Journal of Asia-Pacific Biodiversity 18 (2025) 457–462

Contents lists available at ScienceDirect

Journal of Asia-Pacific Biodiversity

journal homepage: http://www.elsevier.com/locate/japb

Short Communication

The complete chloroplast genome and phylogenetic position of *Cymbidium hartinahianum*, an endemic and endangered orchid of Sumatra, Indonesia

Aninda Retno Utami Wibowo [a,*], Lina Susanti Juswara [a], Destri Destri [a], Diego Bogarín [b,c], Imam Bagus Nugroho [d,e], Izu Andry Fijridiyanto [a], Richa Kusuma Wati [a,b]

[a] Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), Cibinong, Bogor, West Java, Indonesia
[b] Evolutionary Ecology Group, Naturalis Biodiversity Center, Leiden, The Netherlands
[c] Jardín Botánico Lankester, Universidad de Costa Rica, P.O. Box 302-7050, Cartago, Costa Rica
[d] Bioindustry Lab., Department of Agro-industrial Technology, Faculty of Agricultural Technology, Universitas Gadjah Mada, Yogyakarta, Indonesia
[e] Genetic Engineering Lab., Research Center for Biotechnology, Universitas Gadjah Mada, Yogyakarta, Indonesia

ARTICLE INFO

Article history:
Received 19 July 2024
Received in revised form
3 December 2024
Accepted 3 December 2024
Available online 21 December 2024

Keywords:
endemic
next-generation sequencing
orchids
phylogeny
plant taxonomy
Sumatra

ABSTRACT

Cymbidium (Orchidaceae: Epidendroideae) includes species of significant horticultural and ecological importance, yet limited research exists on the distribution and phylogeny of *Cymbidium hartinahianum*, an orchid endemic to North Sumatra, Indonesia. Here, we present the complete plastid genome of *Cymbidium hartinahianum*. The whole plastid genome was sequenced using the Illumina platform, followed by a pipeline for plastome assembly, annotation, and visualization. The circular genome with a length of 150,370 bp possesses the typical quadripartite structure, consisting of a large single-copy (LSC) region of 84,189 bp, a small single-copy (SSC) region of 14,611 bp, and a pair of inverted repeats (IRs) regions of 25,785 bp. The plastid genome has a total GC content of 37.0%. Annotation revealed 148 gene species, including 96 protein-coding genes, 52 transfer RNA, and four ribosomal RNA (rRNA) genes. Bayesian and maximum likelihood phylogenetic analyses based on 52 plastomes confirmed that *C. hartinahianum* is sister to *C. floribundum* and belongs to the *Cymbidium* section *Floribundum*. This newly sequenced plastome data will be valuable for future *Cymbidium* phylogenomic and biogeographical studies. It also provides new molecular data for conservation efforts, including DNA barcoding to help prevent the illegal trade of this endangered orchid.

Introduction

The genus *Cymbidium* Sw. comprises 88 accepted species, with a distribution ranging from India, Central and South China, Korea, and Japan to Southeast Asia and Western Australia (Govaerts et al. 2024). Indonesia is home to approximately 13 species of *Cymbidium*, including one endemic species. *Cymbidium hartinahianum* J.B.Comber & Nasution is a terrestrial orchid endemic to the region from Siborong-borong to Sidikalang, North Sumatra, Indonesia. Plants area was found in areas with elevations ranging from 1680 to 2600 meters (Comber 2001). This species was named in honor of the former first lady Siti Hartinah Soeharto, the wife of Indonesia's second President, Soeharto, and it is listed as one of the protected species based on the Indonesian government regulation of PP No 7 in 1999. Since it was found in 1976 in Baniara village, Siborong-borong, only two additional records have been documented, one in Mount Leuser (Comber 2001) and another recently in Mount Sibuatan (Prakasa et al. 2021).

