



OPEN ACCESS

EDITED BY

Stefan Kusch,
RWTH Aachen University, Germany

REVIEWED BY

Alex Zaccaron,
Oregon State University, United States
Pablo Andrés Pérez Mesa,
Helmholtz Association of German Research
Centres (HZ), Germany

*CORRESPONDENCE

Jie Gao

✉ gaojie@gdaas.cn

Zhiqiang Wu

✉ wuzhiqiang@caas.cn

†These authors have contributed equally to
this work and share first authorship

RECEIVED 28 November 2025

REVISED 30 December 2025

ACCEPTED 09 January 2026

PUBLISHED 12 February 2026

CITATION

Lin S, Song C, Wang J, Kong J, Peng D, Gao J
and Wu Z (2026) Comparative organellar
genomics of *Arundina graminifolia*:
mitochondrial complexity and plastid
conservation in Orchidaceae.
Front. Plant Sci. 17:1756243.
doi: 10.3389/fpls.2026.1756243

COPYRIGHT

© 2026 Lin, Song, Wang, Kong, Peng, Gao and
Wu. This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums
is permitted, provided the original author(s)
and the copyright owner(s) are credited and
that the original publication in this journal is
cited, in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Comparative organellar genomics of *Arundina graminifolia*: mitochondrial complexity and plastid conservation in Orchidaceae

Shengnan Lin^{1,2,3,4,5†}, Chao Song^{6†}, Jie Wang^{3,4,5}, Jiali Kong^{3,4,5},
Dan Peng^{3,4,5,7}, Jie Gao^{8*} and Zhiqiang Wu^{1,2,3,4,5*}

¹Kunpeng Institute of Modern Agriculture at Foshan, Chinese Academy of Agricultural Sciences, Shenzhen, China, ²Guangdong Laboratory of Lingnan Modern Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China, ³State Key Laboratory of Tropical Crop Breeding, Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China, ⁴Guangdong Laboratory of Lingnan Modern Agriculture, Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China, ⁵Key Laboratory of Synthetic Biology, Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China, ⁶Chongqing Key Laboratory for Germplasm Innovation of Special Aromatic Spice Plants, College of Smart Agriculture, Chongqing University of Arts and Sciences, Chongqing, China, ⁷School of Ecology, Hainan University, Haikou, China, ⁸Guangdong Key Laboratory of Ornamental Plant Germplasm Innovation and Utilization, Environmental Horticulture Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, China

Orchidaceae is one of the largest and most diverse angiosperm families, exhibiting remarkable morphological and ecological diversity. Organellar genomes, including mitochondrial and plastid genomes, play essential roles in energy metabolism, photosynthesis, and adaptive evolution, yet their structural evolution in orchids remains unclear. In this study, the mitochondrial and plastid genomes of *Arundina graminifolia* were assembled and compared with 15 representative orchid species to investigate genome architecture, repeat content, gene composition, and evolutionary dynamics. The mitochondrial genome of *A. graminifolia* displayed a complex multi-circular structure with extensive rearrangements and abundant repeats, including simple sequence repeats and long repeats, contributing to genome expansion and structural complexity. In contrast, the plastid genome was highly conserved, exhibiting a typical quadripartite structure. Plastid-to-mitochondrion DNA transfers (MTPTs) varied among orchid species in both number and composition, mainly involving complete plastid genes associated with photosynthesis and translation. Analyses of selection pressure and nucleotide diversity revealed strong purifying selection and low sequence variability in mitochondrial genes, whereas plastid genes exhibited higher diversity and evolutionary rates. Phylogenetic analyses based on conserved single-copy genes from both organellar genomes yielded congruent topologies, with *A. graminifolia* clustering closely with *Bletilla striata*. Synteny analysis indicated low conservation of mitochondrial genome organization across lineages, particularly in mycoheterotrophic orchids,

reflecting dynamic genome evolution. Gene loss and duplication further enhanced genome plasticity and potential functional redundancy. Overall, this study highlights contrasting evolutionary constraints and structural dynamics between mitochondrial and plastid genomes in orchids, providing valuable genomic resources for understanding organelle genome evolution and phylogenetic relationships in Orchidaceae.

KEYWORDS

genome evolution, mitochondrial genome, Orchidaceae, organelle genomes, plastid genome

Introduction

The Orchidaceae is one of the largest and most diverse families of angiosperms, encompassing over 28,000 species that are widely distributed across tropical, subtropical, and temperate regions (Chase et al., 2015; Christenhusz, 2016; Zhang et al., 2022). Orchid species exhibit remarkable morphological and ecological diversity, including complex reproductive strategies, diverse life forms (terrestrial, epiphytic, or mycoheterotrophic), and highly specialized floral structures, making them an important model for studying plant evolution, adaptation, and ecological function (Dressler, 1981; Dressler, 1993; Rasmussen and Rasmussen, 2009; Givnish et al., 2015; Zhang et al., 2018). Therefore, investigating their genome features not only enhances understanding of Orchidaceae evolution but also provides molecular foundations for conservation, breeding, and sustainable utilization.

Organellar genomes, including mitochondrial and plastid genomes, play central roles in plant energy metabolism, photosynthesis, and essential biochemical pathways (Gray et al., 1999; Neuhaus and Emes, 2000; Woodson and Chory, 2008). Plastid genomes are typically highly conserved, exhibiting a quadripartite structure composed of large single-copy (LSC) and small single-copy (SSC) regions separated by two inverted repeats (IRs), typically with stable gene content and limited gene loss or duplication events (Raubeson and Jansen, 2005; Wicke et al., 2011; Jansen and Ruhlman, 2012; Wang et al., 2024b). In contrast, plant mitochondrial genomes are highly dynamic, characterized by complex structures, frequent gene rearrangements, abundant repeat sequences, and incorporation of foreign DNA fragments (Wang et al., 2007; Park et al., 2015; Sun et al., 2022; Wu et al., 2022; Wang et al., 2024c). These features result in substantial interspecific variation in size, structure, and gene content of mitochondrial genome, shaped by complex selective pressures over the course of evolution.

Recent advances in high-throughput sequencing technologies have facilitated the assembly and annotation of numerous orchid organelle genomes, providing valuable data for comparative genomics and phylogenetic studies (Park et al., 2015; Kim et al., 2020; Li et al., 2024a; Shen et al., 2024; Wu et al., 2024; Soe et al.,

2025). However, most studies have focused either on plastid genomes or individual mitochondrial genomes, and few have systematically analyzed the evolutionary dynamics of both organelle genomes simultaneously. In particular, the roles of repeat expansion, plastid-to-mitochondrion DNA transfer (MTPTs), gene loss, and gene duplication in shaping genome size, structural diversity, and genetic variation across species remain largely unexplored (Cole et al., 2018; Yang et al., 2023; Wang et al., 2024a). A comprehensive investigation of these aspects is essential for understanding the evolution, functional maintenance, and ecological adaptation of organelle genomes in Orchidaceae.

Arundina graminifolia, a tropical orchid with high ornamental value (Ahmad et al., 2021; Li et al., 2025a), lacks complete organelle genome information. To address this gap, the present study assembled and annotated the mitochondrial and plastid genomes of *A. graminifolia* and conducted comparative analyses with 15 representative orchid species. Specifically, the study aimed to: (1) characterize the overall genome structure, repeat sequences, and plastid DNA transfer events to assess their contributions to genome expansion and structural complexity; (2) evaluate selection pressures (Ka/Ks) and nucleotide diversity (Pi) of conserved single-copy genes in mitochondrial and plastid genomes to reveal evolutionary patterns and functional constraints; and (3) reconstruct phylogenetic relationships using single-copy genes and assess genome synteny to explore structural evolution. Collectively, these analyses provide a genomic framework for understanding organelle genome evolution, functional adaptation, and phylogenetic relationships in Orchidaceae, offering valuable insights for further comparative genomics, molecular breeding, and ecological adaptation studies.

Materials and methods

Data source

The *A. graminifolia* sample used for sequencing was obtained from the greenhouse of the Environmental Horticulture Research Institute, Guangdong Academy of Agricultural Sciences. Whole-

genome HiFi sequencing was performed by Kindstar Sequenon Biotechnology (Wuhan) Co., Ltd, and the mitochondrial and plastid genomes were subsequently assembled by our team from the resulting unpublished data. Organelle genome sequences of 15 additional Orchidaceae species were downloaded from national center for biotechnology information (NCBI) (Supplementary Table S1). These sequences were used for subsequent analyses, including repeat identification, MTPTs, Ka/Ks and Pi analyses, phylogeny reconstruction, and synteny analyses.

A. *graminifolia* organelle genome assembly and annotation

The mitochondrial genome of *A. graminifolia* was assembled using two complementary approaches. The full set of whole-genome HiFi sequencing reads was processed using HiMT 1.1.1 (Tang et al., 2025) for organelle genome assembly from long-read data. A 30% random subset of the HiFi reads (~19 Gb) was subjected to assembly using Tippo 1.3.0 (Xian et al., 2025) for efficient reconstruction of organelle genomes with high fidelity. Candidate contigs were filtered with CD-HIT 4.8.1 (Li and Godzik, 2006; Fu et al., 2012) with default parameters and then validated against conserved mitochondrial genes and related genomes using BLAST 2.15.0+ (Madden, 2013) with an E-value cutoff of 1×10^{-5} . Contigs supported by both methods were merged and polished with Pilon 1.24 (Walker et al., 2014). Plastid assemblies were generated concurrently, with consistency verified through collinearity analysis.

A combined strategy of automated prediction using PMGA (<http://www.1kmpg.cn/pmga/>) (Li et al., 2025b) and manual curation in Apollo 1.11.8 (Lewis et al., 2002) was adopted for mitochondrial genome annotation with published mitochondrial genes and rRNAs as references. tRNAs were predicted using tRNAscan-SE 2.0 (Chan et al., 2021) with default parameters. plastid annotation was performed with CPGAVAS2 (<http://47.96.249.172:16019/analyzer/home>) (Shi et al., 2019) and manually verified in Apollo 1.11.8 (Lewis et al., 2002). Genome maps were visualized using ShinyCircos 2.0 (<https://venyao.xyz/shinycircos/>) (Wang et al., 2023) for the mitochondrial genome and OGDRAW 1.3.1 (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>) (Greiner et al., 2019) for the plastid genome.

