoptera* Lindl., with a total of 470 SSRs detected across the *Lepanthes* plastomes. Three types of SSRs (mono-, di-, and complex-nucleotide repeats) were identified, with 436 (92.9%) being mono-nucleotide repeats, predominantly composed of A and T motifs. Additionally, 13 di-nucleotide repeats (2.77%) and 21 complex-nucleotide repeats (4.47%) were identified (Table 3).

A total of 215 tandem repeats were detected, with counts ranging from 17 in *L. nicolasii* to 38 in *L. caprimulgus*. Notably, all plastomes contained 50 long repeats each, comprising palindromic, forward, reverse, or complementary long repeats. Palindromic repeats were the most abundant, ranging from 33 in the *L. manabina* plastome to 23 in *L. narcissus*. Forward repeats were also abundant, with *L. felis*, *L. monoptera*, and *L. caprimulgus* each containing 18 forward repeats, while *L.*

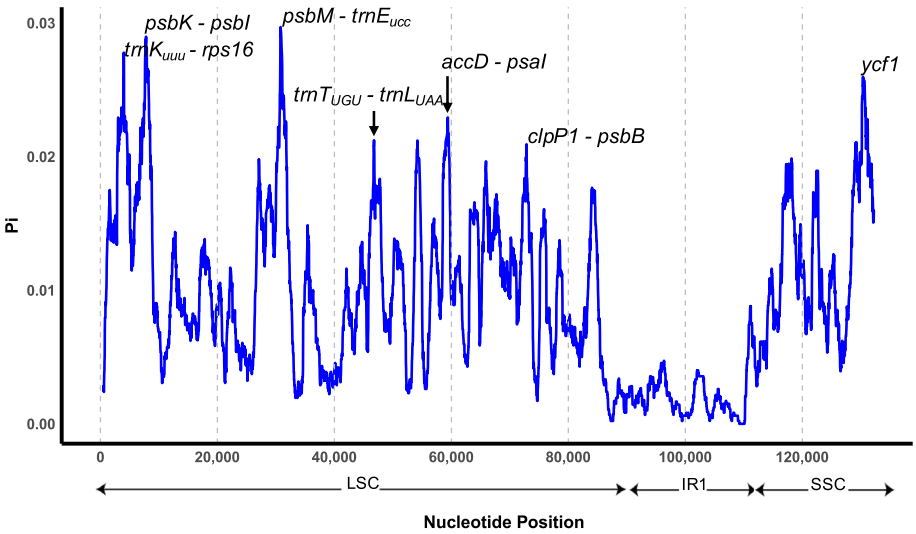

Fig. 4 Nucleotide variability (Pi) of whole plastomes from nine *Lepanthes* species including a single Inverted Repeat

hexapus contained 11. Reverse and complementary repeats accounted for only 59 out of 450 repeats, representing 13.1% of all long repeats (Table S3).

To assess codon usage bias, the relative synonymous codon usage (RSCU) ratios were calculated for *Lepanthes* plastomes. Less frequently used codons (RSCU < 1) were generally consistent across species, with most having 32, except *L. hexapus* and *L. narcissus*, which had 33. Most preferred codons ended with A or U, except for UGG and AGG. The codons AGA and UCU exhibited the highest RSCU values, averaging 2.11 and 1.58, respectively, while CGC and GAC had the lowest RSCU values, with averages of 0.41 and 0.54 (Fig. S4).

Phylogenetic relationships

Topologies based on whole plastomes, coding regions, and intergenic regions were identical, with similar resolutions using both BI and ML methods. Species of *Lepanthes* formed a well-supported monophyletic group (PP = 1.00, BS = 100). Within *Lepanthes*, two sister clades were identified: *L. narcissus* + *L. caprimulgus* (PP = 1.00, BS = 100) and the remaining *Lepanthes* species (PP = 1.00, BS = 100). In the latter clade, *L. monoptera* was recovered as sister to the rest, followed by *L. nicolasii*, which is sister to a clade (PP = 1.00, BS = 100) comprising *L. manabina* sister to *L. mucronata* + *L. hexapus*, and *L. ribes* + *L. felis* in the Bl phylogeny and not resolved in the ML one (Fig. 3).

For the phylogenies combining our newly generated data with sequences mined from GenBank (Fig. 5), the mean alignment sequence length was 37,725 bp, ranging from a maximum of 72,282 bp to a minimum of 7042 bp. The alignment spanned a total length of 85,750 bp, exhibiting a pairwise identity of 46.6%, and 5.3% identical sites. Half of the species had between 50 and 77 chloroplast genes present in the alignment (Fig. 6). We found all taxonomic groups below the genus level to be non-monophyletic, except for series *Mucronatae*. However, the monophyly of *Mucronatae* requires further confirmation as the phylogeny did not fully resolved relationships among species included here (Figs. 3 and 5, Fig. S5).

Six main clades were recovered. An early-divergent and monospecific lineage, comprising *L. tachirensis*, was sister to the remaining species. Clade 1, sister to the rest of *Lepanthes* included *L. narcissus*, as sister to *L. caprimulgus* + *L. katleri*. Clade 2 consists of *L. monoptera*, *L. rhynchion*, and *L. speciosa*, sister to the remaining *Lepanthes* species. A polytomy was observed, including *L. acarina* and *L. clowesii* remaining outside major clades, Clade 3 comprising *L. gargantua* as sister to subclades *L. imitator*, *L. clarae*, *L. nicolasii* + *L. adrianae*, *L. orion* and *L. auriculata*, as well as Clade 4. Clade 4 included Central America and the Caribbean species *L. turialvae*, *L. urbaniana*, and *L. grandiflora*, sister to the remaining species. Clade 5 included *L. tentaculate* and *L. calodyction*, and a polytomy (designated Clade 6) encompassing several smaller subclades: *L. felis* + *L. ribes*, *L. hexapus* + *L. mucronata*, *L. lucifer* + *L. niesseniae*, *L. heteroloba* + *L. pelyx*, and a subclade of *L. dodsonii*, *L. manabina*, *L. terpsichore* and *L. meniscophora* (Fig. 5).

Further BI and ML analyses, focusing on nine of the eleven *Marsipanthes* species and using a combined *rITS* + *matK* dataset, recovered several of the clades identified in our main phylogeny, confirming *Marsipanthes* is non-monophyletic (Figures S6, S7). Main clades to highlight from this phylogeny include: (1) Eastern Andes