This endemic orchid faces significant threats due to habitat degradation, particularly from the extensive expansion of oil palm plantations (Abood et al. 2015; Juniyanti et al. 2021). *Cymbidium* species known for their arching and pendulous inflorescences and high flower longevity, are widely used as parents for hybridization (Pal et al. 2020). *Cymbidium hartinahianum* belongs to section *Floribundum* Seth & Cribb section based on molecular evidence and clusters with the type species *C. floribundum* Lindl. (Zhang et al. 2021; Zhang et al. 2022; Chen et al. 2024). A previous study on

* Corresponding author: ORCID.: 0000-0003-3860-6481
E-mail address: aninda.wibowo@gmail.com (ARU Wibowo).
Peer review under responsibility of National Science Museum of Korea (NSMK) and Korea National Arboretum (KNA).

https://doi.org/10.1016/j.japb.2024.12.003

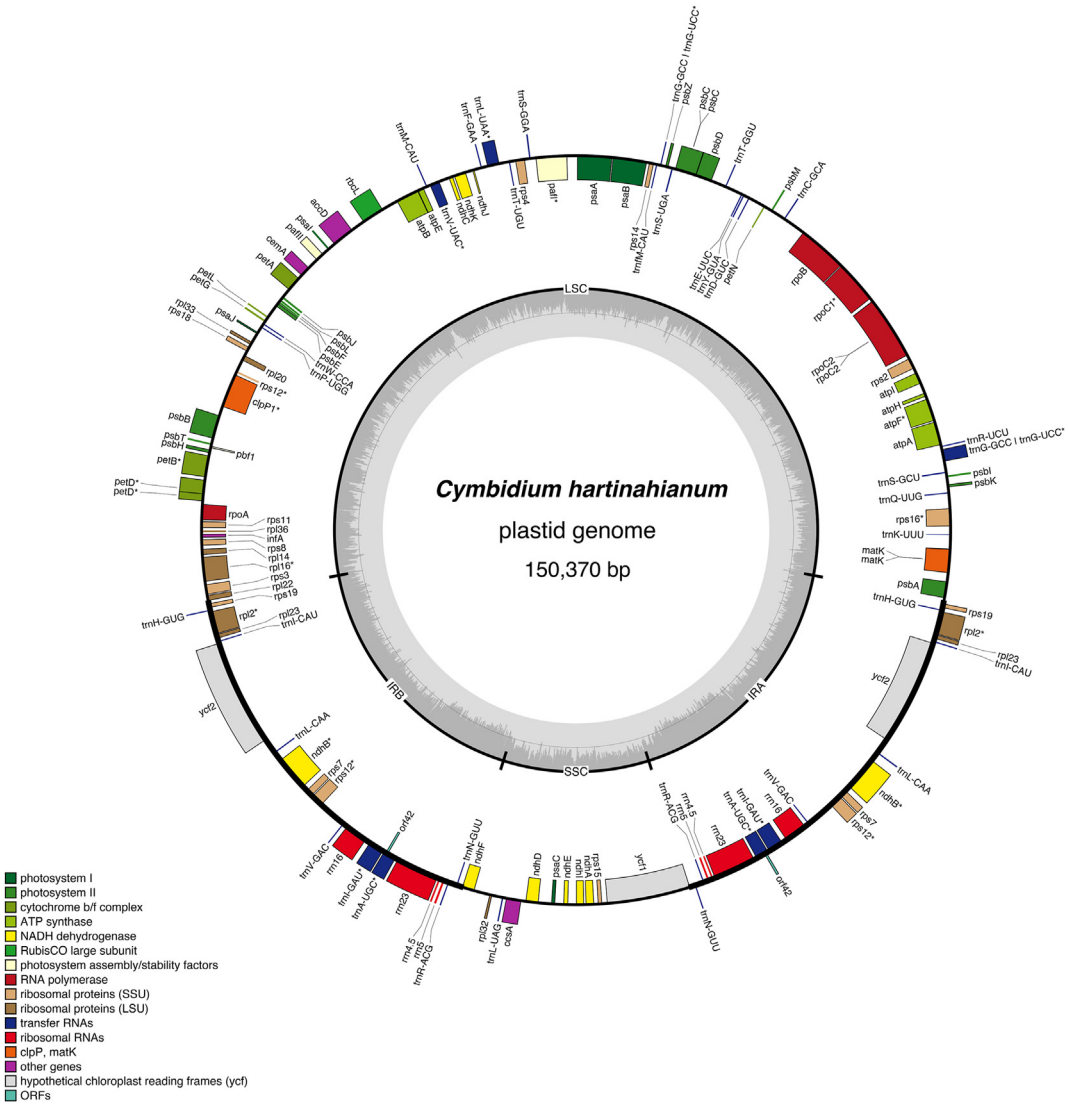

Figure 1. The complete chloroplast genome map of *Cymbidium hartinahianum*. The complete chloroplast genome map of *Cymbidium hartinahianum*. Genes indicated inside the circle are transcribed clockwise, and the outside is counterclockwise. Genes belonging to different functional groups are colored. The inner circle shows the LSC (large-single copy), SSC (small-single copy), and IR (inverted repeat) regions. The darker gray in the inner circle corresponds to GC content, and the lighter gray corresponds to AT content.

the plastid phylogenomics of *Cymbidium* positioned *C. floribundum* as a sister to *C. hartinahianum* (Chen et al. 2024). This study reports the plastid genome of *C. hartinahianum* sequenced from a herbarium specimen. However, wild populations of *C. hartinahianum* are endangered and poorly understood at the molecular and phylogenetic levels. Here, we present a complete, verified plastid