Codon usage and nucleotide composition analysis

The sequences of protein-coding genes (PCGs) were extracted from the mitochondrial genome using PhyloSuite 2.0 (Zhang et al., 2020). Codon usage and nucleotide composition analyses were performed with CodonW 1.4.4 (Sharp and Li, 1986) under default settings to calculate GC content, relative synonymous codon usage (RSCU), and the effective number of codons (ENC). ENC-GC3 plots were generated using the R package ggplot2 (Wickham, 2011), with ENC values on the y-axis and GC content at the third codon position (GC3) on the x-axis. The expected ENC curve was

calculated according to previous research (Romero et al., 2000), using the formula:

$$ENC = 2 + GC3 + \frac{29}{(GC3^2 + (1 - GC3)^2)}$$

Genes located on or near the expected curve indicate that codon usage bias is mainly shaped by mutational pressure, whereas genes below the curve suggest that natural selection plays a dominant role (Sueoka, 1988). Additionally, the scatter plots of GC content were generated using ggplot2 (Wickham, 2011), with the mean GC content at the first and second codon positions (GC12) on the y-axis and GC3 on the x-axis.

Repeat sequence and MTPTs identification

Repeat sequences in *A. graminifolia* and other 15 Orchidaceae mitochondrial genomes were identified by self-alignment using BLAST 2.15.0+ (word size = 7, E-value $\leq 1 \times 10^{-5}$) (Madden, 2013). The repeat sequences with length ≥ 30 bp and sequence similarity $\geq 80\%$ were retained, and classified into four retained and classified into five categories: Class I ($30 \leq L \leq 100$ bp), Class II ($100 < L \leq 200$ bp), Class III ($200 < L \leq 500$ bp), Class IV ($500 < L \leq 1000$ bp), and Class V (> 1000 bp). Simple sequence repeats (SSRs) were detected using MISA (<https://webblast.ipk-gatersleben.de/misa/index.php?action=1>) (Beier et al., 2017), with minimum repeat thresholds of 10, 8, 5, 4, 3, and 3 for mono-, di-, tri-, tetra-, penta-, and hexanucleotide motifs, respectively. For each species, the total number and cumulative length of repeats were calculated, excluding shorter repeats nested within longer ones. Spearman correlations between repeat abundance, cumulative repeat length, and mitochondrial genome size were analyzed and visualized using ggplot2 (Wickham, 2011).

MTPTs were identified by aligning the high-quality genome of plastid of each species against its corresponding mitochondrial assemblies using BLASTN 2.15.0+ (Madden, 2013) with the threshold: E-value $\leq 1 \times 10^{-5}$, alignment length ≥ 250 bp, sequence identity $\geq 80\%$, and word size = 7. Candidate MTPTs were filtered to remove redundancy by merging overlapping or adjacent fragments. Each retained MTPT was annotated to determine genomic position, length, sequence similarity, and whether it contained full-length or partial plastid coding genes.

Evolutionary and comparative analyses

The rates of nonsynonymous (Ka) and synonymous (Ks) substitutions in mitochondrial genomes were estimated using single-copy orthologs from *A. graminifolia* and other 15 Orchidaceae species. Homologous sequences were aligned with ParaAT 2.0 (Zhang et al., 2012) with default parameters, and Ka/Ks values were calculated using the 'Simple Ka/Ks Calculator' in TBtools-II (Chen et al., 2023). Nucleotide diversity (Pi) in plastid genomes was assessed from single-copy orthologous genes aligned with MAFFT 7.429 (Katoh and Standley, 2013) under default

settings, and Pi values were calculated using DnaSP 6.0 (Rozas et al., 2017) with a sliding window of 100 bp and step size of 50 bp. Results were visualized in R using ggplot2 (Wickham, 2011).

Phylogenetic relationships were inferred using 11 mitochondrial and 8 plastid single-copy genes with *Allium cepa* as the outgroup. Nucleotide sequences were translated to amino acids, aligned with MAFFT 7.429 (Kato and Standley, 2013), and back-translated using pal2nal.pl (Suyama et al., 2006). Low-quality regions were trimmed with TrimAl (Capella-Gutiérrez et al., 2009), and sequences were concatenated into a supermatrix with Python script. Maximum likelihood phylogenies were generated using IQ-TREE 3.0.1 (Wong et al., 2025), with ultrafast bootstrap method to assess nodal support. Amino acid-based phylogenies were generated following the same workflow above.

Synten among organelle genomes was analyzed using 16 mitochondrial genomes. Whole-genome alignments were performed with nucmer (–maxmatch), filtered with delta-filter for one-to-one matches, and coordinates were extracted with showcoords (Marçais et al., 2018). Genome synteny link files were generated using GetTwoGenomeSyn.pl and visualized in NGenomeSyn 1.43 (He et al., 2023), enabling systematic comparison of structural conservation and rearrangements across Orchidaceae mitochondrial genomes.

Results

Organellar genome features of *A. graminifolia*

To characterize the organelle genomes of *A. graminifolia*, we assembled and annotated its mitochondrial and plastid genomes using two independent *de novo* assemblers, HiMT and Tippo. The mitochondrial genome assembly with HiMT produced 20 circular molecules, whereas the Tippo assembly yielded 20 circular molecules and 3 linear fragments (Figures 1A, B). Synteny analysis revealed that edge_2166 in HiMT corresponded to the concatenation of edge_176 and edge_158 in Tippo, while edge_21 and edge_157 from Tippo lacked corresponding sequences in HiMT (Figure 1C, Supplementary Table S2). The fragments edge_21 and edge_157, which contained no identifiable coding genes, showed markedly lower sequencing support compared with other mitochondrial contigs, and were thus excluded from downstream analyses. Overall, mitochondrial assemblies were highly consistent when using different methods, and the plastid genome assemblies were identical. Consequently, the HiMT assembly was adopted as the reference for subsequent analyses.

The mitochondrial genome displayed a multicircular structure with a total length of 532,569 bp and a GC content of 44.0% (Figure 1D, Supplementary Table S3). It encoded 27 core protein-coding genes, 13 variable protein-coding genes, 22 tRNA genes, and 3 rRNA genes, including ATP synthase (*atp1–atp9*), cytochrome *c* biogenesis (*ccmB*, *ccmC*, *ccmFC*, *ccmFN*), cytochrome *b* (*cob*), cytochrome *c* oxidase (*cox1–cox3*), NADH dehydrogenase (*nad1–nad9*), succinate dehydrogenase (*sdh4*), maturase (*matR*), and

membrane transport protein (*mttB*) genes (Table 1). Several genes contained 1–4 introns, including *ccmFC*, *rps3*, *rps10*, *cox2*, *nad1*, *nad2*, *nad4*, *nad5*, and *nad7*.

The plastid genome exhibited the typical quadripartite structure, consisting of LSC, SSC, and a pair of IRs, with a total length of 159,503 bp and a GC content of 37.10% (Figure 1E). The genome contained 81 protein-coding genes, 27 tRNA genes, and 8 rRNA genes. Several of these genes are duplicated in the IR regions, including *ndhB*, *rpl2*, *rpl23*, *rps7*, *rps12*, and *rps19*. Nine genes contained a single intron, and three genes contained two introns (Figure 1F, Table 2). These high-quality assemblies provide a foundation for comparative and evolutionary analyses in Orchidaceae.

Codon usage in *A. graminifolia* mitochondrial genome

Codon usage analysis of the 40 mitochondrial protein-coding genes yielded 10,213 codons. Leucine (Leu) was the most abundant amino acid (10.58%), followed by serine (Ser, 8.38%), while the stop codons were the least frequent (0.13%) (Supplementary Table S4). RSCU analysis indicated that the most preferred codons ended with A/U, with strong preferences for Ala (GCU, RSCU = 1.97), Ser (UCU, RSCU = 1.73), and Arg (GCU, RSCU = 1.53) (Figure 2A, Supplementary Table S4).

GC content analysis showed uneven distribution across codon positions: GC1 (21.57%), GC2 (22.13%), GC3 (34.86%), and overall GC (43.47%) (Supplementary Table S5). ENC values ranged from 32.22 to 61, with an average of 46.92, and the two *atp9* copies displayed the lowest (32.22) and second lowest ENC values, reflecting strong codon usage bias. ENC scatterplot analysis suggested that natural selection, rather than mutational pressure, primarily shapes codon usage (Figure 2B), further supported by a weak correlation between GC3 and GC12 (slope = 0.12, $R^2 = 0.031$, Figure 2C).

Gene loss and duplications in Orchidaceae organelle genomes

To explore organellar gene evolution in Orchidaceae, we examined mitochondrial genomes from 16 species together with plastid genomes from 14 species (plastomes were not available for two taxa), and assessed patterns of gene loss and multi-copy gene distribution (Supplementary Table S1). In the mitochondrial genomes, most protein-coding genes were highly conserved, particularly those associated with energy metabolism and electron transport, including ATP synthase genes (*atp1*, *atp4*, *atp6*, *atp8*, and *atp9*), cytochrome *c* biogenesis genes (*ccmB*, *ccmC*, *ccmFC*, and *ccmFN*), the maturase gene (*matR*), and cytochrome *b* (*cob*) (Figure 3A). In contrast, ribosomal protein genes and succinate dehydrogenase subunit genes displayed greater variability, with *rpl2*, *rps1*, *rps2*, *rps3*, *rps4*, *rps7*, *rps10*, *rps11*, and *sdh4* absent in some species. Notably, *rps11* was missing in *A. graminifolia*, *cox2*

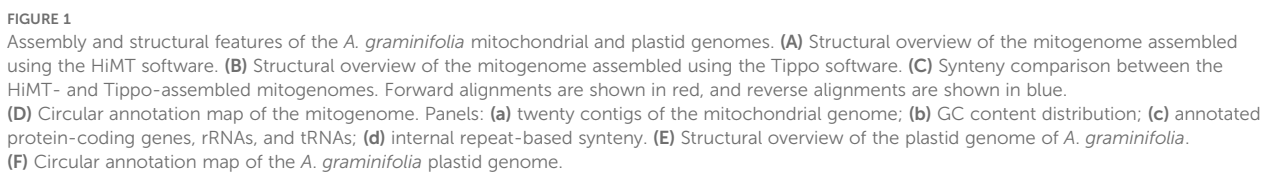


TABLE 1 Gene composition of the *A. graminifolia* mitochondrial genome.