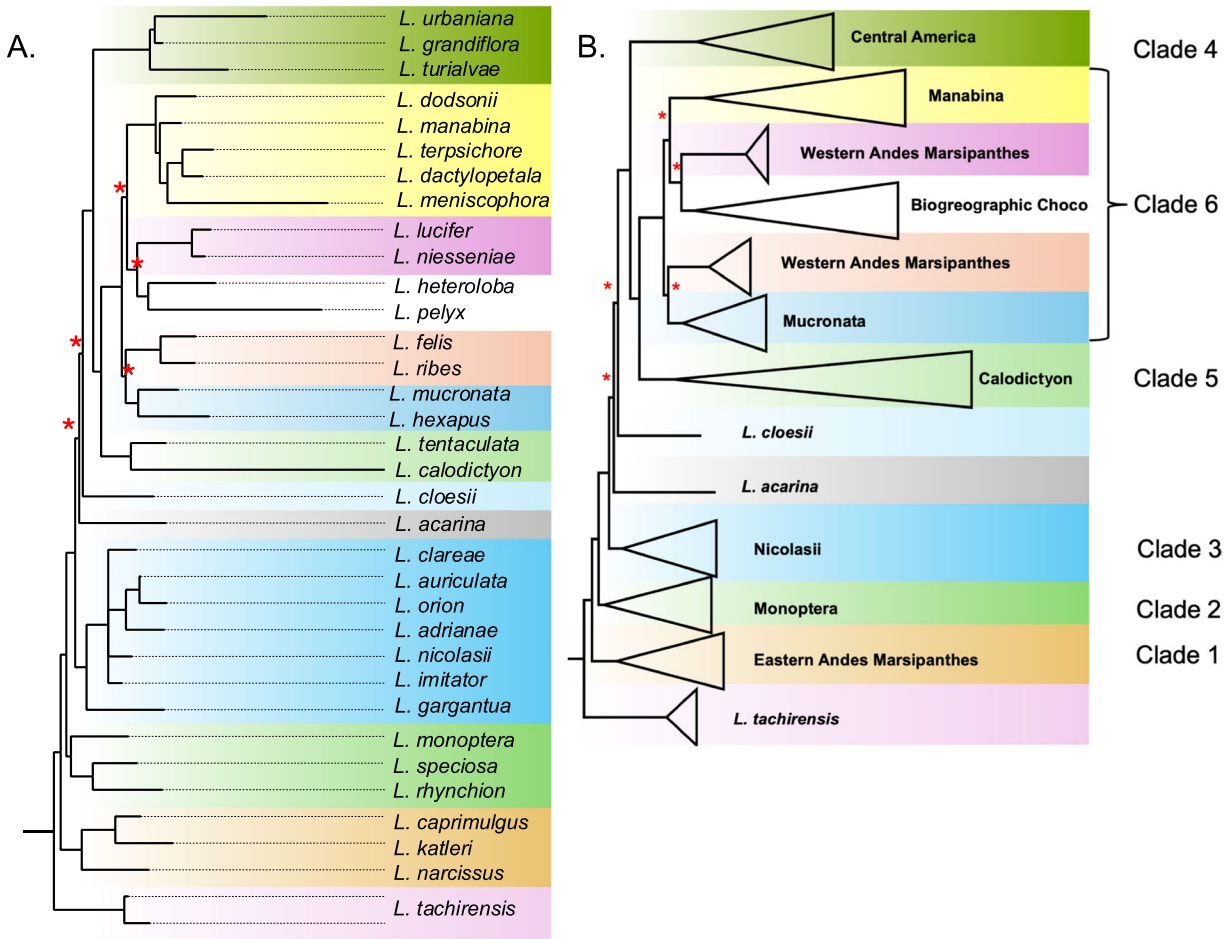

Fig. 5 Inferred phylogenies for the *Lepanthes* backbone using Bayesian Inference based on 17-86 coding plastid regions. **A**. Molecular phylogeny of 35 *Lepanthes* species and two outgroups. **B**. Molecular phylogeny of Lepanthes highlighting the main clades and subclades. Colors represent equivalent clades and subclades in both phylogenies. Red asterisk shows support value below 0.8 PP

Clade (Ecuador, Perú and Bolivia): This clade contains *L. caprimulgus, L. attenboroughii* (both representing subgen. *Marsipanthes*, sect. *Caprimulginae*) along with *L. martinae, L. nycteris* and *L. terbochii,* which belong to subgen. *Lepanthes* (BS = 75, PP = 1). (2) Western Andean *Marsipanthes* clade: This clade includes *L. felis*, *L. ribes*, *L. lucifer* and *L. niesseniae* (BS = 94, PP = 1). The phylogenetic position of *Lepanthes tulcanesis* remains uncertain. Additional clades recovered in this phylogeny include *L. calodictyon* and its closely related species *L. barbellifera, L. saltatrix* and *L. volador* (BS = 100, PP = 1).

Lastly, a *matK* phylogeny for 104 *Lepanthes* species, including both sequences produced here and those available in GenBank, lacked resolution (Fig. S8).

Variation in morphological continuous characters

PCA analysis indicated that continuous morphological measurements did not form distinct groups according to clades (Figs. 7A, B). PC1 (35.49%) and PC2 (26.47%) account for about 61.96% of the total variance in the dataset. Dorsal sepals (Length Min.-Max., Width Min.-Max.) and Lateral sepals (Length Min.-Max., Width Min.-Max.) had the highest positive loading in this PCA suggesting sepal dimensions are critical for distinguishing among the species in our dataset. Traits like Peduncle Length Min. and Pedicel Length Min. also contribute positively. Traits with negative loadings on PC1 include Ramicauls Length Min. and Number of sheaths Min., indicating they might show an inverse relationship with the traits positively contributing to PC1. Traits with positive loadings in PC2 included Dorsal sepal Length Min., Lateral sepal Length Min., and Peduncle Length Min. Ramicauls Length Min., Leaf Width Min., Lip Length Min., and Column Length Max. show negative contributions to the data distribution (Fig. S9).

Fig. 6 Plastid coding genes representation in *Lepanthes* and outgroups. **A**. Number of plastid coding genes per species included in the alignment. **B**. Overall completeness of the alignment for each gene across species. Yellow color indicates the absence of a coding gene. **C**. Percentage coverage of each gene in the total number of species within the alignment