genome of *C. hartinahianum* from a wild specimen and clarify its phylogenetic position, offering additional data for future *Cymbidium* phylogenomic and biogeographical studies. Furthermore, this information provides critical insights for conservation efforts, including DNA barcoding to help prevent the illegal trade of this endangered orchid.

Material and methods

Plant materials, genomic DNA isolation, and sequencing

Fresh plant leaves were collected from a specimen in the orchid collection of Bogor Botanic Gardens. The wild specimen originated from a small population in Baniara Tele village, Siborong-borong, Sumatra, Indonesia, that was collected in scientific exploration by LIPI in 2004 under the permission of the local Natural Resources Conservation Agency (Balai Konservasi dan Sumber Daya Alam). The voucher specimen is kept at the Herbarium Bogoriense under accession number BO200412429. Total genomic DNA was extracted from fresh plant leaves using the modified CTAB procedure from Doyle and Doyle (1987). DNA purity was measured using Nanodrop 2000 (Thermo Scientific) with the purity results from 1.92 to 1.95 on A260/280. DNA concentration was measured using Qubit dsDNA HS Assay Kits (Thermo Scientific) with concentration results from a minimum of 50ng/μl, and DNA integrity was measured by 4150 TapeStation (Agilent) with the results of DNA integrity number 5–8. The library preparation and high throughput sequencing were conducted by Novogene (Beijing, China) on the Illumina Novaseq 6000 platform for PE150 sequencing (Illumina, San Diego, CA). The whole genome sequence resulted in 10 Gb raw reads for each sample.

Genome assembly and post-assembly analyses

Sequencing adapter sequences were removed from the raw reads using TrimGalore-0.6.6 (https://github.com/FelixKrueger/

