



OPEN ACCESS

EDITED BY

Wen-Bin Yu,
Chinese Academy of Sciences (CAS), China

REVIEWED BY

Peng-Fei Ma,
Chinese Academy of Sciences (CAS), China
Xiaohua Jin,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE

Zhitao Niu
✉ niuzhitaonj@163.com
Xiaoyu Ding
✉ dingxyjn@263.net

RECEIVED 23 July 2024

ACCEPTED 01 October 2024

PUBLISHED 22 October 2024

CITATION

Hou Z, Wang M, Jiang Y, Xue Q, Liu W, Niu Z and Ding X (2024) Mitochondrial genome insights into the spatio-temporal distribution and genetic diversity of *Dendrobium hancockii* Rolfe (Orchidaceae). *Front. Plant Sci.* 15:1469267. doi: 10.3389/fpls.2024.1469267

COPYRIGHT

© 2024 Hou, Wang, Jiang, Xue, Liu, Niu and Ding. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Mitochondrial genome insights into the spatio-temporal distribution and genetic diversity of *Dendrobium hancockii* Rolfe (Orchidaceae)

Zhenyu Hou^{1,2}, Mengting Wang^{1,2,3}, Yu Jiang^{1,2}, Qingyun Xue^{1,2}, Wei Liu^{1,2}, Zhitao Niu^{1,2*} and Xiaoyu Ding^{1,2*}

¹College of Life Science, Nanjing Normal University, Nanjing, China, ²Jiangsu Provincial Engineering Research Center for Technical Industrialization for Dendrobiums, Nanjing, China, ³Ningbo Key Laboratory of Agricultural Germplasm Resources Mining and Environmental Regulation, College of Science and Technology, Ningbo University, Cixi, China

Introduction: With its distinctive evolutionary rate and inheritance patterns separate from the nuclear genome, mitochondrial genome analysis has become a prominent focus of current research. *Dendrobium hancockii* Rolfe, a species of orchid with both medicinal and horticultural value, will benefit from the application of the fully assembled and annotated mitochondrial genome. This will aid in elucidating its phylogenetic relationships, comparative genomics, and population genetic diversity.

Methods: Based on sequencing results from Illumina combined with PacBio and Nanopore, the mitochondrial genome map of *D. hancockii* was constructed. Comparative analysis was conducted from the perspectives of phylogeny across multiple species, selection pressure on protein-coding genes, and homologous segments. The population diversity of *D. hancockii* was analyzed using single nucleotide polymorphism (SNP) data from the mitochondrial genome and single-copy nuclear genes.

Results and discussion: This research constructed a circular mitochondrial map for *D. hancockii*, spanning 523,952 bp, containing 40 unique protein-coding genes, 37 transfer RNA genes, and 4 ribosomal RNA genes. Comparative analysis of mitochondrial genes from 26 land plants revealed a conserved gene cluster, "*rpl16-ccmFn-rps3-rps19*," particularly within the *Dendrobium* genus. The mitochondrial genome of *D. hancockii* exhibits a lower point mutation rate but significant structural variation. Analysis of 103 resequencing samples identified 19,101 SNP sites, dividing *D. hancockii* into two major groups with limited gene flow between them, as supported by population diversity, genetic structure analysis, principal component analysis, and phylogenetic trees. The geographical distribution and genetic differentiation of *D. hancockii* into two major groups

suggest a clear phytogeographical division, likely driven by ancient geological or climatic events. The close alignment of mitochondrial data with nuclear gene data highlights the potential of the mitochondrial genome for future studies on genetic evolution in this species.

KEYWORDS

Dendrobium hancockii, mitochondrial genome, SNP, genetic diversity, nuclear single copy gene

1 Introduction

The *Dendrobium hancockii*, known for its alkaloids (Bhardwaj et al., 2024), polysaccharides (Lai et al., 2024), and flavonoids (Meng et al., 2023) found in its stems, is commonly regarded as a herbal medicine (Dendrobii Herba) in East Asia. Additionally, belonging to the Orchidaceae family, the unique floral morphology of *Dendrobium* gives it potential as an ornamental plant (Zhang et al., 2022). As a species widely distributed throughout China (Pan et al., 2022), its range extends across the Sino-Himalayan Forest Subkingdom and the Sino-Japanese Floristic Region, showcasing its extensive environmental adaptability as both a horticultural and medicinal plant. However, samples from these different phytogeographic regions exhibit phenotypic variations, particularly in leaf shape, plant height, and floral morphology. Previous research has predominantly focused on the development and utilization of medicinal components, with limited studies on species evolution, phylogeny, and genetic diversity.