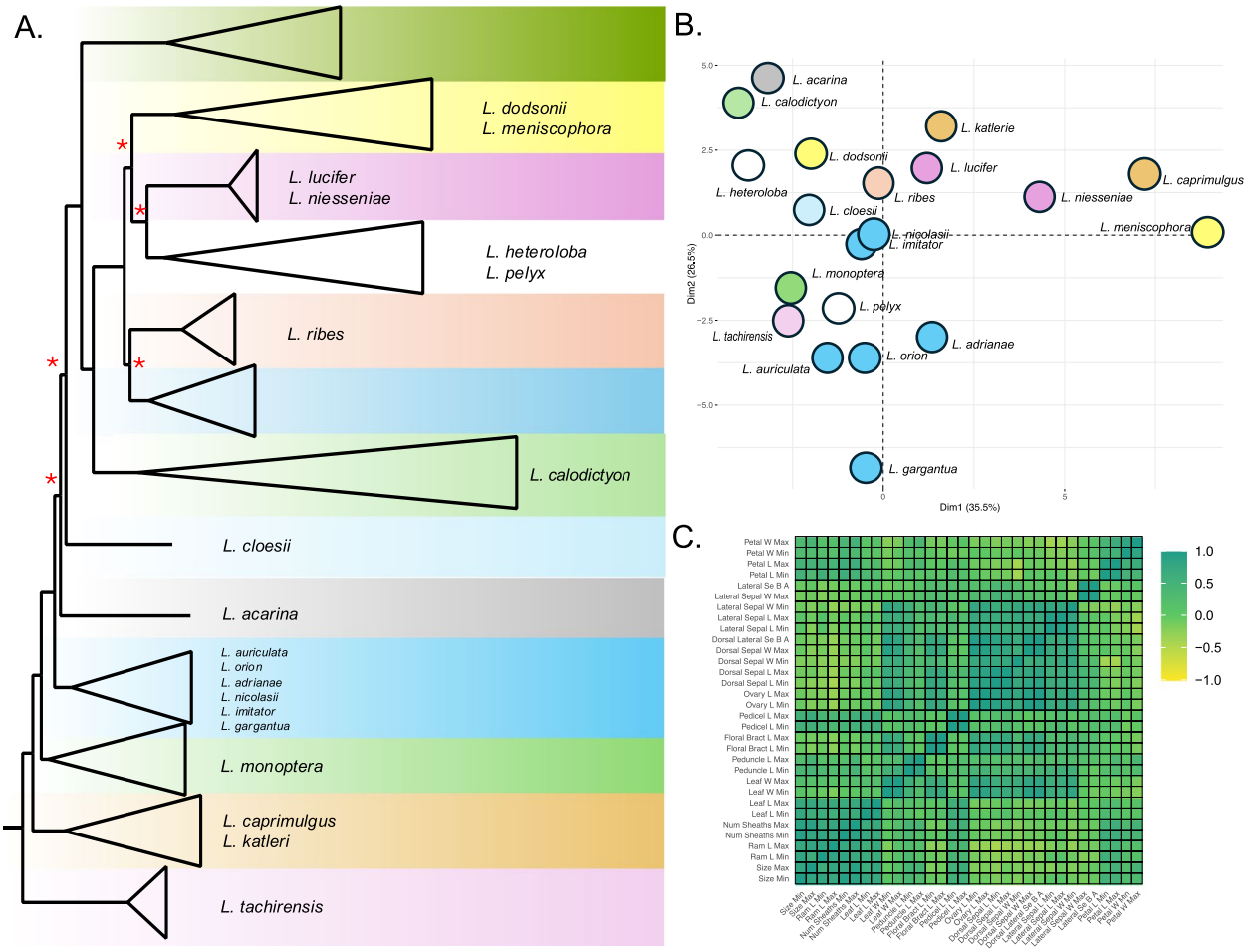

Fig. 7 Morphological analysis of continuous vegetative and reproductive characters in *Lepanthes*. **A**. Recovered phylogenetic relationships and main *Lepanthes* clades, this panel highlights species included in the PCA, with colors representing different clades. **B**. PCA, colored dots represents the clades where the species belong in the phylogeny. **C**. Heat map of a correlation matrix showing either positive or negative correlations among all continuous morphological characters studied here

Continuous morphological characters showed high positive correlations among several floral and vegetative traits (Fig. 7C). Strong correlations (above 0.8) were found between Peduncle Length and Floral Bract Length (0.930). Leaf Length and Ovary Length (0.888), Ovary Length and Floral Bract Length (0.849), Dorsal Sepal Length and Dorsal Sepal Width (0.849). None of the negative correlations were strong, for example Dorsal Sepal Width and Leaf Width were negatively correlated (−0.244) (Fig. 7C).

Discussion

Our genome skimming approach produced nine new chloroplast genomes for *Lepanthes* species, providing valuable molecular resources for the scientific community. By also incorporating plastome genes from GenBank, we present the first moderately resolved phylogenetic reconstruction for the genus. Our findings demonstrate that the current taxonomic groups proposed by Carl Luer are not monophyletic, indicating significant morphological homoplasy within *Lepanthes*. We identified six main clades; however, the moderate resolution of the phylogenetic backbone underscores the necessity of incorporating nuclear markers and a more extensive species sampling to further refine these relationships. Despite the limited species sampling, this study encompasses substantial taxonomic, geographic, and morphological diversity, thereby establishing a strong foundation for future evolutionary research on *Lepanthes*.

Plastome evolution of *Lepanthes* species

A range of plastome lengths, averaging 157,710 base pairs (bp) was observed in *Lepanthes* (Table 2). The G/C content across all plastomes was consistent at approximately 37%, reflecting typical characteristics for angiosperm

Table 2 Summary of complete chloroplast genomes statistics for *Lepanthes* species. Comparison of plastome subunit lengths and GC content for the assembled plastomes

Species	Length (bp)	LSC (bp)	SSC (bp)	IR (bp)	GC%
L. mucronata	157,848	89,756	19,116	24,488	37.1
L. hexapus	157,629	89,567	21,363	23,363	37.1
L. manabina	157,208	89,953	22,091	22,582	37.1
L. ribes	157,648	89,500	23,401	21,346	37.2
L. felis	157,665	89,735	21,432	23,249	37.2
L. nicolasii	158,213	90,128	19,095	24,495	37.1
L. monoptera	157,920	90,056	21,314	23,275	37.1
L. caprimulgus	158,260	90,089	21,305	23,433	37.2
L. narcissus	157,185	90,156	21,325	22,852	37.1

Table 3 Repetitive motif abundance in nine. *Lepanthes* species computed using MISA. Distribution and types of simple sequence repeats

	Mononucleotides	Dinucleotides	Complex repeats	SSRs Total
L. caprimulgus	50	1	2	53
L. narcissus	47	1	3	51
L. monoptera	45	1	2	47
L. nicolasii	47	1	3	51
L. ribes	46	1	3	50
L. felis	47	1	2	50
L. manabina	48	3	2	53
L. mucronata	52	2	2	56
L. hexapus	54	2	2	58

and orchid plastomes [52]. The plastome size observed in this study aligns with the broad range reported for Orchidaceae, spanning from the reduced 19,047 bp in the mycoheterotrophic *Epipogium roseum* [53] to the largest to our knowledge 212,688 bp in the autotrophic *Cypripedium subtropicum* [54], reflecting the family's diversity in photosynthetic function and evolutionary adaptation.

Lepanthes species exhibit consistency in plastome length and structure that aligns with the general characteristics of monocot plastomes [52, 55–57]. This stability is important for epiphytes often growing in nutrient-poor soils, as maintaining a stable chloroplast genome may be crucial to efficiently capture and use light and other limited resources [58]. Despite the generally conserved nature of chloroplast genomes, the presence of unique open reading frames (ORFs) in *Lepanthes* plastomes suggests possible lineage-specific evolutionary processes to be further investigated in the future [52, 59]. For example, our analysis identified an ORF within *trnC-GCA* in all species except *L. nicolasii* and *L. caprimulgus* (408 bp, 136 amino acids), indicating potential losses of this ORF at least twice in the genus evolution. Furthermore, an ORF in *L. manabina*, partially overlapping with the *clpP1* gene, demonstrated strong similarity to a *clp* protease subunit in *Epidendrum ciliare*, suggesting either horizontal gene transfer (HGT) or convergent evolution.

The *ycf1* and *ndhF* genes exhibit considerable molecular diversity and multiple copies across the *Lepanthes* species sequenced here (Figs. 4, S1). *L. felis* and *L. nicolasii* have five and four *ycf1* copies respectively, while *L. narcissus* has two *ndhF* copies (Fig. S1). This copy number variation does not seem to have a clear phylogenetic pattern indicating multiple and independent gene gains and losses. Gene duplications and losses pertaining to these regions have been reported in other angiosperms and orchids [60–62]. The *ycf1* has been showed to encode products that are essential for cell survival [60]. Multiple copies of this gene observed here may perform slightly different functions.

An examination of the IR boundaries and gene positions reveals slight patterns and variation among different species (Figs. 3, S1). Chloroplast region boundaries are not only fundamental to the structural integrity and function of the chloroplast genome but also provide insights into evolutionary processes, functional genomics, and adaptation strategies. The positions of the LSC-IR junctions vary slightly within groups, but usually this has only negligible effects on plastome size [52]. SSC-IR boundaries were not fixed in *Lepanthes*; instead, and like in other plants, they undergo a dynamic and stochastic process that primarily results in conservative expansions and contractions of the IR [55]. Several consistencies were found across all sequenced *Lepanthes* species: the gene *rpl23* is positioned at the end of the LSC, the IR contains a *trnH-rps19* cluster as reported in most monocots [56], and an overlap between the genes *ycf1* and *ndhF* occurs near or at the IRa-SSC boundary. In contrast, there was some variation at the IR and the SSC junctions: the position of *ycf1* copies varied among species, both in the SSC and IRb (JSB) or IRa (JSA) (Fig. 3). For example, the position of *ycf1* in *L. mucronata* and *L. nicolasii* expands from the SSC into the IRa (JSA). The IR can terminate within genes, leading to nonfunctional and potentially disruptive 5' or 3' truncated genes, whose translated products may negatively impact essential processes [55]. In Pleurothallidinae, the IR/SSC boundary interrupts the *ycf1* gene, resulting in a truncated *ycf1* fragment within the IR. Conversely, in *Cattleya crispata*, the SSC encompasses nearly the entire *ycf1* gene [37]. These plastome differences between Pleurothallidinae and Laeliinae may be characteristic of these subtribes.

Collinearity analysis further support the stability of these plastomes, as no gene rearrangements or inversions were detected (Fig. S2). This lack of major structural

changes reinforces the notion that while there are some variations in gene positioning, the overall plastome architecture of *Lepanthes* remains relatively stable across species.