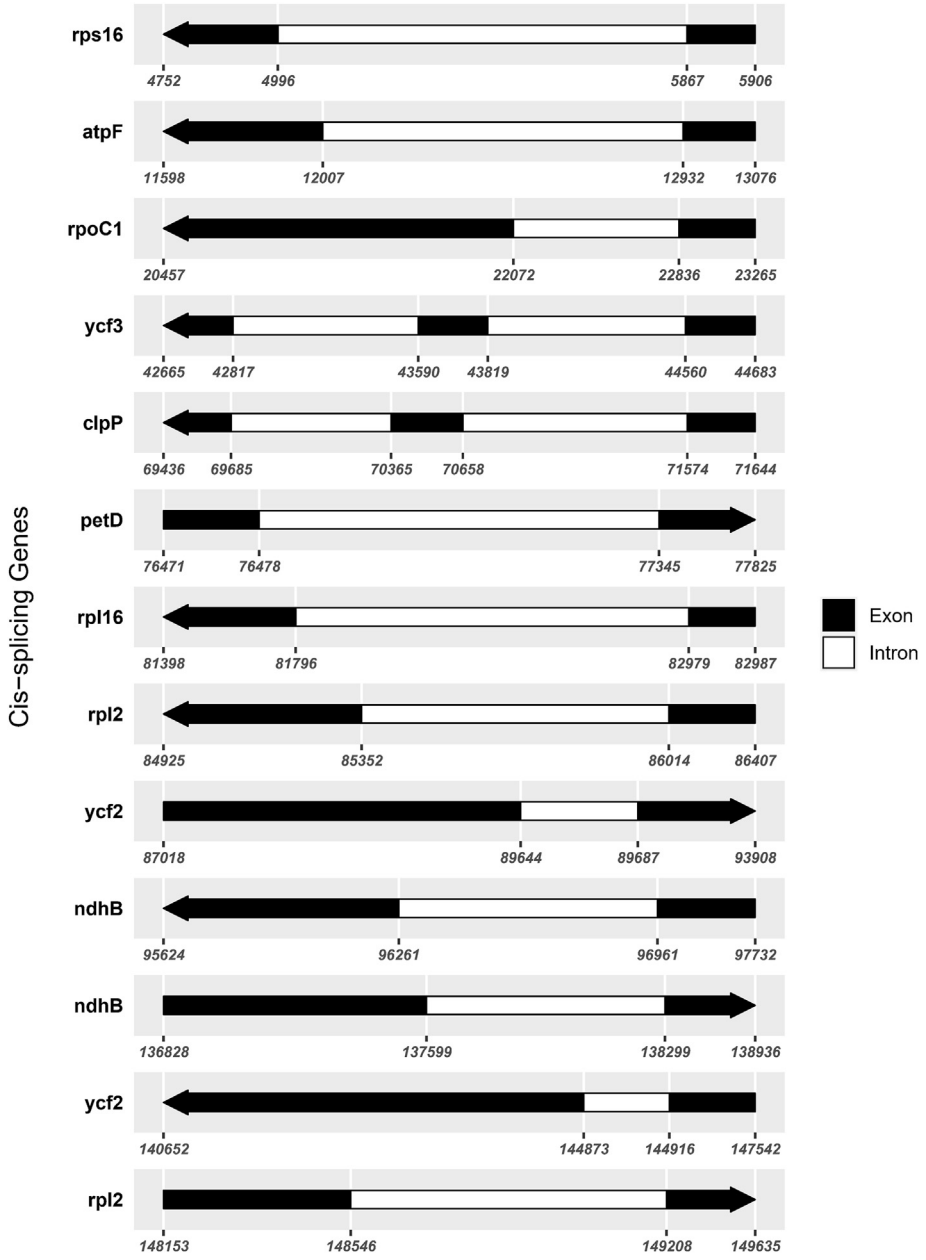

Figure 2. List of cis-splicing genes in the chloroplast of *C. hartinahianum*. The arrows depict the direction of transcription. The numbers are the starting and ending points of each exon in the genome. *C. hartinahianum, Cymbidium hartinahianum.*

TrimGalore), followed by a quality control assessment step using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) (Cock et al. 2009). After filtering approximately 10 gigabytes (GB) of paired-end clean reads, they were assembled with GetOrganelle (Jin et al. 2020). The assembled genome was checked using Bandage (Wick et al. 2015). Annotation of the plastid genome was performed using the online Chlorobox program GeSeq (Tillich et al. 2017) and refined using CPGVAS2 (Shi et al. 2019) to determine repeat elements followed by visualization using CPGView (Liu et al. 2023). The raw reads were deposited in NCBI BioProject (https://www.ncbi.nlm.nih.gov/bioproject) under the accession number PRJNA1118446 and the annotated plastome under GenBank accession number PQ397568.1. For the plastome assembly pipelines, see the supplementary file.