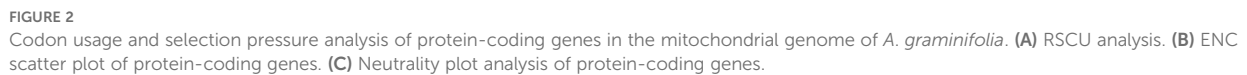
Gene category	Functional group	Name of genes
Core genes	ATP synthase	<i>atp1</i> (2), <i>atp4</i> , <i>atp6</i> , <i>atp8</i> , <i>atp9</i> (2)
	Cytochrome c biogenesis	<i>ccmB</i> , <i>ccmC</i> , <i>ccmFC*</i> , <i>ccmFN</i>
	Ubiquinol cytochrome c reductase	<i>cob</i>
	Cytochrome c oxidase	<i>cox1</i> , <i>cox2**</i> , <i>cox3</i>
	Maturases	<i>matR</i>
	Transport membrane protein	<i>mttB</i>
	NADH dehydrogenase	<i>nad1****</i> , <i>nad2****</i> , <i>nad3</i> , <i>nad4***</i> , <i>nad4L</i> , <i>nad5*</i> , <i>nad6</i> , <i>nad7***</i> , <i>nad9</i>
	Subunit of succinate dehydrogenase	<i>sdh4</i>
Variable genes	Ribosomal proteins (LSU)	<i>rpl2</i> , <i>rpl5</i> , <i>rpl16</i>
	Ribosomal proteins (SSU)	<i>rps1</i> , <i>rps2</i> , <i>rps3*</i> , <i>rps4</i> , <i>rps7</i> , <i>rps10*</i> , <i>rps12</i> , <i>rps13</i> , <i>rps14</i> , <i>rps19</i>
rRNA	Ribosomal RNAs	<i>rrn18</i> , <i>rrn26</i> , <i>rrn5</i>
tRNA	Transfer RNAs	<i>trnE-UUC*</i> (2), <i>trnM-CAU*</i> (3), <i>trnT-GGU*</i> , <i>trnC-GCA</i> (2), <i>trnD-GUC</i> , <i>trnF-GAA</i> (2), <i>trnH-GUG</i> (2), <i>trnK-UUU</i> , <i>trnL-UAG</i> , <i>trnI-CAU</i> , <i>trnJ-M-CAU</i> , <i>trnN-GUU</i> , <i>trnQ-UUG</i> , <i>trnW-CCA</i> , <i>trnY-GUA</i> , <i>trnA-UGC*</i>

Asterisks (*) indicate the number of introns, and numbers in parentheses indicate the copy number of multi-copy genes.

TABLE 2 Gene composition of the *A. graminifolia* plastid genome.

Gene category	Functional group	Name of genes
Photosynthesis	Subunits of ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF*</i> , <i>atpH</i> , <i>atpI</i>
	Subunits of photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	Subunits of photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
	Subunits of NADH-dehydrogenase	<i>ndhA*</i> , <i>ndhB*</i> (2), <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
	Subunits of cytochrome b/f complex	<i>petA</i> , <i>petB*</i> , <i>petD*</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	Subunits of rubisco	<i>rbcL</i>
Self-replication	Large subunit of ribosome	<i>rpl2*</i> (2), <i>rpl14</i> , <i>rpl16*</i> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (2), <i>rpl33</i> , <i>rpl36</i>
	DNA dependent RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1*</i> , <i>rpoC2</i>
	Small subunit of ribosome	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (2), <i>rps8</i> , <i>rps11</i> , <i>rps12**</i> (2), <i>rps14</i> , <i>rps15</i> , <i>rps16**</i> , <i>rps18</i> , <i>rps19</i> (2)
	Ribosomal RNAs	<i>rrn4.5S</i> (2), <i>rrn5S</i> (2), <i>rrn16S</i> (2), <i>rrn23S</i> (2)
	Transfer RNAs	<i>trnA-UGC*</i> (2), <i>trnC-GCA</i> , <i>trnD-GUC</i> , <i>trnE-UUC*</i> (3), <i>trnF-GAA</i> , <i>trnG-GCC</i> , <i>trnH-GUG</i> (2), <i>trnL-CAA</i> (2), <i>trnL-UAA*</i> , <i>trnL-UAG</i> , <i>trnM-CAU</i> (4), <i>trnN-GUU</i> (2), <i>trnP-UGG</i> , <i>trnQ-UUG</i> , <i>trnR-ACG</i> (2), <i>trnR-UCU</i> , <i>trnS-CGA*</i> , <i>trnS-GCU</i> , <i>trnS-GGA</i> , <i>trnS-UGA</i> , <i>trnT-GGU</i> , <i>trnT-UGU</i> , <i>trnV-GAC</i> (2), <i>trnW-CCA</i> , <i>trnY-GUA</i>
	c-type cytochrome synthesis gene	<i>ccsA</i>
	Envelope membrane protein	<i>cemA</i>
	Translational initiation factor	<i>infA</i>
	Protease	<i>clpP**</i>
	Maturase	<i>matK</i>
Unkown	Conserved open reading frames	<i>ycf3**</i> , <i>ycf4</i>

Asterisks (*) indicate the number of introns, and numbers in parentheses indicate the copy number of multi-copy genes.



Plastid genomes showed greater variability than mitogenomes across selected orchid species, with total gene numbers ranging from 19 to 89 (Figure 3B). *B. striata* contained the highest number of plastid genes, including nine duplicated genes (*ndhB*, *rpl2*, *rpl23*, *rps7*, *rps12*, *rps19*, *ycf1*, *ycf2*, and *ycf15*), whereas *Epipogium aphyllum* and *Gastrodia* species had the fewest genes. Across all 14 analyzed species, *clpP*, *rpl14*, *rpl36*, *rps2*, *rps4*, *rps8*, *rps11*, *rps14*, and *rps18* were consistently retained, indicating their high conservation. In contrast, lineage-specific losses and gains were observed for some genes. For instance, *accD* was absent only in *A. graminifolia*, *rpl16*

Repetitive sequence analysis of Orchidaceae mitogenomes

A total of 1,156 SSRs were identified from 16 orchid mitogenomes (Supplementary Table S6), with mononucleotide repeats being the most abundant (66.1%), followed by pentanucleotide repeats (18.3%), and compound SSRs (2.2%) (Figure 4A). Among those Orchid species, *Nervilia mackinnonii* possessed the highest number of SSRs, followed by *G. angusta* (Figure 4B, Supplementary Table S6). Most species contained seven SSR types, such as *Apostasia shenzhenica*, *B. striata*, *G. angusta*, and *G. pubilabiata*, whereas *C. ensifolium*, *D. amplum*, *D. hancockii*, and *P. micranthum* had only four types (Figure 4B, Supplementary

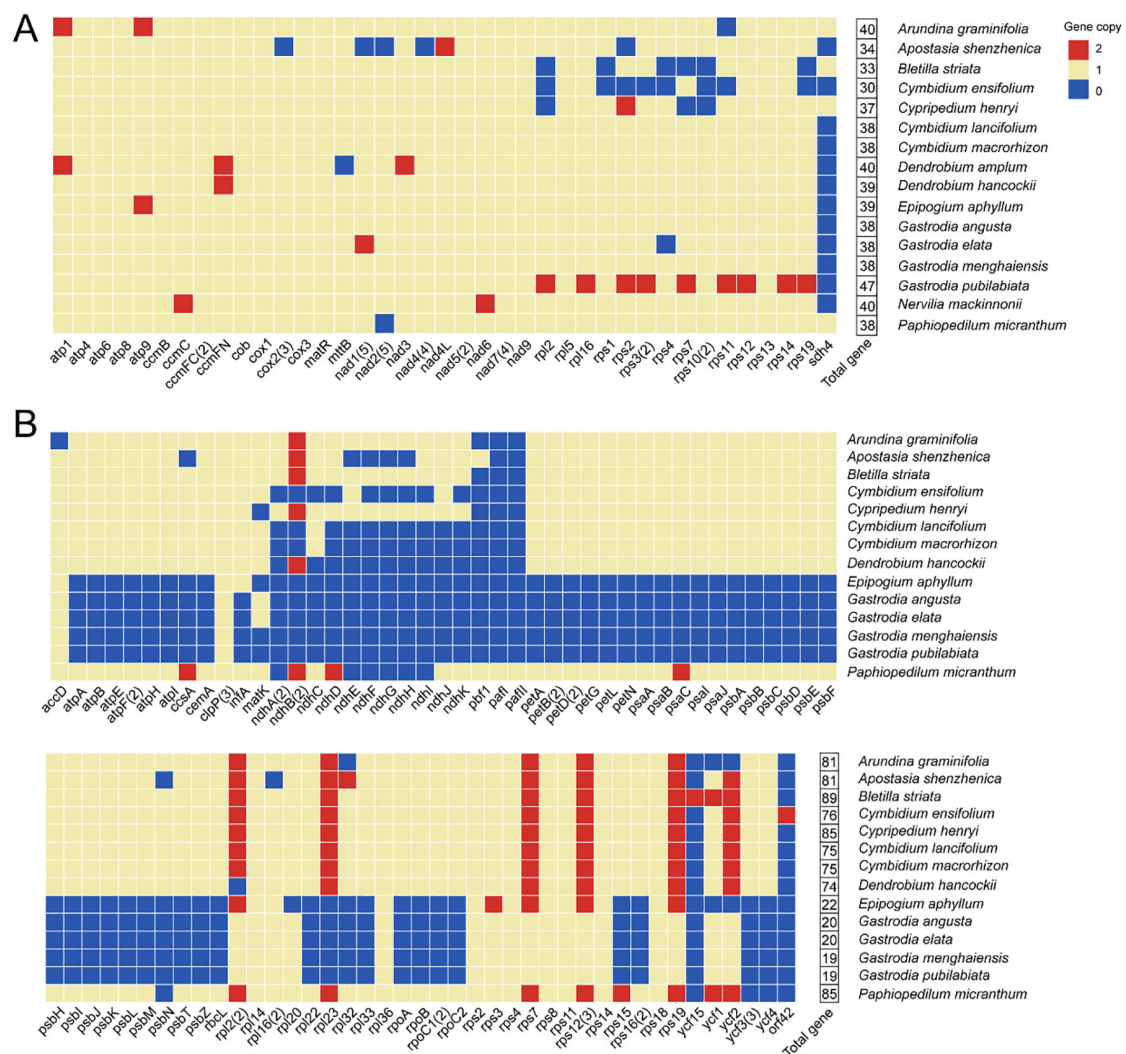


FIGURE 3

Distribution of protein-coding genes in Orchidaceae organelle genomes. Distribution of protein-coding genes in Orchidaceae mitochondrial (A) and plastid (B) genomes; numbers in parentheses indicate exon counts.