Last, the analysis of relative synonymous codon usage (RSCU) revealed variability in codon preferences across the *Lepanthes* plastomes. The presence of both preferred (RSCU >1) and non-preferred codons (RSCU =1) highlights possible selective pressures influencing plastome evolution. Codons ending in A or U are prevalent, with AGA and UCU exhibiting the highest RSCU values (Fig. S4). This pattern is consistent with the trend observed in other plant plastomes, where specific codon biases can reflect functional constraints and evolutionary pressures [63].

New molecular resources for *Lepanthes*

The identification of hypervariable regions in the plastomes of *Lepanthes* species reveals significant insights into the genetic variability across different regions of the plastome. We observed that sequence divergence was notably higher in the conserved non-coding regions compared to the protein-coding regions. Specifically, non-coding regions such as *pskB-psbI, atpB-rbcL, petA-psbJ, psbB-psbN*, and *rpl32-trnL* exhibited higher variability, while rRNA genes remained highly conserved. The analysis of nucleotide diversity (Pi) values further highlights the distribution of genetic variation across different plastome regions. The IR regions showed the lowest nucleotide diversity (Pi= 0.00229), suggesting relative stability in these regions. In contrast, the LSC and SSC regions exhibited higher diversity, (Pi values of 0.01050 and 0.01286, respectively). This indicates these regions are more prone to variation, which could be useful for distinguishing between species or for evolutionary studies.

Parts of the *ycf1* gene are identified as potential good barcode regions for the genus, its utility has also been confirmed for many other angiosperms [61]. With Pi values between 0.02 and 0.03, these regions offer the highest degree of variability, making them effective markers for species identification and phylogenetic studies (Figs. 4, S3). In addition to *ycf1*, our study specifically identified numerous SSRs within the chloroplast genomes that can be utilized to discover genome polymorphisms, enabling detailed population genetics analyses both within and between the species investigated in this study [64].

Plastome phylogeny challenges traditional classifications and revealed new lineages in *Lepanthes*

Although our phylogenies have limited taxonomic representation (9–35/1164 spp.) they include a large number of informative molecular characters (≤15,834bp) and capture species representing unique morphological groups identified in previous taxonomic studies [65].

All phylogenies reconstructed here confirmed the monophyly of *Lepanthes* as proposed by Bogarín *et al* [22] (Fig. 5). However, both *Lepanthes* subgenera recognized by Carl Luer [19] were found to be non-monophyletic (Figs. 5, S5-7). This indicates that perianth characters are homoplastic, and the fusion of petals and number of veins are not informative at this taxonomic level.

Additionally, most of the subclades recovered do not support the classifications below subgenus level proposed Luer [6]. An exception, which requires confirmation in further studies is the subclade comprising *L. hexapus* + *L. mucronata.* This subclade was recovered as monophyletic in the whole chloroplast phylogeny (Fig. 3), and belongs to series *Mucronatae* from the sect. *Lepanthes*. However, the monophyly of series *Mucronatae* was not confirmed in the 35-tip phylogeny, largely because of the lack of resolution at that scale (Figs. 5 and 6).

The series *Filamentosae* (subg. *Lepanthes*, sect. *Lepanthes*, subsect. *Breves*) is not represented in this phylogeny by any species. Those species will remain to be included in future studies to be properly placed in a systematic context for the genus.

Lepanthes tachirensis emerges as the sister lineage to all other *Lepanthes* species, representing, an isolated, monotypic lineage. Luer's taxonomic classification initially place it within subg. *Lepanthes*, subsect. *Breves* [8]. Originally described from Venezuela's Tachira region, the species is now recognized across northern South America, including Colombia and Ecuador. Morphologically, *L. tachirensis* is distinguished by its triangular-shaped open sepals and a tightly sheathed ramicaul, with the ostia, closely adhering to the ramicaul, unlike the funnel-like shape seen in other *Lepanthes*. Given the current phylogenetic sampling, this monotypic status remains provisional.

Our phylogenetic analysis indicates that a GenBank sequence attributed to *Lepanthes juninensis* is actually *L. tachirensis*. Accordingly, we have treated this sample as *L. tachirensis* in our study. The misidentification of *L. juninensis*, originally described from Peru, appears to have been historically propagated among collectors and commercial nurseries, despite its minimal morphological similarity to *L. tachirensis*. This confusion likely stems from the limited availability of diagnostic information on *L. juninensis* and the challenges in interpreting its original description and illustration [66]. Resolving this taxonomic uncertainty will require targeted fieldwork in the Peruvian type locality to confirm the true identity of *L. juninensis.*

An early divergent clade, the Eastern Andean Marsipanthes clade (Clade 1), is identified in our phylogenies as sister to the rest of *Lepanthes*. This clade comprises

two species from the subg. *Marsipanthes* sect. *Caprimulginae* (*L. attenboroughii* and *L. caprimulgus)*, both native to the Eastern Andes of Ecuador and Perú. These species are more closely related to members of *Lepanthes* sect. *Elongatae* (*L. martineae*, *L. nycteris*, *L. terbochii*, and *L. katleri*) and sect. *Bilabiatae* (*L. narcissus*)— also distributed across the Eastern Andes of Ecuador, Perú, and Bolivia— than to other *Marsipanthes* (Figs. 6, S5–7, 10). Most species in this clade, except for *L. narcissus*, have shortly pubescent, multi-veined (5–7 at the dorsal sepal and 2–4 in each lateral sepal) sepals, that are erose to lacerate, adaxially carinate, and long pubescent to fringed at the margins, deeply to shallowly concave, and stripped. Their petals are frequently lunate, with upper and lower lobes of similar shape and size. The lip has long, densely pubescence, narrow to very narrowly triangular lobes that surround the long column (Figs. 1 and S10). Based on their floral morphology and distribution in the Eastern Andes, molecular studies should corroborate whether species from subg. *Lepanthes* series *Elongatae* such as *L. echinata* Luer & Cloes, *L. mulleriana* Luer, *L. tigrina* Luer & Thoerle, in addition to *L. portillae* (sect. *Marsipanthes*) discovered in El Condor Mountain range of Ecuador, belong to Clade 1. The monophyly of the remaining species in sect. *Bilabiatae* (approximately seven species) and their status as the sister group to Clade 1, as suggested by our phylogeny with *L. narcissus*, remains to be tested.

The Monoptera clade (Clade 2), is resolved as sister to the remainder of *Lepanthes.* This clade comprises species from subsect. *Breves*, which Dodson and Luer [23] and Luer and Thoerle [20] distinguished by their single-veined lateral sepals. These high-altitude species typically occur above 2500 m and form two distinct groups: (1) the *L. monoptera* group is characterized by a bilaminate lip with oblong to ovate, long-ciliate blades, rounded blade bases, narrowly obtuse apices, short connectives, and a distinctive ovary with a single wing or carina. Other species that could be part of this group include *L. cornualis*, *L. cyrtostele*, *L. ferax*, and *L. jucas*. (2) the *Rhynchion/Speciosa* group, comprising species with large stature, large leaves, and prominent, adnate, or flattened lepanthiform sheaths. Intriguingly, *L. tachirensis* also belongs to subsect. *Breves,* and shares morphological traits with *L. rhynchion* and *L. speciosa*, yet it is phylogenetically placed at the base of the phylogeny. This highlights the non-monophyly of subsect. *Breves*.

Two main polytomies are present in this phylogeny, indicating that the molecular data used were insufficient to fully resolve parts of the backbone. Future studies will requiere additional molecular markers or more extensive species sampling to improve resolution, similar to how increased data led to a better-resolved *Cypripedium* phylogeny [67]. These polytomies might also reflect recent divergence or rapid speciation events. The first polytomy contains: *L. cloesii*, *L. acarina*, the Nicolasii clade (clade 3) (comprising *L. nicolassi*, *L. gargantua*, *L. clareae*, *L. adrianae*, *L. auriculata*, and *L. orion),* and a clade containing the remaining *Lepanthes* species sampled here (Fig. 5).