Phylogenetic analysis

To determine the phylogenetic relationships of *C. hartinahianum*, we retrieved 50 complete plastid genome sequences directly from NCBI GenBank when available as annotated files or assembled plastomes using GetOrganelle with raw reads downloaded from the Sequence Read Archive of the NCBI GenBank repository (https://www.ncbi.nlm.nih.gov/sra) via the fastq-dump tool from the Sequence Read Archive Toolkit v.2.10.5. The complete plastid genomes were aligned using HomBlocks (https://github.com/fenghen360/HomBlocks) (Bi et al. 2018), a tool tailored for the alignment of homologous gene sequences. Then the best-fit nucleotide substitution model was calculated using PartitionFinderV1.1.1, maximum likelihood (ML) phylogenetic analyses were inferred with

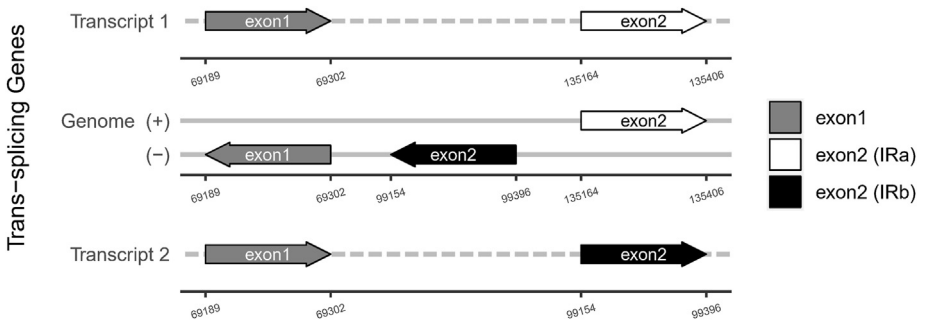

Figure 3. The trans-splicing *rps12* gene in the chloroplast of *C. hartinahianum*. The diagram shows duplicated exons within the inverted repeat (IR) regions. The arrows depict the direction of transcription. The numbers are the starting and ending points of each exon in the genome. *C. hartinahianum, Cymbidium hartinahianum.*

RAxML-VI-HPC Version 2.2.3 using 1,000 bootstrap replicates and *Grammatophyllum scriptum* (GenBank accession no. OP142289) as an outgroup. Bayesian inference (BI) was performed in MrBayes v.3.2.6. The dataset was partitioned into five subsets. Each subset was assigned a general time-reversible model with gamma-distributed rate variation, and invariant sites were included for subsets 1, 2, 4, and 5. Partition-specific rates were allowed to vary, and parameters such as state frequencies, substitution matrices, and rate heterogeneity were unlinked across partitions. The analysis was run for 10 million generations with four chains, sampling every 1000 generations. The first 2500 samples were discarded as burn-in. Convergence diagnostics were monitored every 1000 generations, and a majority-rule consensus tree was generated.

Results and discussion

Complete chloroplast genome of Cymbidium hartinahianum

The complete chloroplast genome sequence of *C. hartinahianum* was 150,370 bp in length, comprising a large single-copy (LSC) legion of 84,189 bp, a small single-copy (SSC) region of 14,611 bp, and a pair of IRs regions of 25,785 bp each (Figure 1). The overall GC content was 37.0%, while the corresponding values of the LSC, SSC, and IR regions were 34.22%, 29.76%, and 43.52%, respectively. A total of 148 genes, including 96 protein-coding genes and 52 transfer RNA, in addition to four ribosomal RNA genes in two copies, were annotated (Table 1). Additionally, a total of 62 exons