Table S6). In *A. graminifolia*, 47 mononucleotide, 2 dinucleotide, 3 trinucleotide, 5 pentanucleotide, and 1 compound SSR were detected.

Mitochondrial repeats were further classified by length into five categories: Class I (30–100 bp), Class II (100–200 bp), Class III (200–500 bp), Class IV (500–1000 bp), and Class V (> 1000 bp). Across the 16 mitochondrial genomes (416,775–1,339,825 bp), 16,539 repeats were identified with a cumulative length of 1,252,601 bp (Supplementary Tables S7, S8). Class I repeats were the most numerous (90.4%) and accounted for the largest proportion of repeat length (55.7%), followed by Class V (18.0%) (Figure 4C). Although Class I repeats dominated in most species, Class V contributed disproportionately to total repeat length in *D. hancockii*, *D. amplum*, *Cyrtopodium henryi*, and *G. pubilabiata* (Figure 4D). Some repeat classes were absent in specific species:

Class IV repeats were missing in *Apostasia shenzhenica* and *Cyrtopodium henryi*, and Class V repeats were absent in *A. graminifolia*, *Apostasia shenzhenica*, *C. ensifolium*, *E. aphyllum*, *G. menghaiensis*, and *P. micranthum*. Notably, the mitochondrial genome of *G. pubilabiata* (898,047 bp) contained 65 Class V repeats, far exceeding all other species (< 10), highlighting species-specific expansion of long repeats.

Correlation analysis revealed that mitochondrial genome size in Orchidaceae was strongly positively correlated with the number of Class I–III repeats ($R > 0.7$, $p < 0.001$), and also with the total lengths of Class I and II repeats ($R > 0.7$, $p < 0.001$). Genome size was additionally positively correlated with Class III repeat length ($R > 0.5$, $p < 0.05$) (Figures 4E, F). These results highlight the role of repeats in the expansion of the mitochondrial genome and its structural complexity.



FIGURE 4

Repeat sequence analysis in mitochondrial genomes of Orchidaceae. **(A)** Overview of SSR types in Orchidaceae mitochondrial genomes. **(B)** SSR types and numbers in each Orchidaceae mitochondrial genome; p1–p6 denote mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats, respectively, while c represents compound SSRs. **(C)** Overview of repeat sequence types in Orchidaceae mitochondrial genomes. **(D)** Repeat sequence types and numbers in each Orchidaceae mitochondrial genome. **(E, F)** Spearman correlations between repeat abundance or total repeat length and mitochondrial genome size. Species abbreviations used in panels **(B, D)** Agr (*Arundina graminifolia*), Ash (*Apostasia shenzhenica*), Bst (*Bletilla striata*), Cen (*Cymbidium ensifolium*), Che (*Cypripedium henryi*), Cla (*Cymbidium lancifolium*), Cma (*Cymbidium macrorhizon*), Dam (*Dendrobium amplum*), Dha (*Dendrobium hancockii*), Epa (*Epipogium aphyllum*), Gan (*Gastrodia angusta*), Gel (*Gastrodia elata*), Gme (*Gastrodia menghaiensis*), Gpu (*Gastrodia pubilabiata*), Nma (*Nervilia mackinnonii*), Pmi (*Paphiopedilum micranthum*).

Plastid-derived DNA fragments in Orchidaceae mitogenomes

MTPTs varied widely among 14 orchid species, ranging from 0 in *G. menghaiensis* to 65 in *C. lancifolium* (Figure 5A, Supplementary Table S9). Analysis of MTPTs length revealed species-specific distribution patterns. Short fragments (0–500 bp)

were most abundant in *Cymbidium* species and *D. hancockii*, whereas *P. micranthum* primarily contained 1000–2000 bp fragments. In *C. ensifolium* and *B. striata*, the fragments of 2000–5000 bp were predominant. MTPTs longer than 5000 bp were detected in *C. ensifolium*, *B. striata*, other *Cymbidium* species, and *G. pubilabiata*. The total MTPTs length also varied among species: *B. striata* had the greatest cumulative length, followed by *C.*

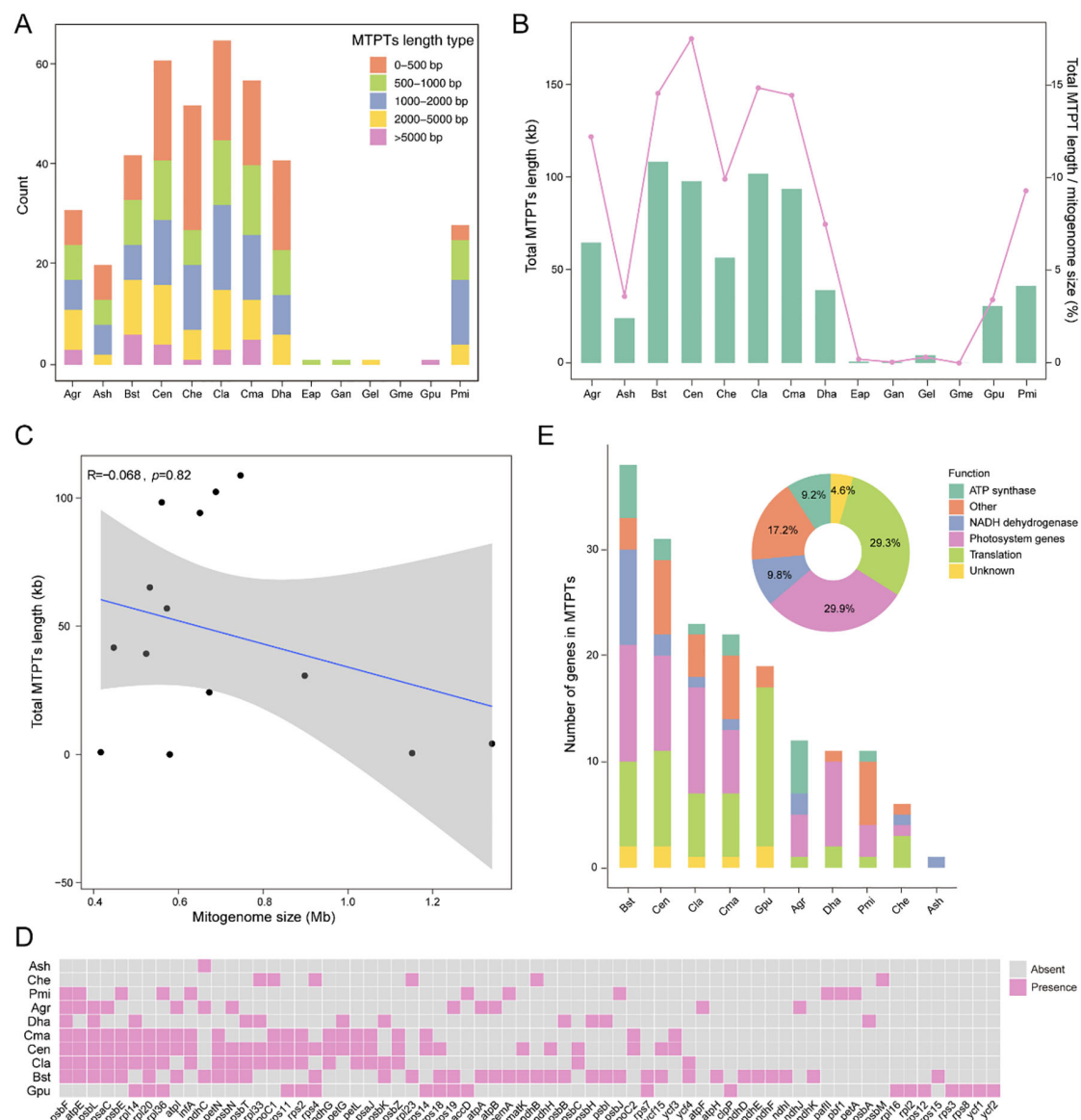


FIGURE 5

Analysis of MTPTs in Orchidaceae mitochondrial genomes. (A) Number of MTPTs of different lengths in Orchidaceae mitochondrial genomes. (B) Total length of MTPTs (left) and their ratio to mitochondrial genome size (right). (C) Spearman correlation between total MTPT length and mitochondrial genome size. (D) Distribution of plastid protein-coding genes within MTPTs. (E) Functional classification of plastid protein-coding genes within MTPTs. Pie charts show the overall functional composition, and bar charts display the functional distribution for each mitochondrial genome. Species abbreviations used in panels (A, B, D) Agr (*Arundina graminifolia*), Ash (*Apostasia shenzhenica*), Bst (*Bletilla striata*), Cen (*Cymbidium ensifolium*), Che (*Cypripedium henryi*), Cla (*Cymbidium lancifolium*), Cma (*Cymbidium macrorhizon*), Dha (*Dendrobium hancockii*), Epa (*Epipogium aphyllum*), Gan (*Gastrodia angusta*), Gel (*Gastrodia elata*), Gme (*Gastrodia menghaiensis*), Gpu (*Gastrodia pubilabiata*), Pmi (*Paphiopedilum micranthum*).

lancifolium, whereas *G. angusta* had the shortest total length (excluding *G. menghaiensis*, which lacked MTPTs) (Figure 5B). The proportion of MTPTs in mitochondrial genomes ranged from 0 to 17.53%, with *C. ensifolium* exhibiting the highest fraction. No significant correlation was observed between total MTPTs length and mitochondrial genome size ($r = -0.068$, $p = 0.82$) (Figure 5C).

Functional assessment showed that some species contained 1–38 intact plastid genes within MTPTs, most frequently *psbF*, *atpE*, and *psbL* (Figures 5D, E). Photosystem-related genes were the most abundant (29.9%), followed by translation-related genes (29.3%) (Figures 5D, E). These results highlight the dynamic patterns of organelle DNA transfer and retention in orchid mitochondria.