Compared to nuclear and plastid genomes, the mitochondrial genome has long been considered a key focus in evolutionary studies within genetic systems. This recognition stems from several characteristics, including its large size, complex and highly adaptable structure (Møller et al., 2021), varied nucleotide substitution rates (Drouin et al., 2008; Chevigny et al., 2020), and widespread horizontal gene transfer events (Waltz and Giegé, 2020). The exploration of mitochondrial genomes in green plants and their genetic reservoirs can be traced back to the pioneering work of Unsel (Unsel et al., 1997), which initially elucidated the mitochondrial genome of *Arabidopsis thaliana*. With advances in sequencing technologies and mitochondrial genome research, a vast structural diversity has now been revealed within plant mitochondrial genomes. This diversity is evident in single-ring mitochondrial genomes (Pinard et al., 2019) and multi-ring

mitochondrial genomes (Lloyd et al., 2019). The variety of structures and differing evolutionary rates endow plant mitochondrial genomes with significant potential and applications in structural properties, homologous gene transfer, genetic architecture, and phylogeography.

Mitochondrial conformations often exhibit structural diversity mediated by repetitive sequences (Sun et al., 2022) or the insertion of foreign sequences (Choi et al., 2019). This structural diversity can occur even among closely related species within the same genus. Research on *Fragaria* (strawberries) indicates that such structural variability, rather than variability at gene loci, contributes to the rapid evolution of mitochondrial genomes (Fan et al., 2022). Furthermore, horizontal gene transfer among chloroplast, mitochondrial, and nuclear genomes, leading to derived fragments, also promotes species evolution (Kleine et al., 2009). The combination of foreign fragments with diverse structures, alongside recombinational changes, enables the mitochondrial genome to better adapt to evolving environments (Ma et al., 2022; Liu et al., 2013).

Furthermore, the dissimilarity in synonymous substitution rates between nuclear and plastid genomes also enhances the promise and value of mitochondrial-related systematic geography and biological evolution studies (Drouin et al., 2008). Utilizing molecular markers derived from the mitochondrial genome to investigate the origin, evolution, and dispersal patterns of mustard greens shows significant potential and prospects (You et al., 2022). However, research on the mitochondrial genome and its relevance in population evolution remains limited in most angiosperms, particularly in the case of orchids. To elucidate the mechanisms behind plant mitochondrial structural variations, replication, gene transfer events and genetic diversity, as well as their evolutionary history, a greater abundance of mitochondrial sequence data and comprehensive analyses are imperative.

Therefore, we initiated *de novo* sequencing to unravel the enigmatic mitochondrial genome of the rare *D. hancockii*, aiming to address the following critical inquiries: (a) furnish a comprehensive, high-quality mitochondrial genome assembly of *D. hancockii*, elucidating its structural attributes, sequence composition, and gene distribution within the species; (b) engage in comparative genomics analyses, leveraging our previously published *Dendrobium* mitochondrial genomes, to unearth

Abbreviations: PCR, Polymerase chain reaction; SSR, Simple sequence repeat; ML, Maximum-likelihood; NJ, Neighbor-joining; NCBI, National Center for Biotechnology Information; BLAST, Basic Local Alignment Search Tool; PCGs, Protein-coding gene sequences; MTPT, Mitochondrial plastid; DNA lncRNA, Long noncoding RNA; SNP, Single nucleotide polymorphism; Ka/Ks, the nonsynonymous to synonymous substitution ratio; PCA, Principal component analysis.

potential molecular marker fragments and shed light on speciation events; (c) employ mitochondrial genome SNP data along with nuclear single-copy genes to conduct genetic structure investigations of *D. hancockii*, assessing the feasibility of employing plant mitochondrial genomes in genetic differentiation and systematic evolution.