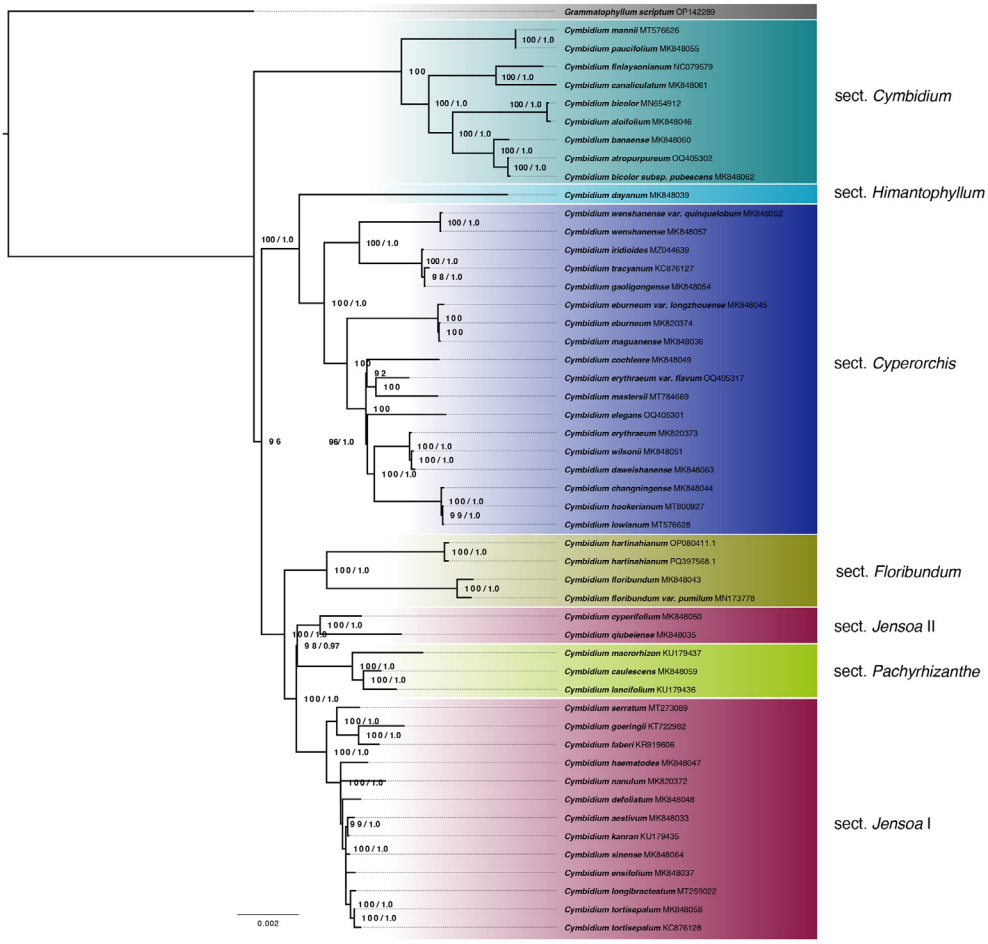

Figure 4. Maximum-likelihood consensus tree based on 51 complete plastid genomes of *Cymbidium* and *Grammatophyllum* as outgroup. Bootstrap percentages and posterior probabilities are shown at the nodes with high support (ML=95, BI=0,9). ML, maximum likelihood; BI, Bayesian inference.

Table 1. Summary of gene groups and their associated genes in the *Cymbidium hartinahianum* plastid genome.