Lepanthes cloesii and *L. acarina* did not form a clade, despite their previous classification in the artificial series *Elongatae*. This finding supports Luer and Thoerle's [2] decision to abandon this grouping. In the *ITS/matK* phylogeny, *L. cloesii* forms a clade with *L. elata* and *L. ballatrix* (Figs. S6–7). Other species with morphological similarities to *L. cloesii*, such as *L. elegantula*, *L. pastoensis*, and *L. alexandroi*, share features like thick, fleshy leaves and relatively large flowers compared to their smaller leaves. Within this phylogeny, *L. cloesii* is the sole representative of a larger group that requires further investigation. In contrast, *L. acarina*—a widely distributed and common species ranging from Colombia to Bolivia—is distinguished by its extremely small flowers and unique morphology, which complicates its association with other species. These traits suggest that *L. acarina* may represent a cleistogamous lineage, potentially following an independent evolutionary pathway within the genus.

The Nicolasii clade (Clade 3) [6] includes large-sized species with erect to suberect leaves, supported by relatively long, robust ramicauls, compact rachis, conspicuous appendix and typically with yellow flowers produced in a raceme on the leaf's abaxial surface. This group might also include species such as *L. monitor*, *L. auriculata*, and *L. steyermarkii*.

The remaining *Lepanthes* species form a polytomy, resolving into a clade with two subclades. A Central American and the Caribbean clade (Clade 4) includes *L grandiflora*, *L. turialvae*, *L. urbaniana*, which are sister to a clade represented by *L calodyction* + *L. tentaculata* and the rest of *Lepanthes*. Most species from both the *Calodictyon* clade (Clade 5) and the remaining *Lepanthes* are distributed in the Western Cordillera of Colombia and northern Ecuador, an area known as the Chocó biogeographic region [68, 69].

The Calodictyon clade (Clade 5) was also recovered in the combined *ITS/matK* phylogeny, including species such as *L. saltatrix*, *L. barbelifera*, and *L. volador* (Figs. S6–7). These species are distributed from Central to South America. While most species in this clade have reticulated leaves, those from Mesoamerica lack this feature. Their flowers emerge above the leaf, with widely spaced and reflexed Sepals that rest on the leaf surface. The petals feature a distinct upper and lower lobe. The upper lobe is often elongated, sometimes with

a filamentous extension at the apex (though not always present), while the lower lobe is typically expanded into a broader segment, occasionally with a basal filamentous extension. The labellum is simple, lacking the typical *Lepanthes* structure of a body, connectives, and laminae. Instead, it takes on reniform, lunate, ovate, or bilobed shapes and does not include an appendage. Some species have two filaments on their petals, others have one, and some lack filaments entirely. They grow at low to mid elevations 200 to 1800 m. Calodictyon-related species occur mostly in Central America (*L. arachnion*, *L. pantomima*) and the western Andes (*L. bibarbullata*, *L. arachnion*, *L. kayii*, *L. microcalodictyon*, *L. pantomima*, *L. pretiosa*, *L. tentaculata*, *L. tortuosa*) [2, 6].

The second polytomy identified within this phylogeny includes a clade with four strongly supported subclades, primarily with species from the western Andes (Clade 6): *L. mucronate*+*L. hexapus* (Mucronata clade), *L. pelix*+*L. heteroloba* (Chocó biogeographic region Clade), *L. ribes*+*L. felis* (western Andes Marsipanthes clade 1), *L. lucifer*+*L. niesseniae* (western Andes Marsipanthes clade 2) and the Manabina clade (Fig. 5). The last four species formed a clade in the combined *ITS/matk* phylogeny, and we will refer to them as the western Andes Marsipanthes species (Figs. S6-7, 11). We anticipate species such as *L. carunculigera, L. equicalceolata, L. quadricornis* will be part of this clade.

In our phylogeny, *L. hexapus* represents a group of species characterized by elliptic to ovate leaves with reticulate veins and a distribution almost exclusively from the Chocó biogeographic region [67, 68]. Based on morphology, species like *L. satyrica*, *L. hirsutula*, *L. acrogenia*, *L. tetrapus*, and *L. heptapus,* among others, could be part of this clade. A second group of morphologically similar species is represented here by *L. mucronata;* these species might have a mucron or a small central lobe on the petals. Lastly, *L. pelix*+*L. heteroloba* form part of a group of species with very corrugated leaves. Carl Luer's initial grouping termed these as the series *Mucronatae* [23]; however, he later retracted this, possibly due to the complexity and variation in flower and plant morphology.

The phylogenetic position of *L. tulcanensis* in the combined *ITS/matK* phylogenies, along with its morphological similarities to both western Andes *Marsipanthes* species and the Mucronata group, is intriguing (Figures S6–7). The plant closely resembles species from the Mucronata group, such as *L. rodophylla*, *L. pelyx*, and *L. corrugata*. However, its flowers bear a stronger resemblance to those of species in the Marsipanthes group (Fig. S11), particularly *L. niesseniae*. This is evident in features such as the long peduncle, sepals with more than three veins, a simple lip that folds beneath the column, and three-dimensional petals.

Finally, the Manabina clade has already been informally recognized by Baquero et al. [65] based on *Lepanthes manabina* Dodson. All species in this group share centrally concave leaves, ranging from deeply to slightly concave. Their leaves also have slightly to strongly recurved margins, a microscopically to conspicuously pubescent adaxial surface, and congested inflorescences. The flowers rest on the adaxial side of the leaves and are accompanied by a short to long caudate synsepal and a very small, inconspicuous appendix. Other species that might belong to this group include *L. farallonensis* Haelter., Gal.-Tar. & Zuluaga, *L. foreroi* P.Ortiz, O.Pérez & E.Parra, *L. ortiziana* O.Pérez, E.Parra & Kolan., *L. smaragdina* Luer & R.Escobar, *L. tomentosa* Luer, and *L. cincinnata* Luer & R.Escobar.

Our morphological analysis, though limited to a subset of *Lepanthes* taxa, confirmed that continuous characters do not exhibit grouping patterns consistent with phylogenetic clades. Only in a few instances, such as the artificial subgenus *Marsipanthes* (Fig. 7), do these characters show a taxonomic pattern. However, this analysis revealed intriguing trends using continuous morphological characters, which we will test further with a broader taxon sampling and additional datasets.

Measurements such as sepal size and peduncle length are key in distinguishing among *Lepanthes* species. Our PCA highlights the importance of traits like dorsal and lateral sepal dimensions, which show the highest positive PCA loadings for species differentiation (Fig. S9). The strong correlations we observed among *Lepanthes* traits, —for example, between peduncle length and floral bract length, or leaf length and ovary length— likely reflect developmental integration and genetic linkages. Feng et al. [70] similarly found a positive correlation between individual leaf area, peduncle diameter and inflorescence length in other orchids. This suggest that larger leaves can provide more resources, to support the development of thicker peduncles and longer inflorescences. Such finding point to a trade-off between leaf area and flower traits, potentially indicating developmental or genetic constraints.

Studies on phenotypic correlations across plants and animals suggest that traits within the same functional group tend to be more strongly correlated due to shared developmental pathways and functional constraints [71]. Furthermore, research in plants supports the idea that phenotypic correlations can serve as proxies for genetic correlations, indicating that these linked traits in *Lepanthes* may be co-inherited. These findings underscore the evolutionary constraints shaping morphological diversity in the genus, suggesting that certain trait combinations are maintained due to developmental or genetic dependencies [72]. Our results provide insights into

how morphological traits may be co-adapted in response to environmental pressures or reproductive strategies. Conversely, weak negative correlations, such as those between dorsal sepal width and leaf width, might indicate the independent evolution of certain vegetative and floral traits, pointing to differing selective pressures on these traits.