2 Materials and methods

2.1 Materials and DNA extraction

The material used for sequencing the mitochondrial genome is the root tip tissue of tissue-cultured *D. hancockii* seedlings, sourced from the *D. hancockii* population in Pingbian, Yunnan. The remaining individual samples of the *D. hancockii* population are either plant or leaf samples collected by the research team from 2019 to 2021 from various provinces, including Yunnan, Guangxi, Shanxi, Shaanxi, Hubei, etc. All samples have been identified as *D. hancockii*. The leaves are preserved using silica gel desiccant and stored in a -80°C refrigerator. The plants are maintained in the cultivation room at the Institute of Plant Resources and Environment, College of Life Sciences, Nanjing Normal University.

Total DNA was extracted using 0.2 g of fresh plant tissue using the CATB method (Xu et al., 2012), followed by agarose gel electrophoresis to detect the quality of DNA; the concentration of DNA was measured by spectrophotometry (Preuten et al., 2010), and all samples were diluted to the same concentration. According to the *D. hancockii* resequencing data mapped to the *D. officinale* reference genome (Niu et al., 2021) using CLC Genomics Workbench 8.5.1 (CLC Biosciences, Aarhus, Denmark), 8 homologous genes were selected. Subsequently, 11 pairs of primers were designed for these genes or gene fragments using Primer Premier 5.0 (Singh et al., 1998), as listed in Supplementary Table S1. The PCR reaction followed a standard procedure: initial denaturation at 94°C for 5 min, 30 cycles (denaturation at 95°C for 15 s \rightarrow annealing at T_m for 30 s \rightarrow extension at 72°C for 30 s), and a final extension at 72°C for 5 min.

Mitochondria were isolated from 5 g of root tip tissue of *D. hancockii* using a modified method involving differential centrifugation and density gradient centrifugation (Gui et al., 2016). Mitochondrial DNA was extracted from purified mitochondria using a modified SDS method (Chen et al., 2011). Check the quality of mtDNA by agarose gel electrophoresis. DNA samples meeting the quality requirements (concentration ≥ 50 ng/ μL , $A_{260}/A_{280} = 1.8\text{--}2.0$, $A_{260}/A_{230} > 1.7$) were used for sequencing.

2.2 DNA sequencing, mitochondrial genome assembly and annotation

Sequence: Purified mtDNA of *D. hancockii* species was sequenced on Illumina HiSeq. Platform (Illumina, San Diego, CA, USA) generated 150 bp paired-end reads (with an insert size of 450 bp) and obtained approximately 4.2 Gb of raw data. Low-quality

reads were trimmed using CLC Genomics Workbench 8.5.1 (CLC Biosciences, Aarhus, Denmark). In addition, the PacBio (Pacific Biosciences, Menlo Park, CA, USA) and Nanopore (Oxford Nanopore Biosciences, Cambridge, MA, USA) platforms generated approximately 43817 long reads from *D. hancockii*. Corresponding Illumina pair-end reads were used to correct erroneous bases with PacBio (nanopore) and indels were read using LoRDEC (kmer value = 19; abundance threshold = 3) (Salmela and Rivals, 2014).

Assembly: Firstly, the Illumina sequencing data was preliminarily assembled using ABYSS v2.0.2 (<http://www.bcgsc.ca/platform/bioinfo/software/abyss>) (Jackman et al., 2017). Subsequently, the Nanopore third-generation data was aligned using blasR, and based on the alignment results, a correction and error correction were applied to the single-molecule sequencing data. The purpose of this step was to reduce errors such as single-base substitutions, insertions, and deletions in the long single-molecule sequences. Finally, the corrected single-molecule sequencing data was combined with the second-generation data for hybrid assembly using SPAdes-3.13.0 software (kmer values = 33; phred-offset = 33) (Meleshko et al., 2019). High-coverage and long-assembled sequences were selected as candidate sequences, and the mitochondrial scaffold sequences were confirmed by aligning against the NT library. Clean reads were aligned back to the assembled mitochondrial scaffolds. Based on the paired-end and overlap relationships of the reads, local assembly and optimization were performed on the assembly results to determine the order and direction of the scaffolds. Subsequently, the GapCloser software (v1.12, <http://soap.genomics.org.cn/soapdenovo.html>) (Luo et al., 2012) was employed to perform gap filling on the assembly results, resulting in the final mitochondrial genome sequence.