Selection pressure and nucleotide diversity in Orchidaceae organelle genomes

Selection pressure (Ka/Ks) and nucleotide diversity (Pi) were assessed for conserved organelle protein-coding genes. In mitochondrial genomes, 14 shared single-copy genes were examined (Figure 6A, Supplementary Table S10). Among these genes, *nad9* showed the highest Ka/Ks value, followed by *ccmB*, while *cox1* had the lowest. All genes exhibited Ka/Ks < 1, indicating strong purifying selection. Nucleotide diversity was similarly low, with *atp4* displaying the highest Pi value (0.0427), followed by *atp8* (0.0404), whereas *cox1* presented the lowest Pi value (0.0115).

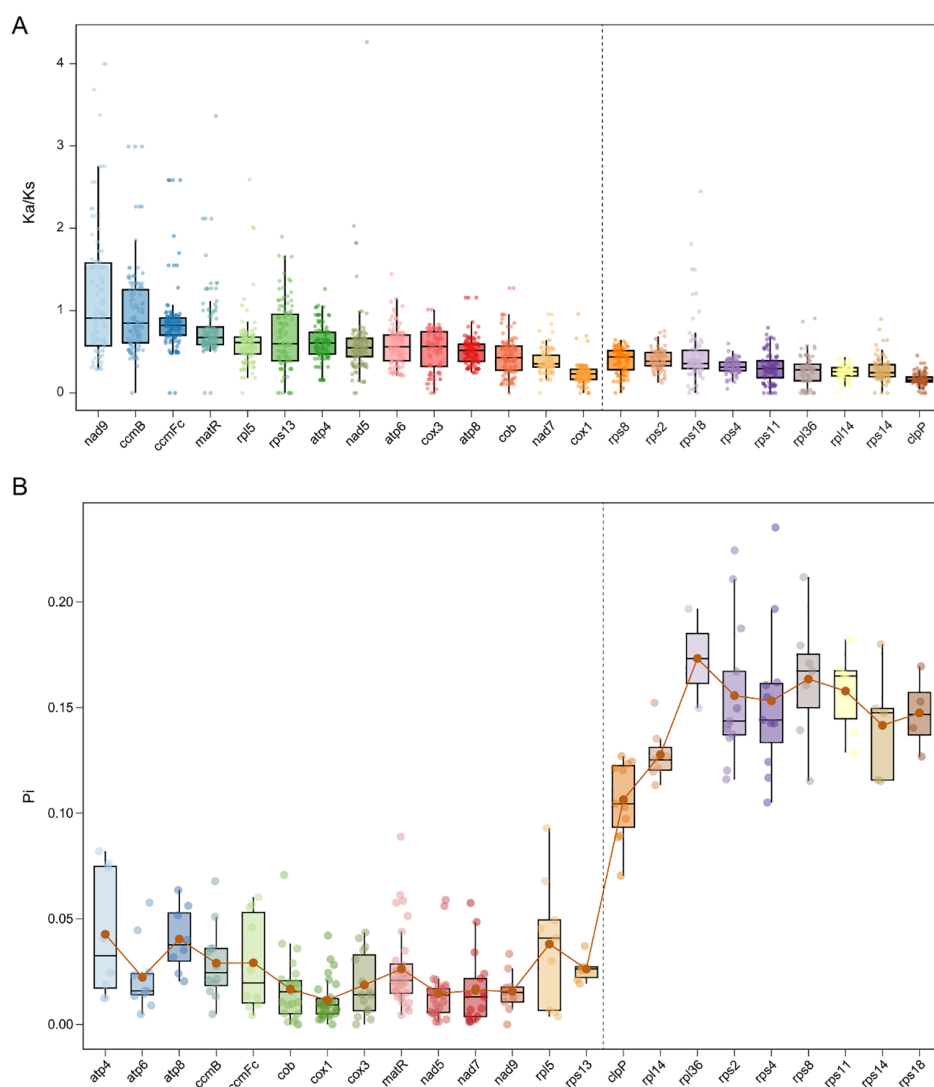


FIGURE 6

Evolutionary analysis of single-copy protein-coding genes shared between mitochondrial and plastid genomes of Orchidaceae. (A) Ka/Ks analysis of shared single-copy protein-coding genes. (B) Nucleotide diversity (Pi) analysis of the same genes. The dashed line separates mitochondrial genes (on the left) and plastid genes (on the right).

(Figure 6B, Supplementary Table S11). Additionally, all P_i values were below 0.05, indicating high sequence conservation in mitochondrial genomes.

In plastid genomes, nine shared single-copy genes were identified. Gene *rps8* displayed the highest K_a/K_s value, followed by *rps2*, whereas *clpP* had the lowest (Figure 6A). Although all genes exhibited $K_a/K_s < 1$, nucleotide diversity was markedly higher than in mitochondrial genes. The genes *rpl36* (0.1733) and *rps8* (0.1635) showed the greatest variability, while *clpP* had the lowest P_i value (0.1064) (Figure 6B). Overall, mitochondrial genes exhibit stronger purifying selection and greater sequence conservation, whereas plastid genes display substantially higher nucleotide diversity and faster evolutionary rates across Orchidaceae species.

Phylogenetic relationships among Orchidaceae species

Phylogenetic relationships among Orchidaceae species were inferred using shared single-copy genes from mitochondrial and plastid genomes. For mitochondrial genomes, 14 conserved single-copy genes from 17 species were analyzed at both nucleotide and protein levels (Figure 7A). The resulting trees showed highly consistent topologies with strong branch support. Basal lineages such as *Apostasia shenzhenica* and *P. micranthum* occupied stable

positions, and *A. graminifolia* consistently clustered with *B. striata*, indicating a close evolutionary relationship.

Plastid phylogenies were constructed using nine shared single-copy genes from 15 species (Figure 7B). Nucleotide- and protein-based trees again produced congruent topologies, with most nodes receiving strong support. Major lineages occupied the same positions across trees, and *A. graminifolia* remained clustered with *B. striata*, consistent with the mitochondrial phylogeny. Overall, single-copy genes from both the mitochondrial and plastid genomes provide stable and reliable phylogenetic signals in Orchidaceae. The strong congruence between organelle-based phylogenies supports their effectiveness for resolving evolutionary relationships among orchid species.

Comparative synteny of Orchidaceae mitogenomes

Structural conservation of mitochondrial genomes across Orchidaceae was assessed through synteny analysis of 16 species (Figure 8). Overall, mitochondrial genomes exhibited low synteny, with highly fragmented homologous blocks, indicating extensive genomic rearrangements across the family. Within the *Cymbidium* clade, synteny was comparatively higher than in other species, with several species sharing long, continuous homologous regions,

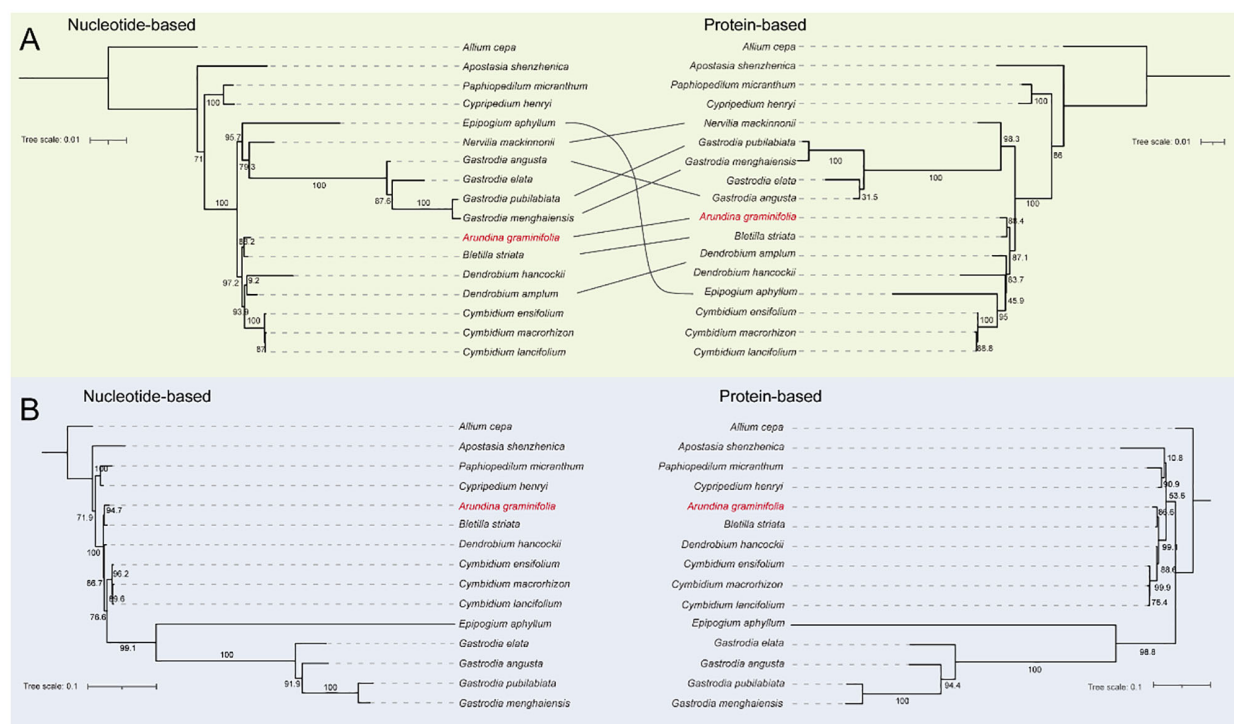


FIGURE 7

Phylogenetic analysis of Orchidaceae based on single-copy orthologs in mitochondrial and plastid genomes, with *Allium cepa* included as the outgroup. Maximum likelihood phylogenies of Orchidaceae species were inferred using nucleotide and protein sequences of single-copy orthologs in (A) mitochondrial genomes (17 species) and (B) plastid genomes (15 species).

suggesting a more conserved mitochondrial genome structure within this genus. However, noticeable block fragmentation and rearrangements were still presented, reflecting ongoing structural reorganization even among closely related species.