Group of Genes	Gene Name
Photosystem I	*psa*A, *psa*B, *psa*C, *psa*I, *psa*J
Photosystem II	*psb*A, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*I, *psb*J, *psb*K, *psb*L, *psb*M, *psb*T, *psb*Z
Cytochrome b6/f	*pet*A, *pet*B, *pet*D, *pet*G, *pet*L, *pet*N
ATP synthase	*atp*A, *atp*B, *atp*E, *atp*F, *atp*H, *atp*I
NADH (Nicotinamide adenine dinucleotide) dehydrogenase	*ndh*A, *ndh*B, *ndh*C, *ndh*D, *ndh*E, *ndh*F, *ndh*G, *ndh*H, *ndh*I, *ndh*J, *ndh*K
Large ribosomal subunit	*rpl*2, *rpl*14, *rpl*16, *rpl*20, *rpl*22, *rpl*23, *rpl*32, *rpl*33, *rpl*36
Small ribosomal subunit proteins	*rps*2, *rps*3, *rps*4, *rps*7, *rps*8, *rps*11, *rps*12, *rps*14, *rps*15, *rps*16, *rps*18, *rps*19
RNA polymerase	*rpo*A, *rpo*B, *rpo*C1, *rpo*C2
tRNAs	*trn*A-UGC, *trn*C-GCA, *trn*D-GUC, *trn*E-UUC, *trn*F-GAA, *trn*G-GCC, *trn*G-UCC, *trn*H-GUG, *trn*I-CAU, *trn*I-GAU, *trn*K-UUU, *trn*L-CAA, *trn*L-UAA, *trn*L-UAG, *trn*M-CAU, *trn*N-GUU, *trn*P-UGG, *trn*Q-UUG, *trn*R-ACG, *trn*R-UCU, *trn*S-GCU, *trn*S-GGA, *trn*S-UGA, *trn*T-GGU, *trn*T-UGU, *trn*V-GAC, *trn*V-UAC, *trn*W-CCA, *trn*Y-GUA
Ribosomal RNAs	*rrn*4.5, *rrn*5, *rrn*16, *rrn*23
Other genes	*acc*D, *clp*P, *mat*K, *ycf*1, *ycf*2, *ycf*3, *ycf*4, *cem*A, *ccs*A, *inf*A

and 35 introns were identified. We also found 13 cis-splicing genes (Figure 2) and a trans-splicing gene for *rps12* with duplicated genes located in the IR region (Figure 3). Plastid genome characteristics of *C. hartinahianum* falls within the *Cymbidium* plastome range (Chen et al. 2024).

Phylogenetic relationships within the genus Cymbidium

The consensus phylogenetic tree based on chloroplast genome sequences was constructed using RAxML and MrBayes on previously built multiple sequence alignment by HomBlocks (Figure 4). The results from ML and BI showed congruent topologies. Most branch nodes were supported with high bootstrap percentages (ML=81%–100%) and posterior probabilities (PP=1.0). Our phylogenetic result is in agreement with the result of Chen et al. (2024), the only publication of the chloroplast genome of *C. hartinahianum,* whereas this species is included in section *Floribundum* with high bootstrap percentage of 100%. Although *C. hartinahianum* and *C. floribundum* are classified into one section of *Floribundum*, the morphology of *C. hartinahianum* can be differentiated from *C. floribundum,* such as by having larger pseudobulbs (7 cm × 3.5 cm vs 1.5 cm × 1 cm), longer inflorescences (50–100 cm vs 10–15 cm) and longer labellum (2.1 cm vs 1.5 cm long) (Comber 2001; Flora of China Editorial Committee 2009). Furthermore, C. *hartinahianum* is also only known as a terrestrial orchid growing in highlands (1680–2600 m). In contrast, *C. floribundum* grows as terrestrial, epiphyte, or lithophyte in lowlands up to highlands 400–3300 m) (Comber 2001; Flora of China Editorial Committee 2009).

From our result, we noticed that sect. Jensoa is a non-monophyletic group as also reported in the previous study (Zhang et al. 2022; Chen et al. 2024). Furthermore, we acknowledge that completing species samples from sect. *Floribundum*, sect. *Jensoa*, sect. *Pachyrhizanthe*, and sect. *Geocymbidium* are urgently needed.

Conclusion

This study presents the complete, annotated plastid genome of *Cymbidium hartinahianum*, an endemic species of Indonesia, obtained from a wild plant cultivated at the Bogor Botanic Gardens. The plastome provides valuable data for future phylogenomic and biogeographical research on *Cymbidium* and confirms its close relationship to *C. floribundum*, placing it within *Cymbidium* section *Floribundum*. Additionally, this data will support conservation efforts by enabling DNA barcoding for precise species identification and helping to prevent the potential illegal trade of *C. hartinahianum*.

CRediT authorship contribution statement

Aninda Retno Utami Wibowo: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Lina Susanti Juswara:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, pProject administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Destri Destri:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Investigation, Formal analysis, Data curation, Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Investigation, Formal analysis, Data curation. **Diego Bogarín:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Imam Bagus Nugroho:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation. **Izu Andry Fijridiyanto:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Investigation, Formal analysis, Data curation. **Richa Kusuma Wati:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work report in this paper.