Annotation: We employed a homology-based prediction method for gene prediction in the sample genome. Initially, the protein sequences of 36 angiosperm mitochondrial reference genomes (Supplementary Table S2) were rapidly aligned to the sample genome using BLAST+ 2.71 software with the parameter set as (e-value $< 1e-5$). After filtering out poor alignment results and removing redundancies, manual correction of the sample genes' integrity and exon/intron boundaries was performed to obtain a highly accurate set of conserved genes. Homology prediction and software such as tRNAscan-SE 2.0.4 (Schattner et al., 2005) were used to predict non-coding RNAs (ncRNAs) contained in the genome.

2.3 Chloroplast to mitochondrion DNA transfer and RNA editing analyses

Prediction of RNA editing sites in protein-coding genes using PREP-mt (Mower, 2009). The cutoff value was set as default. RE sites of double-copy genes were counted twice. To identify plastid-derived fragments, the mitogenome of *D. hancockii* was compared against the plastomes of *D. hancockii* (LC500592.1) (Hou et al., 2019) using BLAST v2.11.0+ (Camacho et al., 2009), with criteria set as a matching rate of $\geq 70\%$, E-value of $\leq 1e-10$, and length of ≥ 40 bp.

2.4 Analysis of repeated sequences

The Microsatellite identification tool facilitated the detection of simple sequence repeats (Beier et al., 2017) (<https://webblast.ipkgatersleben.de/misa/index.php>). Tandem repeats with a repeat unit of more than 6p were identified utilizing Tandem Repeats Finder v4.09 software (Benson, 1999) (<http://tandem.bu.edu/trf/trf.submit.options.html>) employing default parameters. Direct and inverted repeats were identified using REPuter software (Kurtz et al., 2001) (<https://bibiserv.cebitec.uni-bielefeld.de/reputer>) with a minimal repeat size threshold set at 20 bp.

Co-linearity analysis employed two methods. To better illustrate the relationships among the Locally Collinear Blocks (LCBs) of the 26 mitochondria (Supplementary Table S3), we utilized the Mauve software (Darling et al., 2004) (<http://darlinglab.org/mauve/mauve.html>) for co-linearity analysis. For the identification of conserved gene clusters in the *Dendrobium* genus, we employed the MCScanX software (Wang et al., 2012) (<https://github.com/wyp1125/MCScanX>) for co-linearity analysis.

2.5 Analysis of codon usage bias and calculation of selective pressure

The MEGA X software (Kumar et al., 2018) was utilized to calculate the Relative Synonymous Codon Usage (RSCU) values of Protein-Coding Genes (PCGs) and their corresponding amino acid composition. Codon preferences were established through the implementation of Perl scripts. The non-synonymous (K_a) and synonymous (K_s) substitution rates were computed using DnaSP 5.10 (Librado and Rozas, 2009), considering a total of 13 shared PCGs among the mitogenomes of *D. hancockii* and 25 other land plants.

2.6 Read mapping and SNP calling

Before read mapping, Trimmomatic v.0.38 (Lohse et al., 2012) was employed to eliminate adapters and trim low-quality bases. Bases with a quality score below 20 were removed from the beginning or end of reads, and reads shorter than 36 bases after trimming were discarded. The quality of reads was evaluated using FastQC v.0.11.7 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, last accessed October 23, 2019) both before and after trimming. The remaining read pairs were aligned to the complete mitochondrial genome of *D. hancockii* using BWA-MEM v.0.7.17 (parameter: mem -t 4, -k 32 -M) (Li, 2013) and subsequently sorted with SAMtools v.1.9 (parameter: r m d u p) (Li et al., 2009). The MarkDuplicates tool from Picard v.2.18.21 (<http://broadinstitute.github.io/picard/>, last accessed March 16, 2019) was then applied to identify and mark PCR duplicates, while realignment around indels was performed using GATK v.3.8.1 (DePristo et al., 2011). Ultimately, GATK v.3.8.1 HaplotypeCaller was utilized for per-sample variant detection, and the

GenotypeGVCFs tool was employed for joint genotyping. Given that our dataset comprises individuals from multiple species, we utilized the “EMIT_ALL_SITES” flag to generate a “mutation sites” VCF, which includes both variant and nonvariant sites, during SNP calling and genotyping.