In contrast, mycoheterotrophic orchids such as *Gastrodia* species and *E. aphyllum* exhibited extremely low synteny with other taxa. Only a few short and discontinuous homologous blocks were detected, indicating that these lineages have undergone more extensive and complex mitochondrial genome restructuring. These patterns highlight substantial variation in mitochondrial genome architecture among Orchidaceae lineages, demonstrating that large-scale rearrangements have been pervasive throughout their evolutionary history.

Discussion

Robustness and reliability of organelle genome assemblies in *A. graminifolia*

The assembly of *A. graminifolia* organelle genomes using two independent *de novo* assemblers demonstrated overall high consistency, highlighting the robustness of the mitochondrial and plastid genome reconstructions. Minor differences were observed between the HiMT (Tang et al., 2025) and Tiplo (Xian et al., 2025) assemblies, such as the presence of small linear fragments in Tiplo and unaligned edges. This difference likely resulted from assembly artifacts rather than biologically meaningful variation, as these

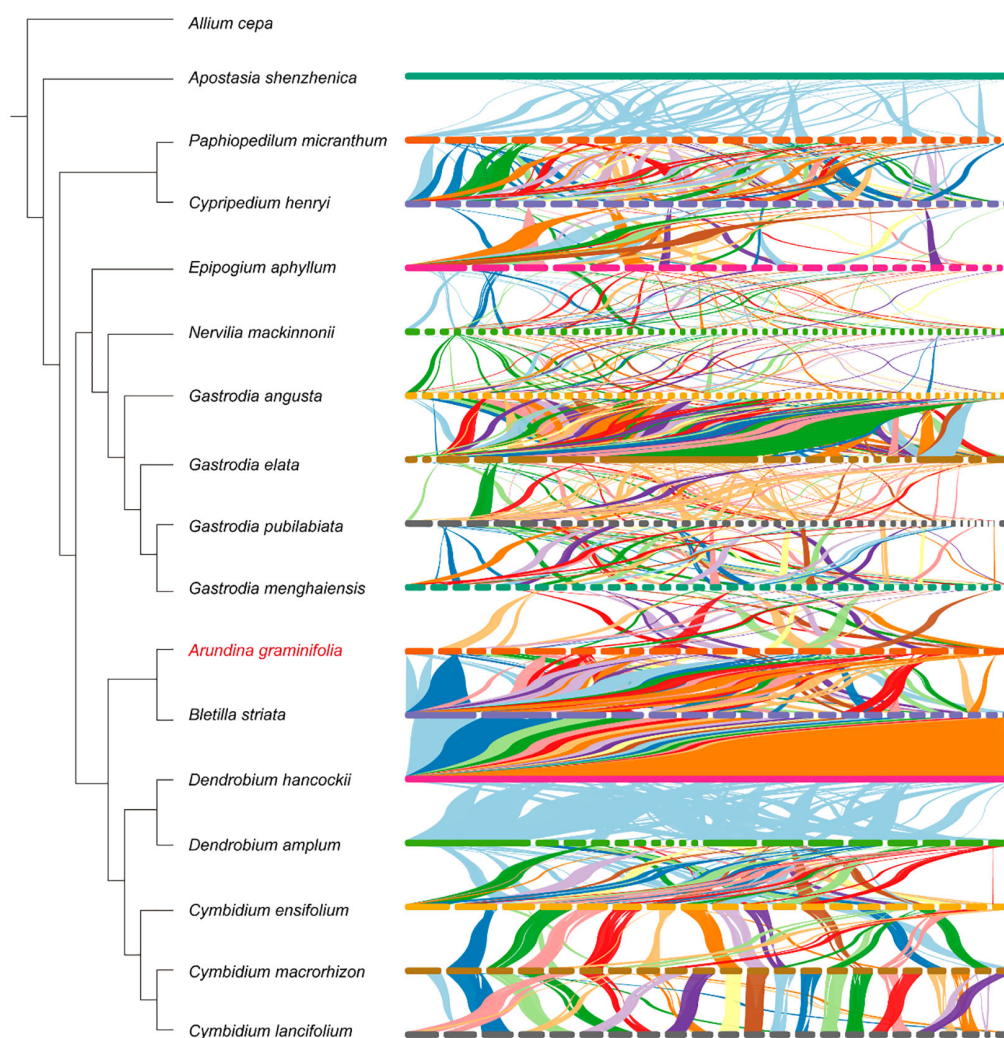


FIGURE 8

Synteny analysis of 16 Orchidaceae mitochondrial genomes. The synteny map highlights conserved and rearranged regions among the 16 orchid mitogenomes. Note that genomes and contigs are normalized to a uniform length for visualization purposes and are not to scale. Due to the integration of multiple species, some overlapping syntenic ribbons may appear clustered; for a detailed, deconstructed visualization of the homologous blocks between *B. striata* and *D. hancockii*, please refer to Supplementary Figure S1.

fragments lacked identifiable coding sequences and showed low sequencing support. Because both assemblers produced highly concordant and stable structural configurations, we did not further apply additional tools such as PMAT2 (Han et al., 2025) or Oatk (Zhou et al., 2025) for fine-scale assembly refinement. The strong agreement between HiMT and TipPO supports the reliability of the mitochondrial genome structure and gene content reported here. Furthermore, the identical plastid genome assemblies suggest a conserved plastome organization in *A. graminifolia*, reinforcing the suitability of the HiMT-based assembly as the reference for downstream comparative and evolutionary analyses.

Contrasting evolutionary constraints of mitochondrial and plastid genomes

The mitochondrial genome of *A. graminifolia* exhibits a multichromosomal structure with extensive rearrangements, consistently across the selected Orchidaceae species (Yuan et al., 2018; Li et al., 2023; Yang et al., 2023; Shen et al., 2024; Soe et al., 2025; Wang et al., 2025). Highly fragmented homologous blocks indicate frequent recombination events and large-scale structural evolution. Notably, mycoheterotrophic lineages, including *Gastrodia* genus and *E. aphyllum*, display extremely low synteny with other taxa, suggesting lineage-specific expansions and complex restructuring of mitochondrial genomes. In contrast, plastid genomes of photosynthetic orchids tend to be conserved in overall architecture and gene content, although heterotrophic lineages often exhibit extensive genomic reduction (Raubeson and Jansen, 2005; Peng et al., 2022; Wen et al., 2022; Zhao et al., 2024). This stark difference reflects distinct evolutionary constraints that mitochondrial genomes are more prone to structural rearrangements, whereas plastid genomes are constrained by the functional requirements of photosynthesis and gene expression. The observed pattern aligns with previous findings in angiosperms, where plastid genomes evolve slowly in structure compared with recombination-active mitochondrial genomes (Sloan et al., 2018; Xu et al., 2021; Peng et al., 2022). Furthermore, the distribution of conserved versus variable genes supports the notion that functional essentiality contributes to evolutionary stability: genes involved in energy metabolism, such as ATP synthases and cytochrome c biogenesis genes, remain highly conserved in mitochondrial genome, whereas ribosomal protein genes show lineage-specific variation, reflecting relaxed constraints or potential functional redundancy (Gray et al., 1999; Maier et al., 2013; Skippington et al., 2017).

Repetitive sequences and plastid DNA transfers as drivers of mitochondrial genome expansion

Repetitive elements play a crucial role in determining the size and structural complexity of the mitochondrial genome (Ouyang et al., 2025; Su et al., 2025). The analysis of SSRs and long repeats

across 16 orchid species revealed substantial interspecific variation in number, length, and distribution, with the repeats of Class I being the most abundant. Species such as *G. pubilabiata* exhibit notable expansion of long repeats, which disproportionately contributes to genome size and likely facilitates homologous recombination and structural rearrangements. Additionally, MTPTs exhibit remarkable diversity among species, ranging from 0 to 65 fragments, with functional enrichment in genes related to photosynthesis and translation (Wang et al., 2012; Nhat Nam et al., 2024). The combined presence of repeats and MTPTs suggests a synergistic role in promoting genome expansion and rearrangement, which may provide a selective advantage by increasing genomic plasticity and buffering against deleterious mutations. The variation in repeat content and MTPTs abundance across species also indicates that mitochondrial genome evolution is influenced by both lineage-specific and ecological influences. For instance, mycoheterotrophic orchids typically contain fewer MTPTs but exhibit remarkable structural rearrangements, emphasizing the interplay between genome architecture and their life history traits (Zhao et al., 2024; Wang et al., 2025).

Purifying selection, genetic diversity, and phylogenetic utility of single-copy genes

Selection analyses reveal that Orchidaceae mitochondrial single-copy genes are under strong purifying selection ($Ka/Ks < 1$) with low nucleotide diversity ($\pi < 0.05$), reflecting functional constraints and evolutionary stability. Conversely, Orchidaceae plastid single-copy genes, although under purifying selection, tend to show greater sequence variability compared with other regions, suggesting faster evolutionary rates in these loci (Jansen et al., 2007; Cole et al., 2018; Jiang et al., 2024). Phylogenetic reconstruction using single-copy genes generated congruent topologies for mitochondrial and plastid genomes, with *A. graminifolia* consistently clustering with *B. striata*, supporting their close evolutionary relationship (Li et al., 2019). Low synteny and fragmented homologous blocks in mitochondrial genomes, particularly in mycoheterotrophic orchids, underscore the dynamic nature of structural evolution (Lee et al., 2020; Zhao et al., 2024; Wang et al., 2025). Gene loss and duplication contribute to functional redundancy, potentially enhancing environmental adaptability (Costello et al., 2020; Li et al., 2024b). Collectively, these findings indicate that single-copy genes serve as reliable molecular markers for resolving Orchidaceae phylogeny (Li et al., 2019), whereas mitochondrial genome architecture reflects lineage-specific structural plasticity driven by both ecological adaptation and intrinsic genome dynamics (Wang et al., 2024a; Zheng et al., 2024).

Conclusions

Orchidaceae organelle genomes exhibit contrasting evolutionary patterns. Mitochondrial genomes are structurally

complex, multichromosomal, and under strong purifying selection, while plastid genomes are highly conserved yet genetically more diverse. Repetitive sequences and plastid DNA transfers drive mitochondrial genome expansion and rearrangement, whereas gene loss and duplication increase genomic plasticity. Single-copy genes provide robust markers for phylogenetic inference. These results provide a comprehensive understanding of the structural, functional, and evolutionary dynamics in Orchidaceae organelle genomes, laying a foundation for further studies on adaptation and diversification mechanisms.