The hyperdiverse nature of *Lepanthes* presents significant challenges in resolving its phylogenetic relationships, particularly given the rapid diversification of the genus, which is estimated to have occurred 5–10 million years ago [30, 73]. The discovery of non-monophyletic subgeneric classifications highlights the need for a thorough reassessment of the taxonomic framework for *Lepanthes*. However, we strongly discourage any taxonomic revisions until a more comprehensively sampled phylogeny is available, particularly including representatives from all major clades and informal morphological and geographical groups. Taxon sampling is a critical factor in phylogenetic reconstructions, as demonstrated by numerous studies [74–76]. Enhanced taxon sampling, even when some data are missing, has consistently been shown to improve branch support, as seen in studies combining whole plastome sequences with datasets containing only a few chloroplast genes [29]. While this study provides some insights, the current dataset reflects the limitations of incomplete taxon sampling and the constrained informativeness of chloroplast markers. Currently our efforts are in an expanded taxon sampling, both geographically and taxonomically, to capture the *Lepanthe*'s full diversity and refine its taxonomic framework. Ongoing work on the sequencing of whole nuclear genomes of 100 additional taxa aims to address these gaps.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12862-025-02396-6.

Supplementary Material 1. Table S1. List of primers used for sequence amplification on *Marsipanthes* and other selected *Lepanthes* species.

Supplementary Material 2. Table S2. Raw and trimmed reads and percentages of reads retain after data filtering for genome skimming readings of ten *Lepanthes* species representing main taxonomic groups.

Supplementary Material 3. Table S3. Repeat sequence analysis found nine *Lepanthes* species representing main taxonomic groups.

Supplementary Material 4.Table S4. List of species, codes, herbaria and accession of orchids used in this study.

Supplementary Material 5. Figure S1. Number of *ycf1* and *ndhF* gene copies and localization in the chloroplast genome of nine *Lepanthes* species analyzed here. A. Gene copies located between IRa-SSC. B. Genes copies located between IRb-SSC.

Supplementary Material 6. Figure S2. Collinearity analysis using Mauve v1.1.3 plugin in Geneious v9.1.4 showed no gene rearrangements or inversions in the Lepanthes plastomes.

Supplementary Material 7. Figure S3. Sequence divergence of *Lepanthes* plastomes examined using the online program mVISTA Whole-genome alignment revealed that sequence variation in conserved non-coding regions (highlighted in pink bars) was greater than in protein-coding regions (highlighted in purple bars).

Supplementary Material 8. Figure S4. Heatmap of the Relative Synonymous Codon Usage (RSCU) ratio for nine *Lepanthes* plastomes was estimated using CodonW v1.4.2 [44].

Supplementary Material 9. Figure S5. Inferred ML phylogeny for the *Lepanthes* backbone using 35 species with different chloroplast gene sampling size.

Supplementary Material 10. Figure S6. Reconstructed phylogenetic tree of concatenated markers r *ITS* and *matK* with BI analysis.

Supplementary Material 11. Figure S7. Reconstructed phylogenetic tree of concatenated markers r *ITS* and *matK* with ML analysis.

Supplementary Material 12. Figure S8. *MatK* Inferred ML phylogeny for all *Lepanthes* species sequenced here and available in GenBank.

Supplementary Material 13. Figure S9. PCA loading contributions for PCA analysis of morphological continuous characters in *Lepanthes* species.

Supplementary Material 14. Figure 10. *Marsipanthes* species of the Eastern Andes. A. *Lepanthes caprimulgus* Luer. B. *Lepanthes attenboroughii* Baquero & Yeager. C. *Lepanthes portillae* Luer. D. *Lepanthes martinae* Luer & Cloes. E. *Lepanthes mulleriana* Luer. F. *Lepanthes terborchii* Luer & Sijm. G. *Lepanthes tigrina* Luer & Thoerle. Photos: A & B by Luis E. Baquero, C, E & G by Lourens Grobler and D & F by Ron Parsons

Supplementary Material 15. Figure 11. *Marsipanthes* species of the Western Andes. A. *Lepanthes ribes* Luer. B. *Lepanthes carunguligera* Rchb.f. C. *Lepanthes quadricornis* D. *Lepanthes equicalceolata* E. *Lepanthes felis* F. *Lepanthes lucifer*. G. *Lepanthes nisseniae* Luer. H. *Lepanthes tulcanensis* Baquero & Yeager. Photos: A, C & D by Lourens Grobler, B by Ron Parsons and E, F, G & H by Luis E. Baquero.

Acknowledgements

We extend our gratitude to the Sociedad Colombiana de Orquideología and Colomborquídeas in Colombia for their support with resources and access to plant collections used in this research. We also thank the Universidad de las Américas (UDLA) for funding research on Orchidaceae in Ecuador. Our appreciation goes to the Corporación para Investigaciones Biológicas for providing collection permits for TA and ASS during their affiliation with the institution. We are grateful to Idea Wild for their support in supplying equipment to ASS. Special thanks to Juan Felipe Posada, Luis Eduardo Mejía, Jean Mark Palandre, and Sebastián Vieira for their valuable input, collaboration, and assistance with plant materials. We also acknowledge Juliana Arcila and Jorge Muñoz for their expertise in bioinformatics, which was crucial to the success of this study. Finally, we thank Lourens Grobler and Ron Parsons for kindly allowing the use of their photographs of several *Lepanthes* species, which greatly enriched this research. We also appreciate Justin Yeager for his help in refining the English for this manuscript.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT, OpenAI, December 2024 to improve the readability and language of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Authors' contributions

TA: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review and editing. JSM: Conceptualization, Data curation, Investigation, Writing – original draft, Writing – review and editing. SR: Formal analysis, Methodology, Writing – review and editing. ASS: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Writing – review and editing. MLL: Formal analysis, Writing – review and editing. GAI: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration,

Resources, Supervision, Visualization, Writing – original draft, Writing – review and editing. JV: Data curation, Formal analysis, Writing – review and editing. LB: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review and editing. AZ: Data curation, Formal analysis, Methodology, Project administration, Resources, Supervision, Visualization, Writing – review and editing.

Funding
Sociedad Colombiana de Orquideología in Colombia for their support with resources. Universidad de las Américas (UDLA) for funding research on Orchidaceae in Ecuador. Idea Wild for their support in supplying equipment to ASS.

Data availability
The datasets generated and/or analysed during the current study are available in the NCBI repository, GenBank accession numbers: Bioproject PRJNA1209138, BioSample accessions: SAMN46208730, SAMN46208731, SAMN46208732, SAMN46208733, SAMN46208734, SAMN46208735, SAMN46208736, SAMN46208737, SAMN46208738, SAMN46208739.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Marie Selby Botanical Gardens, 1534 Mound Street, Sarasota, FL 34236, USA. [2]Jardín Botánico de Cali-FZC, Cra. 2 Oe. #21, Valle del Cauca, Cali, Colombia. [3]Programa de Biología, Facultad de Ciencias Exactas y Naturales, Universidad de Caldas, Calle 65 #26-10, Manizales, Colombia. [4]Departamento de Biología, Universidad del Valle, Calle 13 # 100-00, Cali, Colombia. [5]Centro de Investigación Jardín Botánico Lankester, Universidad de Costa Rica, 302–7050, Cartago, Costa Rica. [6]Grupo de Investigación en Biodiversidad, Medio Ambiente y Salud (BIOMAS), Carrera de Ingeniería Agroindustrial, Facultad de Ingeniería y Ciencias Aplicadas, Universidad de Las Américas UDLA, Vía a Nayón, Quito 170124, Ecuador. [7]Facultad de Ingeniería, Max Planck Tandem Group, Universidad Nacional de Colombia, Carrera 30 # 45-03, edificio 615, Bogotá D.C., Colombia. [8]Current address: Orchids for Peace Foundation, Research Group in Evolution, Systematics and Conservation of Neotropical Epiphyte Orchids, Sabaneta, Antioquia, Colombia.