2.7 Phylogenetic analysis

Protein-coding sequences of 13 mitochondrial genes and complete mitochondrial genome that were shared by 26 species (include two outgroups, *Psilotum nudum* and *Funaria hygrometrica*), were used to construct a constraint phylogeny for rate comparisons. Sequence alignments for each gene were performed using MAFFT 7.0 (Kato and Standley, 2013). The maximum likelihood (ML) tree was constructed using Phylogenies were inferred by RAxMLv8.0.0 (Stamatakis, 2014) with (GTR + G + I) model and 1000 bootstrap replicates, and *Dendrobium kingianum* was selected as an outgroup IQ-TREE v.1.6.12 (Nguyen et al., 2015) with non-parametric bootstrap from 1000 pseudoreplicates. The neighbor-joining (NJ) tree (Tamura et al., 2004) was constructed using Phylogenies were inferred by MEGA X (Kumar et al., 2018) with Kimura 2-parameter model and 1000 bootstrap replicates. The ML/NJ tree with bootstrap support values was visualized using FigTree v.1.4.4 (<https://tree.bio.ed.ac.uk/software/figtree/>).

2.8 Population genetic diversity and genetic structure analysis

We identified nuclear gene sequences using version DnaSP 5.10 (Librado and Rozas, 2009) to calculate nucleotide diversity (π). Calculate the genetic differentiation index F_{st} and gene flow Nm ($Nm \approx (1 - F_{st})/4F_{st}$) between populations. Use ArcGIS ver. 10.2 (<http://desktop.arcgis.com>) software was used to create geographic maps. Use PERMUT to calculate intra-population gene diversity (H_S), genetic diversity (H_T) across all populations, interpopulation differentiation (G_{ST}), and number of surrogate types (N_{ST}). The last two indices (G_{ST} and N_{ST}) were analyzed by a permutation test with 1000 permutations. Analysis of molecular variance (AMOVA) using Arlequin ver 3.5.2.246 divided different levels of genetic variation, statistically significant determined by 1,000 permutations, squared deviations (SSDs), Harpending's raggedness index (HRI), and corresponding P values were calculated by Arlequin ver.3.5.2.246.

2.9 Structure and PCA analysis

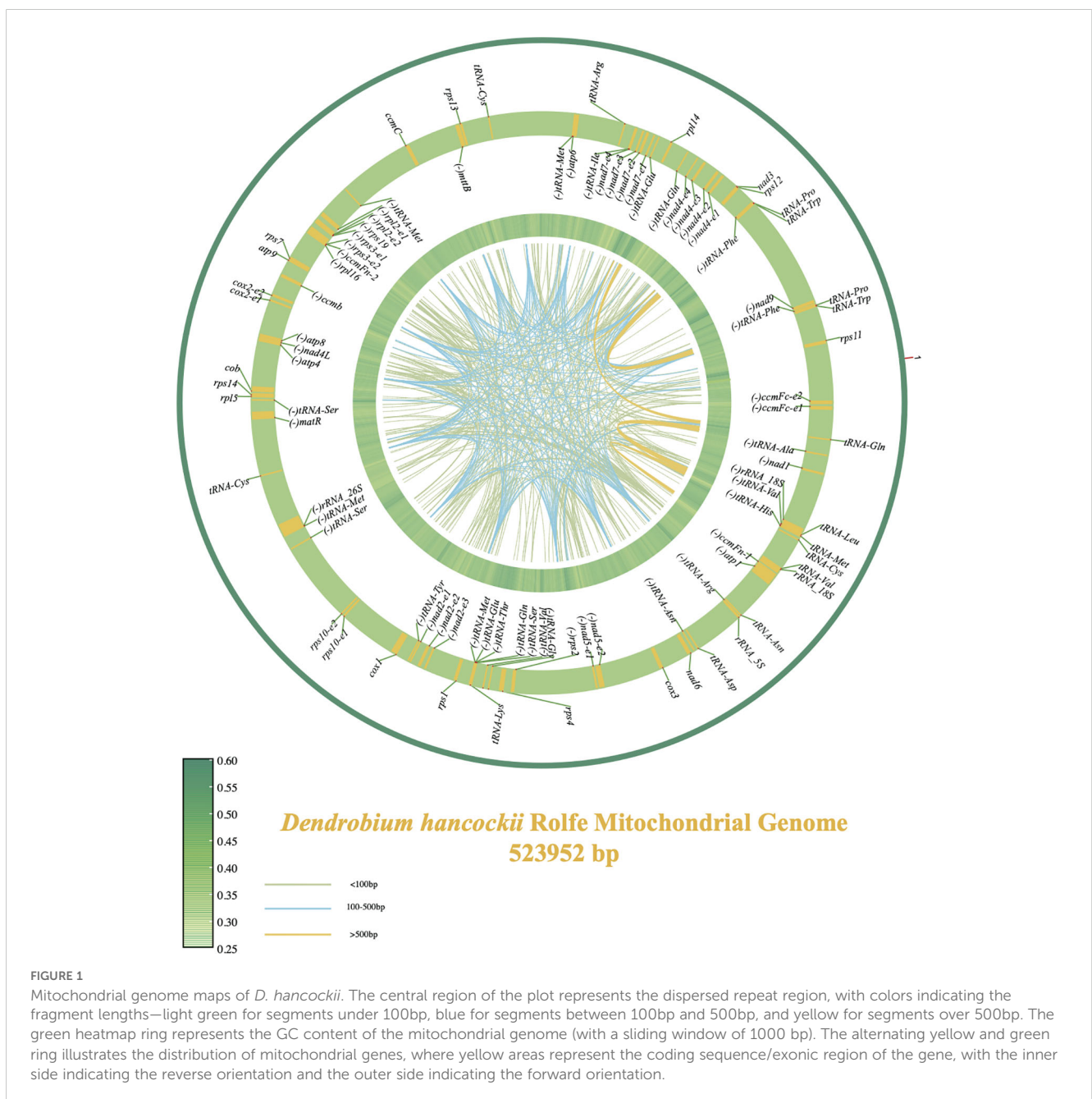
The model-based program STRUCTURE (Raj et al., 2014) was used to determine the genetic structure of five populations of *D.*