Data availability statement

The data presented in the study are deposited in the NCBI GenBank repository, with the mitochondrial genome of *A. graminifolia* available as multiple circular molecules under accession numbers PX694305–PX694324, and the plastid genome under accession number PX694325.

Author contributions

SL: Data curation, Formal analysis, Funding acquisition, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. CS: Writing – review & editing. JW: Writing – review & editing. JK: Writing – review & editing. DP: Data curation, Writing – review & editing. JG: Resources, Writing – review & editing. ZW: Conceptualization, Funding acquisition, Validation, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was financially supported by Guangdong Basic and Applied Basic Research Foundation (2023A1515110239), the National Natural Science Foundation of China (32402621), the Chinese Academy of Agricultural Sciences Elite Youth Program (grant 110243160001007), the Guangdong Pearl River Talent Program

(grant 2021QN02N792), and the Funding of Major Scientific Research Tasks, Kunpeng Institute of Modern Agriculture at Foshan (KIMA-ZDKY2022004).

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author ZW declared that they were an Associate Editor for Frontiers in Plant Science, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2026.1756243/full#supplementary-material>

References

- Ahmad, S., Lu, C., Gao, J., Ren, R., Wei, Y., Wu, J., et al. (2021). Genetic insights into the regulatory pathways for continuous flowering in a unique orchid *Arundina graminifolia*. *BMC Plant Biol.* 21, 587. doi: 10.1186/s12870-021-03350-6
- Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198
- Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Chan, P. P., Lin, B. Y., Mak, A. J., and Lowe, T. M. (2021). tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* 49, 9077–9096. doi: 10.1093/nar/gkab688
- Chase, M. W., Cameron, K. M., Freudenstein, J. V., Pridgeon, A. M., Salazar, G., Van den Berg, C., et al. (2015). An updated classification of Orchidaceae. *Botanical J. Linn. Soc.* 177, 151–174. doi: 10.1111/boj.12234
- Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., et al. (2023). TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Mol. Plant* 16, 1733–1742. doi: 10.1016/j.molp.2023.09.010
- Christenhusz, M. (2016). The number of known plants species in the world and its annual increase. *phytotaxa*. 216, 201–217. doi: 10.11646/phytotaxa.261.3.1
- Cole, L. W., Guo, W., Mower, J. P., and Palmer, J. D. (2018). High and variable rates of repeat-mediated mitochondrial genome rearrangement in a genus of plants. *Mol. Biol. Evol.* 35, 2773–2785. doi: 10.1093/molbev/msy176

- Costello, R., Emms, D. M., and Kelly, S. (2020). Gene duplication accelerates the pace of protein gain and loss from plant organelles. *Mol. Biol. Evol.* 37, 969–981. doi: 10.1093/molbev/msz275
- Dressler, R. L. (1981). *The orchids: natural history and classification* (Cambridge, Massachusetts: Harvard University Press).
- Dressler, R. L. (1993). *Phylogeny and classification of the orchid family* (Cambridge, UK: Cambridge University Press).
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Givnish, T. J., Spalink, D., Ames, M., Lyon, S. P., Hunter, S. J., Zuluaga, A., et al. (2015). Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc. R. Soc. B: Biol. Sci.* 282, 20151553. doi: 10.1098/rspb.2015.1553
- Gray, M. W., Burger, G., and Lang, B. F. (1999). Mitochondrial evolution. *Science* 283, 1476–1481. doi: 10.1126/science.283.5407.1476
- Greiner, S., Lehwark, P., and Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 47, W59–W64. doi: 10.1093/nar/gkz238
- Han, F., Bi, C., Chen, Y., Dai, X., Wang, Z., Wu, H., et al. (2025). PMAT2: An efficient graphical assembly toolkit for comprehensive organellar genomes. *Imeta* 4, e70064. doi: 10.1002/imt2.70064
- He, W., Yang, J., Jing, Y., Xu, L., Yu, K., and Fang, X. (2023). NGenomeSyn: an easy-to-use and flexible tool for publication-ready visualization of syntenic relationships across multiple genomes. *Bioinformatics* 39, btad121. doi: 10.1093/bioinformatics/btad121
- Jansen, R. K., Cai, Z., Raubeson, L. A., Daniell, H., Depamphilis, C. W., Leebens-Mack, J., et al. (2007). Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc. Natl. Acad. Sci.* 104, 19369–19374. doi: 10.1073/pnas.0709121104
- Jansen, R. K., and Ruhlman, T. A. (2012). “Plastid genomes of seed plants,” in *Genomics of chloroplasts and mitochondria* (Dordrecht: Springer), 103–126.
- Jiang, L.-Q., Drew, B. T., Arthan, W., Yu, G.-Y., Wu, H., Zhao, Y., et al. (2024). Comparative plastome analysis of Arundinelleae (Poaceae, Panicoideae), with implications for phylogenetic relationships and plastome evolution. *BMC Genomics* 25, 1016. doi: 10.1186/s12864-024-10871-5
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kim, Y.-K., Jo, S., Cheon, S.-H., Joo, M.-J., Hong, J.-R., Kwak, M., et al. (2020). Plastome evolution and phylogeny of Orchidaceae, with 24 new sequences. *Front. Plant Sci.* 11, 22. doi: 10.3389/fpls.2020.00022
- Lee, S. Y., Meng, K., Wang, H., Zhou, R., Liao, W., Chen, F., et al. (2020). Severe plastid genome size reduction in a mycoheterotrophic orchid, *Danxiaorchis singhiana*, reveals heavy gene loss and gene relocations. *Plants* 9, 521. doi: 10.3390/plants9040521
- Lewis, S. E., Searle, S., Harris, N., Gibson, M., Iyer, V., Richter, J., et al. (2002). Apollo: a sequence annotation editor. *Genome Biol.* 3, research0082–1. doi: 10.1186/gb-2002-3-12-research0082
- Li, W., and Godzik, A. (2006). Cd-Hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinf. (Oxford England)* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158
- Li, Y.-X., Li, Z.-H., Schuitman, A., Chase, M. W., Li, J.-W., Huang, W.-C., et al. (2019). Phylogenomics of Orchidaceae based on plastid and mitochondrial genomes. *Mol. Phylogenet. Evol.* 139, 106540. doi: 10.1016/j.ympev.2019.106540
- Li, J., Ni, Y., Lu, Q., Chen, H., and Liu, C. (2025b). PMGA: A plant mitochondrial genome annotator. *Plant Commun.* 6, 101191. doi: 10.1016/j.xplc.2024.101191
- Li, Z., Ran, Z., Xiao, X., Yan, C., Xu, J., Tang, M., et al. (2024b). Comparative analysis of the whole mitochondrial genomes of four species in sect. *Chrysanthia* (Camellia L.), endemic taxa in China. *BMC Plant Biol.* 24, 955. doi: 10.1186/s12870-024-05673-6
- Li, J., Wei, Y., Jin, J., Gao, J., Xie, Q., Yang, F., et al. (2025a). A fragment insertion of *AgDFR* results in a white flower phenotype in *Arundina graminifolia* (Orchidaceae). *Plants* 14, 1680. doi: 10.3390/plants14111680
- Li, L., Wu, Q., Zhai, J., Wu, K., Fang, L., Li, M., et al. (2024a). Comparative chloroplast genomics of 24 species shed light on the genome evolution and phylogeny of subtribe Coelogyninae (Orchidaceae). *BMC Plant Biol.* 24, 31. doi: 10.1186/s12870-023-04665-2
- Li, X., Zhe, M., Huang, Y., Fan, W., Yang, J., and Zhu, A. (2023). The evolution of mitochondrial genomes between two *Cymbidium* sister species: dozens of circular chromosomes and the maintenance and deterioration of genome synteny. *Genes* 14, 864. doi: 10.3390/genes14040864
- Madden, T. (2013). The BLAST sequence analysis tool. *NCBI Handb.* 2, 425–436.
- Maier, U.-G., Zauner, S., Woehle, C., Bolte, K., Hempel, F., Allen, J. F., et al. (2013). Massively convergent evolution for ribosomal protein gene content in plastid and mitochondrial genomes. *Genome Biol. Evol.* 5, 2318–2329. doi: 10.1093/gbe/evt181
- Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., and Zimin, A. (2018). MUMmer4: a fast and versatile genome alignment system. *PLoS Comput. Biol.* 14, e1005944. doi: 10.1371/journal.pcbi.1005944
- Neuhaus, H., and Emes, M. (2000). Nonphotosynthetic metabolism in plastids. *Annu. Rev. Plant Biol.* 51, 111–140. doi: 10.1146/annurev.arplant.51.1.111
- Nhat Nam, N., Pham Anh Thi, N., and Do, H. D. K. (2024). New insights into the diversity of mitochondrial plastid DNA. *Genome Biol. Evol.* 16, evae184. doi: 10.1093/gbe/evae184
- Ouyang, L., Li, X., Wang, R., Chen, Y., Wang, S., Wang, J., et al. (2025). Comprehensive analysis of the mitochondrial genome of *Iris domestica* emphasizing multichromosomal organization and repeat-mediated homologous recombination. *Front. Plant Sci.* 15, 1520033. doi: 10.3389/fpls.2024.1520033
- Park, S., Grewe, F., Zhu, A., Ruhlman, T. A., Sabir, J., Mower, J. P., et al. (2015). Dynamic evolution of geranium mitochondrial genomes through multiple horizontal and intracellular gene transfers. *New Phytol.* 208, 570–583. doi: 10.1111/nph.13467
- Peng, H.-W., Lian, L., Zhang, J., Erst, A. S., and Wang, W. (2022). Phylogenomics, plastome degradation and mycoheterotrophy evolution of Neottieae (Orchidaceae), with emphasis on the systematic position and Loess Plateau-Changbai Mountains disjunction of *Diplandrorchis*. *BMC Plant Biol.* 22, 507. doi: 10.1186/s12870-022-03906-0
- Rasmussen, H. N., and Rasmussen, F. N. (2009). Orchid mycorrhiza: implications of a mycophagous life style. *Oikos* 118, 334–345. doi: 10.1111/j.1600-0706.2008.17116.x
- Raubeson, L. A., and Jansen, R. K. (2005). “Chloroplast genomes of plants,” in *Plant diversity and evolution: genotypic and phenotypic variation in higher plants* (CABI publishing, Wallingford UK), 45–68.
- Romero, H., Zavala, A., and Musto, H. (2000). Codon usage in *Chlamydia trachomatis* is the result of strand-specific mutational biases and a complex pattern of selective forces. *Nucleic Acids Res.* 28, 2084–2090. doi: 10.1093/nar/28.10.2084
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248
- Sharp, P. M., and Li, W.-H. (1986). Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for ‘rare’ codons. *Nucleic Acids Res.* 14, 7737–7749. doi: 10.1093/nar/14.19.7737
- Shen, B., Shen, A., Liu, L., Tan, Y., Li, S., and Tan, Z. (2024). Assembly and comparative analysis of the complete multichromosomal mitochondrial genome of *Cymbidium ensifolium*, an orchid of high economic and ornamental value. *BMC Plant Biol.* 24, 255. doi: 10.1186/s12870-024-04962-4
- Shi, L., Chen, H., Jiang, M., Wang, L., Wu, X., Huang, L., et al. (2019). CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47, W65–W73. doi: 10.1093/nar/gkz345
- Skippington, E., Barkman, T. J., Rice, D. W., and Palmer, J. D. (2017). Comparative mitogenomics indicates respiratory competence in parasitic *Viscum* despite loss of complex I and extreme sequence divergence, and reveals horizontal gene transfer and remarkable variation in genome size. *BMC Plant Biol.* 17, 49. doi: 10.1186/s12870-017-0992-8
- Sloan, D. B., Warren, J. M., Williams, A. M., Wu, Z., Abdel-Ghany, S. E., Chicco, A. J., et al. (2018). Cytonuclear integration and co-evolution. *Nat. Rev. Genet.* 19, 635–648. doi: 10.1038/s41576-018-0035-9
- Soe, T., Kong, J., Xu, B., Kan, J., Tembrock, L. R., Nie, L., et al. (2025). Mitochondrial genomes of four *Dendrobium* species (Orchidaceae): a comprehensive analysis of structural diversity, synteny, and plastid-derived sequences. *Genomics Commun.* 2, e018. doi: 10.48130/gcomm-0025-0018
- Su, X., Dong, Z., Yu, H., Yi, S., Gao, S., Liu, J., et al. (2025). Pan-mitogenome in Dipterocarpaceae: mitochondrial plastid DNAs and repeats shape the dynamic evolution of mitogenomes. *BMC Plant Biol.* 25, 1–13. doi: 10.1186/s12870-025-07506-6
- Sueoka, N. (1988). Directional mutation pressure and neutral molecular evolution. *Proc. Natl. Acad. Sci.* 85, 2653–2657. doi: 10.1073/pnas.85.8.2653
- Sun, M., Zhang, M., Chen, X., Liu, Y., Liu, B., Li, J., et al. (2022). Rearrangement and domestication as drivers of Rosaceae mitogenome plasticity. *BMC Biol.* 20, 181. doi: 10.1186/s12915-022-01383-3
- Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34, W609–W612. doi: 10.1093/nar/gkl315
- Tang, S., Liang, Y., Wu, F., Wu, Y., Li, J., Lin, L., et al. (2025). HiMT: An integrative toolkit for assembling organellar genomes using HiFi reads. *Plant Commun.* 6, 101467. doi: 10.1016/j.xplc.2025.101467
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9, e112963. doi: 10.1371/journal.pone.0112963
- Wang, Y., Jia, L., Tian, G., Dong, Y., Zhang, X., Zhou, Z., et al. (2023). shinyCircos-V2.0: leveraging the creation of circos plot with enhanced usability and advanced features. *Imeta* 2, e109. doi: 10.1002/imt2.109
- Wang, J., Kan, S., Liao, X., Zhou, J., Tembrock, L. R., Daniell, H., et al. (2024b). Plant organellar genomes: much done, much more to do. *Trends Plant Sci.* 29, 754–769. doi: 10.1016/j.tplants.2023.12.014
- Wang, L., Liu, X., Wang, Y., Ming, X., Qi, J., and Zhou, Y. (2024a). Comparative analysis of the mitochondrial genomes of four *Dendrobium* species (Orchidaceae)