Acknowledgments

This work was supported by the Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), under Grant Rumah Program ORHL- BRIN No 9/ III/ HK/2022. This research was partially supported by computational resources provided by the Kabré supercomputer at the Costa Rica National High Technology Center.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.japb.2024.12.003.

References

Abood SA, Lee JSH, Burivalova Z, et al. 2015. Relative Contributions of the Logging, Fiber, Oil Palm, and Mining Industries to Forest Loss in Indonesia. *Conservation Letters* 8:58–67. https://doi.org/10.1111/conl.12103.

Bi G, Mao Y, Xing Q, et al. 2018. HomBlocks: A multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching. *Genomics* 110:18–22. https://doi.org/10.1016/j.ygeno.2017.08.001.

Cock PJA, Fields CJ, Goto N, et al. 2009. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research* 38:1767–1771. https://doi.org/10.1093/nar/gkp1137.

Chen HY, Zhang ZR, Yao X, et al. 2024. Plastid phylogenomics provides new insights into the systematics, diversification, and biogeography of *Cymbidium*

(Orchidaceae). *Plant Diversity* 46:448–461. https://doi.org/10.1016/j.pld.2024.03.001.
Comber JB. 2001. *Orchids of Sumatra*. Kew: Royal Botanic Gardens, Kew Publishing.
Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11–15.
Flora of China Editorial Committee. 2009. Flora of China Orchidaceae. In: Wu CY, Raven H, Hong DY, editors. *Flora of China*. St.Louis: Missouri Botanical Garden Press. pp. 260–280.
Govaerts R, editor. 2024. *WCVP: World Checklist of Vascular Plants*.http://sftp.kew.org/pub/data-repositories/WCVP/. [Accessed 21 May 2024].
Jin JJ, Yu W Bin, Yang JB, et al. 2020. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology* 21:1–31. https://doi.org/10.1186/s13059-020-02154-5.
Juniyanti L, Purnomo H, Kartodihardjo H, et al. 2021. Understanding the driving forces and actors of land change due to forestry and agricultural practices in sumatra and kalimantan: A systematic review. *Land* 10. https://doi.org/10.3390/land10050463.
Liu S, Ni Y, Li J, et al. 2023. CPGView: A package for visualizing detailed chloroplast genome structures. *Molecular Ecology Resources* 23:694–704. https://doi.org/10.1111/1755-0998.13729.
Pal R, Meena NK, Pant RP, et al. 2020. Cymbidium: Botany, Production, and Uses. In: Mérillon JM, Kodja H, editors. *Ref Ser Phytochem*. Switzerland: Springer Nature. pp. 1–37. Switzerland.
Prakasa H, Syafitri FI, Pulungan IAH, et al. 2021. Identifikasi Habitat Anggrek Tien Soeharto (*Cymbidium Hartinahianum* JB Comber & Nasution) Di Gunung Sibuatan Sumatera Utara. *Journal of Natural Science Research* 2:8–17.
Shi L, Chen H, Jiang M, et al. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Resources* 47:W65–W73. https://doi.org/10.1093/nar/gkz345.
Tillich M, Lehwark P, Pellizzer T, et al. 2017. GeSeq - Versatile and accurate annotation of organelle genomes. *Nucleic Acids Resources* 45:W6–W11. https://doi.org/10.1093/nar/gkx391.
Wick RR, Schultz MB, Zobel J, et al. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 31:3350–3352.
Zhang GQ, Chen GZ, Chen LJ, et al. 2021. Phylogenetic incongruence in *Cymbidium* orchids. *Plant Diversity* 43:452–461. https://doi.org/10.1016/j.pld.2021.08.002.
Zhang L, Huang YW, Huang JL, et al. 2022. DNA barcoding of *Cymbidium* by genome skimming: Call for next generation nuclear barcodes. *Molecular Ecology Resources* 23:424–439. https://doi.org/10.1111/1755-0998.13719.