hancockii by using mtSNPs and nuclear gene single-copy fragment combinations data (Falush et al., 2003). We specified a burn-in of 500,000 and set the MCMC iteration number to 1,200,000. We assumed an admixture model and correlated allele frequencies, and no prior information on taxon identity was included; default values were applied for all other parameters. The number of groups (K) ranged from 1 to 10 with 10 independent runs. The most likely number of clusters was determined by computing the natural logarithm of the likelihood function, and the ΔK statistic was calculated using Structure Harvester (Earl and vonHoldt, 2012). Conversely, for Principal Component Analysis (PCA) on the pruned SNPs, we utilized the smartpca program from EIGENSOFT v.7.2.1 (Patterson et al., 2006).

3 Results

3.1 Mitogenome structure and gene content

The mitochondrial genome of tissue-cultured *D. hancockii* root tip cells was meticulously assembled into a single, large circular molecule, boasting a total length of 523,952 base pairs (bp) and an overall glycine cysteine content of 44.28% (Accession number: LC771085). Thorough annotation efforts resulted in the prediction and annotation of a total of 81 genes within the mitochondrial genome, including 40 protein-coding genes (PCGs), 4 ribosomal RNA (rRNA) genes, and 37 transfer RNA (tRNA) genes (Figure 1).



Of the 40 conserved PCGs present in all plant mitochondrial genomes, 17 genes encode vital components related to electron transport proteins and ATP synthase. This set includes nine subunits of complex I (*nad1*, 2, 3, 4, 5, 6, 7, 9, and 4L), one subunit of complex III (*cob*), three subunits of complex IV (*cox1*, 2, 3), and five subunits of complex V (*atp1*, 4, 6, 8, and 9). Furthermore, four genes were identified for large ribosomal proteins (*rpl2*, 5, 14, 16), eleven genes for small ribosomal proteins (*rps1*, 2, 3, 4, 7, 10, 11, 12, 13, 14, and 19), four genes for cytochrome c biogenesis (*ccmB*, *ccmC*, *ccmFn*, and *ccmFc*), and two genes for maturase and transporter functions (*matR* and *mttB*). The combined length of the 40 PCGs, ranging from 273 bp (*nad4L*) to 1,905 bp (*matR*), totaled 34,116 bp, accounting for 6.5% of the entire mitochondrial genome. According to the Relative Synonymous Codon Usage (RSCU) analysis presented in (Supplementary Figure S1), it is evident that leucine constitutes the predominant amino acid in the coding genes, accounting for 9.62% of the total codons with 1094 occurrences. Subsequently, serine is the second most prevalent, represented by 1071 codons, contributing to 9.42% of the total. Arginine follows with 894 codons, making up 7.86% of the overall composition, as revealed by the RSCU analysis.