Received: 31 December 2024 Accepted: 13 May 2025
Published online: 07 August 2025

References

1. Karremans AP. Genera Pleurothallidinarum: an updated phylogenetic overview of Pleurothallidinae. Lankesteriana. 2016;16(2):219–41. https://doi.org/10.15517/lank.v16i2.26008.
2. Luer CA, Thoerle L. Icones Pleurothallidinarum XXXI: lepanthes of Bolivia. Monogr Syst Bot Missouri Bot Garden. 2010;120:1–64.
3. Bogarín D, Cinchilla IF, Cedeno-Fonseca M. Two new species of *Lepanthes* (Orchidaceae: Pleurothallidinae) from Costa Rica and their phylogenetic affinity. Plant Syst Evol. 2020;20(306):1–13. https://doi.org/10.1007/s00606-020-01653-z.
4. POWO. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. 2024. Published on the Internet; https://powo.science.kew.org/Retrieved 02 Dec 2024.
5. Bogarín D, Jiménez D. Lepanthes novae Durikaensis. Orchids. 2015;84:430–40.
6. Luer CA. Icones Pleurothallidinarum XIV. Systematics of draconanthes, lepanthes subgenus marsipanthes and subgenus lepanthes of Ecuador. Monogr Syst Bot Missouri Bot Garden. 1996;61:1–255.
7. Luer CA, Garden LMB. Icones Pleurothallidinarum, I. systematics of the Pleurothallidinae (Orchidaceae). Monogr Syst Bot Missouri Bot Garden. 1986;15:1–81.
8. Luer CA, Vazquez R. New species of lepanthes. Phytologia. 1983;54(5):357.
9. Moreno S. Patrones de diversidad y evaluacion del estado de conservación del género Lepanthes en Colombia. Universidad del Valle, Cali, Colombia: MS Thesis; 2024. p. 152.
10. Karremans AP, Moreno JS, Gil-Amaya K, Gutiérrez Morales N, Espinosa F, Mesa S, Restrepo E, Rincón-González M, Serna A, Sierra-Ariza M, Vieira-Uribe S. Colombian Orchidaceae: a catalogue of the Pleurothallidinae. Lankesteriana. 2023;17(2):235–44. https://doi.org/10.15517/lank.v23i2.56158.
11. Reid WV. Biodiversity hotspots. Trends Ecol Evol. 1998;13(7):275–80.
12. Bogarín D, Pupulin F, Smets E, Gravendeel B. Evolutionary diversification and historical biogeography of the Orchidaceae in Central America with emphasis on Costa Rica and Panama. Lankesteriana. 2016;16(2):189–200.
13. Blanco MA, Barboza G. Pseudocopulatory pollination in Lepanthes (Orchidaceae: Pleurothallidinae) by fungus gnats. Annals of Botany. 2005;95(5):763–72.
14. Pupulin F, Bogarín D. Illustrations and studies in Neotropical Orchidaceae—the lepanthes jimenezii group (Pleurothallidinae). Harvard Papers Bot. 2010;15(1):111–21.
15. Pupulin F and Bogarín D. 2014a Aliae Lepanthes Machogaffenses (Orchidaceae: Pleurothallidinae) Harvard Papers in Botany 19(2), 193–200.
16. Pupulin F, Bogarín D. Illustrations and studies in Neotropical Orchidaceae. 6. The Lepanthes guatemalensis group (Pleurothallidinae) in Costa Rica. In Anales del Jardín Botánico de Madrid. 2014;71(1):1. Real Jardín Botánico.
17. Luer CA. New lepanthes species from Ecuador — 4. Lindleyana. 1987;2(2):95–111.
18. Luer CA. New lepanthes species from Ecuador. 5 Lindleyana. 1987;2(3):125–55.
19. Luer CA. New series and species of lepanthes from Ecuador. Novon. 1993;3:442–54.
20. Luer CA, Thoerle L. Icones Pleurothallidinarum XXXII. Lepanthes of Colombia (Orchidaceae) Monographs in Systematic Botany from the Missouri Botanical. Garden. 2012;123:1–296.
21. Bogarin D, Karremans AP, Fernandez M. Genus-level taxonomical changes in the Lepanthes affinity (Orchidaceae, Pleurothallidinae). Phytotaxa. 2018;340(2):128–36. https://doi.org/10.11646/phytotaxa.340.2.2.
22. Bogarín D, Pérez-Escobar OA, Karremans AP, Fernández M, Kruizinga J, Pupulin F, Smets E, Gravendeel B. Phylogenetic comparative methods improve the selection of characters for generic delimitations in a hyperdiverse Neotropical orchid clade. Sci Rep. 2019;9(1):15098. https://doi.org/10.1038/s41598-019-51360-0.
23. Dodson CH, Luer CA. 2011 Orchidaceae, part 7. Lepanthes and affiliates. In: C. Persson & B. Ståhl (eds), Flora of Ecuador 88. Department of Plant and Environmental Sciences, University of Gothenburg, Göteborg, Sweden
24. Reichenbach HG. Orchideae describuntur. Flora. 1886;69(35):547–62.
25. Baquero LE, Monteros MF. A new tall and exceptional species of Lepanthes from north-west Ecuador (Orchidaceae:Pleurothallidinae). Lankesteriana. 2020;20(3):331–8. https://doi.org/10.15517/lank.v20i3.
26. Yeager J, Iturralde GA, Horna LO, Baltazar AH, Baquero LE. A remarkable new lepanthes (Orchidaceae: Pleurothallidinae) hidden in plain sight. Phytotaxa. 2022;560(3):285–94. https://doi.org/10.11646/phytotaxa.560.3.3.
27. Givnish TJ, Ames M, McNeal JR, McKain MR, Steele PR, dePamphilis CW, Graham SW, Pires JC, Stevenson DW, Zomlefer WB, Briggs BG, Duvall MR, Moore MJ, Heaney JM, Soltis DE, Soltis PS, Thiele K, Lebbens-Mack JH. Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of Poales *Annals of the*. Missouri Bot Gard. 2010;97(4):584–616.
28. Steele PR, Hertweck KL, Mayfield D, McKain MR, Leebens-Mack J, Pires JC. Quality and quantity of data recovered from massively parallel sequencing: examples in Asparagales and Poaceae. Am J Bot. 2012;99(2):330–48.
29. Xi Z, Ruhfel BR, Schaefer H, Amorim AM, Sugumaran M, Wurdack KJ, Endress PK, Matthews ML, Stevens PF, Mathews S, Davis CC. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. Proc Natl Acad Sci USA. 2012;109(43):17519–24.
30. Serna-Sánchez MA, Pérez-Escobar OA, Bogarín D, Torres MF, Alvarez-Yela AC, Arcila JE, Hall CF, de Barros F, Pinheiro F, Dodsworth S, Chase