reveals heterogeneity in structure, synteny, intercellular gene transfer, and RNA editing. *Front. Plant Sci.* 15, 1429545. doi: 10.3389/fpls.2024.1429545

Wang, D., Rousseau-Gueutin, M., and Timmis, J. N. (2012). Plastid sequences contribute to some plant mitochondrial genes. *Mol. Biol. Evol.* 29, 1707–1711. doi: 10.1093/molbev/mss016

Wang, H., Wang, D., Shao, B., Li, J., Li, Z., Chase, M. W., et al. (2025). Unequally abundant chromosomes and unusual collections of transferred sequences characterize mitochondrial genomes of *Gastrodia* (Orchidaceae), one of the largest mycoheterotrophic plant genera. *Mol. Biol. Evol.* 42, msaf082. doi: 10.1093/molbev/msaf082

Wang, D., Wu, Y.-W., Shih, A. C.-C., Wu, C.-S., Wang, Y.-N., and Chaw, S.-M. (2007). Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. *Mol. Biol. Evol.* 24, 2040–2048. doi: 10.1093/molbev/msm133

Wang, J., Zou, Y., Mower, J. P., Reeve, W., Wu, Z., Wang, J., et al. (2024c). Rethinking the mutation hypotheses of plant organellar DNA. *Genomics Commun.* 1, e003. doi: 10.48130/gcomm-0024-0003

Wen, Y., Qin, Y., Shao, B., Li, J., Ma, C., Liu, Y., et al. (2022). The extremely reduced, diverged and reconfigured plastomes of the largest mycoheterotrophic orchid lineage. *BMC Plant Biol.* 22, 448. doi: 10.1186/s12870-022-03836-x

Wicke, S., Schneeweiss, G. M., Depamphilis, C. W., Müller, K. F., and Quandt, D. (2011). The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol. Biol.* 76, 273–297. doi: 10.1007/s11103-011-9762-4

Wickham, H. (2011). ggplot2. *Wiley Interdiscip. reviews: Comput. Stat* 3, 180–185. doi: 10.1002/wics.147

Wong, T. K., Ly-Trong, N., Ren, H., Baños, H., Roger, A. J., Susko, E., et al. (2025). IQ-TREE 3: phylogenomic inference software using complex evolutionary Models.

Woodson, J. D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. *Nat. Rev. Genet.* 9, 383–395. doi: 10.1038/nrg2348

Wu, Z.-Q., Liao, X.-Z., Zhang, X.-N., Tembrock, L. R., and Broz, A. (2022). Genomic architectural variation of plant mitochondria—a review of multichromosomal structuring. *J. systematics Evol.* 60, 160–168. doi: 10.1111/jse.12655

Wu, Y., Zeng, M.-Y., Wang, H.-X., Lan, S., Liu, Z.-J., Zhang, S., et al. (2024). The complete chloroplast genomes of *Bulbophyllum* (Orchidaceae) species: insight into genome structure divergence and phylogenetic analysis. *Int. J. Mol. Sci.* 25, 2665. doi: 10.3390/ijms25052665

Xian, W., Bezrukov, I., Bao, Z., Vorbrugg, S., Gautam, A., and Weigel, D. (2025). TIPPO: a user-friendly tool for *de novo* assembly of organellar genomes with high-fidelity data. *Mol. Biol. Evol.* 42, msae247. doi: 10.1093/molbev/msae247

Xu, Y., Lei, Y., Su, Z., Zhao, M., Zhang, J., Shen, G., et al. (2021). A chromosome-scale *Gastrodia elata* genome and large-scale comparative genomic analysis indicate convergent evolution by gene loss in mycoheterotrophic and parasitic plants. *Plant J.* 108, 1609–1623. doi: 10.1111/tj.15528

Yang, J.-X., Dierckxsens, N., Bai, M.-Z., and Guo, Y.-Y. (2023). Multichromosomal mitochondrial genome of *Paphiopedilum micranthum*: compact and fragmented genome, and rampant intracellular gene transfer. *Int. J. Mol. Sci.* 24, 3976. doi: 10.3390/ijms24043976

Yuan, Y., Jin, X., Liu, J., Zhao, X., Zhou, J., Wang, X., et al. (2018). The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. *Nat. Commun.* 9, 1615. doi: 10.1038/s41467-018-03423-5

Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., et al. (2020). PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* 20, 348–355. doi: 10.1111/1755-0998.13096

Zhang, Z., Xiao, J., Wu, J., Zhang, H., Liu, G., Wang, X., et al. (2012). ParaAT: a parallel tool for constructing multiple protein-coding DNA alignments. *Biochem. Biophys. Res. Commun.* 419, 779–781. doi: 10.1016/j.bbrc.2012.02.101

Zhang, S., Yang, Y., Li, J., Qin, J., Zhang, W., Huang, W., et al. (2018). Physiological diversity of orchids. *Plant Diversity* 40, 196–208. doi: 10.1016/j.pld.2018.06.003

Zhang, D., Zhao, X.-W., Li, Y.-Y., Ke, S.-J., Yin, W.-L., Lan, S., et al. (2022). Advances and prospects of orchid research and industrialization. *Horticulture Res.* 9, uhac220. doi: 10.1093/hr/uhac220

Zhao, Z., Li, Y., Zhai, J.-W., Liu, Z.-J., and Li, M.-H. (2024). Organelle genomes of *Epipogium roseum* provide insight into the evolution of mycoheterotrophic orchids. *Int. J. Mol. Sci.* 25, 1578. doi: 10.3390/ijms25031578

Zheng, Q., Luo, X., Huang, Y., Ke, S.-J., and Liu, Z.-J. (2024). The complete mitogenome of *Apostasia fujianica* Y. Li & S. Lan and comparative analysis of mitogenomes across Orchidaceae. *Int. J. Mol. Sci.* 25, 8151. doi: 10.3390/ijms25158151

Zhou, C., Brown, M., Blaxter, M., Consortium, D. T. of L. P., McCarthy, S. A., and Durbin, R. (2025). OatK: a *de novo* assembly tool for complex plant organelle genomes. *Genome Biol.* 26, 235. doi: 10.1186/s13059-025-03676-6