MW, Antonelli A, Arias T. Plastid phylogenomics resolves ambiguous relationships within the orchid family and provides a solid timeframe for biogeography and macroevolution. Sci Rep. 2021;11:6858.
31. Pérez-Escobar OA, Balbuena JA, Gottschling M. Rumbling orchids: how to assess divergent evolution between chloroplast endosymbionts and the nuclear host. Syst Biol. 2016;65(1):51–65.
32. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.
33. Kasajima I, Ide Y, Ohkama-Ohtsu N, Hayashi H, Yoneyama T, Fujiwara T. A protocol for rapid DNA extraction from Arabidopsis thaliana for PCR analysis. Plant Mol Biol Report. 2004;22(1):49–52. https://doi.org/10.1007/BF02773348.
34. Brown J, Pirrung M, McCue LA. FQC dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics. 2017;33(19):3137–9.
35. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
36. Jin JJ, Yu WB, Yang JB, Song Y, DePamphilis CW, Yi TS, Li DZ. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology. 2020;21:1–31.
37. Mauad AV, do Nascimento Vieira L, Bolson M, Baura V, Balsanelli E, Souza E. Complete chloroplast genome of Anathallis obovata (Orchidaceae: Pleurothallidinae). Brazilian Journal of Botany. 2019;42:345.
38. 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(W1):W59–64.
39. Chen Y, Ye W, Zhang Y, Xu Y. High speed BLASTN: an accelerated MegaBLAST search tool. Nucleic Acids Res. 2015;43(16):7762–8.
40. Gish W, States DJ. Identification of protein coding regions by database similarity search. Nat Genet. 1993;3(3):266–72.
41. Amiryousefi A, Hyvönen J, Poczai P. IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics. 2018;34:3030–1.
42. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. DnaSP 520 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 2017;34:3299–302.
43. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27:573–80.
44. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: the manifold applications of repeat 506 analysis on a genomic scale. Nucleic Acids Res. 2001;29:4633–42.
45. Sun XY, Yang Q, Xia XH. An improved implementation of effective number of codons (nc). Molecular Biology and Evolution. 2013;30:191–6.
46. Katoh K, Toh H. Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics. 2010;26(15):1899–900. https://doi.org/10.1093/bioinformatics/btq224.
47. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML web servers. Syst Biol. 2008;57(5):758–71.
48. Ronquist F, Huelsenbeck JP. MrBayes 3: bayesian phylogenetic inference under mixed models. Bioinformatics. 2003;19(12):1572–4. https://doi.org/10.1093/bioinformatics/btg180.
49. Ortiz EM, Höwener A, Shigita G, Raza M, Maurin O, Zuntini A, Félix F, Baker W, SchaeferH. A novel phylogenomics pipeline reveals complex pattern of reticulate evolution in Cucurbitales. bioRxiv. 2023:2023-10.
50. 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. Systematic biology. 2010;59(3):307–21.
51. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. virus. Evolution. 2018;4(1):vey016. https://doi.org/10.1093/ve/vey016.
52. 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 molecular biology. 2011;76:273–97.
53. Zhao Z, Li Y, Zhai JW, Liu ZJ, Li MH. Organelle genomes of *Epipogium roseum* provide insight into the evolution of mycoheterotrophic orchids. Int J Mol Sci. 2024;25(3): 1578.
54. Guo YY, Yang JX, Li HK, Zhao HS. Chloroplast genomes of two species of Cypripedium: Expanded genome size and proliferation of AT-biased repeat sequences *Frontiers in Plant*. Science. 2021;12:609729.
55. Goulding SE, Wolfe KH, Olmstead RG, Morden CW. Ebb and flow of the chloroplast inverted repeat. Mol Gen Genet MGG. 1996;252:195–206.
56. Wang RJ, Cheng CL, Chang CC, Wu CL, Su TM, Chaw SM. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. BMC Evol Biol. 2008;8:1–14.
57. Wu FH, Chan MT, Liao DC, Hsu CT, Lee YW, Daniell H, Melvin RD, Lin CS. Complete chloroplast genome of Oncidium Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. BMC Plant Biol. 2010;10:1–12.
58. Ruhlman TA, Jansen RK. The plastid genomes of flowering plants. Chloroplast Biotechnol. 2014;1:3–38.
59. Barrett CF, Davis JI. The plastid genome of the mycoheterotrophic Corallorhiza striata (Orchidaceae) is in the relatively early stages of degradation. American Journal of Botany. 2012;99(9):1513–23.
60. Drescher A, Ruf S, Calsa T Jr, Carrer H, Bock R. The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. Plant J. 2000;22(2):97–104.
61. Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, Chen T, Guo J, Zhou S. ycf1, the most promising plastid DNA barcode of land plants. Sci Rep. 2015;5(1):1–5.
62. Kim YK, Jo S, Cheon SH, Joo MJ, Hong JR, Kwak M, Kim KJ. Plastome evolution and phylogeny of Orchidaceae, with 24 new sequences. Front Plant Sci. 2020;11:22.
63. Morton BR. Selection on the codon bias of chloroplast and cyanelle genes in different plant and algal lineages. J Mol Evol. 1998;46(4):449–59. https://doi.org/10.1007/PL00006321.
64. Zhou C, He X, Li F, Weng Q, Yu X, Wang Y, Li M, Shi J, Gan S. Development of 240 novel EST-SSRs in Eucalyptus L'Hérit. Molecular Breeding. 2014;33:221–5.
65. Baquero LE, Galindo R, Haelterman D, Zuluaga A. A new Lepanthes (Pleurothallidinae) from south-western Colombia Lankesteriana 2021 21(3), 341–349.
66. Ocupa-Horna L, Thoerle L, Schuiteman A, Martel C. A revised checklist of lepanthes (Orchidaceae) from Peru, including new country records. Brittonia. 2023;75(2):159–79.
67. Lagou LJ, Kadereit G, Morales-Briones DF. Phylogenomic analysis of target enrichment and transcriptome data uncovers rapid radiation and extensive hybridization in the slipper orchid genus Cypripedium. Ann Bot. 2024;134(7):1229–50.
68. Hazzi NA, Moreno JS, Ortiz-Movliav C, Palacio RD. Biogeographic regions and events of isolation and diversification of the endemic biota of the tropical Andes. Proc Natl Acad Sci. 2018;115(31):7985–90.
69. Pérez-Escobar OA, Lucas E, Jaramillo C, Monro A, Morris SK, Bogarín D, Greer D, Dodsworth S, Aguilar-Cano J, Sanchez Meseguer A, Antonelli A. The origin and diversification of the hyperdiverse flora in the Chocó biogeographic region *Frontiers in Plant*. Science. 2019;10:1328.
70. Feng JQ, Zhang FP, Huang JL, Hu H, Zhang SB. Allometry between vegetative and reproductive traits in orchids. Front Plant Sci. 2021;12:728843.
71. Conner JK, Cooper IA, La Rosa RJ, Pérez SG, Royer AM. Patterns of phenotypic correlations among morphological traits across plants and animals. Philosophical Transactions of the Royal Society B: Biological Sciences. 2014;369(1649):20130246.
72. Waitt DE, Levin DA. Genetic and phenotypic correlations in plants: a botanical test of Cheverud's conjecture. Heredity. 1998;80(3):310–9.
73. Pérez-Escobar OA, Chomicki G, Condamine FL, Karremans AP, Bogarín D, Matzke NJ, Silvestro D, Antonelli A. Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. New Phytologist. 2017;215(2):891–905.
74. Sanderson MJ, McMahon MM, Steel M. Phylogenomics with incomplete taxon coverage: the limits to inference. BMC Evol Biol. 2010;10:1–13.
75. Philippe H, Brinkmann H, Lavrov DV, Littlewood DTJ, Manuel M, Wörheide G, Baurain D. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 2011;9(3): e1000602.
76. Barrett CF, Specht CD, Leebens-Mack J, Stevenson DW, Zomlefer WB, Davis JI. Resolving ancient radiations: can complete plastid gene sets elucidate deep relationships among the tropical gingers (Zingiberales)? Ann Bot. 2014;113(1):119–33.
77. Jansen RK, Cai Z, Raubeson LA, Daniell H, Depamphilis CW, Leebens-Mack J, Muller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee S-B, Peery R, McNeal JR, Kueh JV, Boore JL. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies

genome-scale evolutionary patterns. Proceedings of the National Academy of Sciences. 2007;104(49):19369–74.
78. Bridson D, Forman L. The Herbarium Handbook. Royal Botanic Gardens, Kew; 1992. p. 346.
79. Frazer KA, et al. VISTA: Computational tools for comparative genomics. Nucleic Acids Research, 32 (Web Server issue), W273–W279. 2004.

Publisher's Note