Additionally, within the identified PCGs, six protein-coding genes (*ccmFc*, *cox2*, *rps3*, *nad5*, *rpl2*, *rps10*) contain one intron each. Furthermore, the *nad2* gene is characterized by the presence of two introns, while the *nad4* and *nad7* genes each contain three introns.

3.2 Analyses of repeat elements in *D. hancockii* mitogenome

Traditional analysis of mitochondrial repeats categorizes them into three types: simple repeats (SSR), tandem repeats, and scattered repeats. Through the utilization of an online Microsatellite Identification tool, we identified a total of 50 SSRs in the mitochondrial genome. These consist primarily of 46 mononucleotide repeat sequences, accounting for 92% of the total, alongside a small proportion (8%) of 4 dinucleotide repeat sequences (Table 1). Through the Tandem Repeats Finder tool, we identified 15 tandem repeats distributed evenly within the

intergenic regions, with lengths ranging from 25 bp to 612 bp (Supplementary Table S4). Additionally, we employed the repetur software to quantify dispersed repeats within the mitochondrial genome. This analysis revealed a total of 437 dispersed repeats (larger than 20bp), as outlined in Supplementary Table S5. The combined length of these dispersed repeat sequences is 50,454 bp, accounting for 9.62% of the total length of the mitochondrial genome. Among these, repetitive segments with lengths less than 100 bp dominate, comprising 315 segments, accounting for 72.08% of the total segments. The lengths of 117 repetitive segments range from 100 bp to 500 bp, representing 26.77% of the total number of repeat segments, while the remaining 1.14% of segments have lengths exceeding 500 bp.

3.3 Characterization of *D. hancockii* plastid genome transfer into the mitogenome

To identify chloroplast fragments within the mitochondrial genome, we utilized the mitochondrial genome as a reference and compared it with the published chloroplast genome (GenBank: LC500592.1). A total of 55 similar fragments were compared, ranging in length from 32 bp to 2434 bp, as outlined in Supplementary Table S6. These comparable sequences have a collective length of 31333 bp, constituting approximately 5.98% of the entire mitochondrial genome. Notably, a single plastid gene, *rpl14*, has been transferred and annotated in the mitochondrial genome of *D. hancockii*. We have identified nine intact tRNA genes of plastid origin in the mitochondrial genome. (*trnV-GAC*, *trnQ-UUG*, *trnL-CAA*, *trnM-CAU*, *trnV-GAC*, *trnQ-UUG*, *trnI-CAU*, *trnN-GUU*, *trnS-GCU*). Moreover, We validated the regional boundaries of mitochondrial genomes derived from partial plastids through PCR to ensure the reliability of assembly.

3.4 Comparative analysis of mitochondrial genomes

Currently, we have conducted a comparative analysis of 26 plant mitochondrial genomes based on published data, encompassing a

TABLE 1 Frequency of identified SSR motifs in *D. hancockii* mitogenome.

Motif Type	Number of Repeats							Total	Proportion (%)
	6	7	8	9	10	11	12		
Monomer	0	0	0	0	32	11	3	46	92.00
Dimer	2	2	0	0	0	0	0	4	8.00
Trimer	0	0	0	0	0	0	0	0	0
Tetramer	0	0	0	0	0	0	0	0	0
Pentamer	0	0	0	0	0	0	0	0	0
Hexamer	0	0	0	0	0	0	0	0	0
Total	2	2	0	0	32	11	3	50	100

Monomer: Single nucleotide repeats; Dimer: Single nucleotide repeats.

diverse collection of land plants including bryophytes, ferns, gymnosperms, and angiosperms (Supplementary Table S3). Among land plants, the key functional genes of mitochondria, particularly those associated with electron transport chains and oxidative phosphorylation (*atp1*, 6, 9, *ccmB*, *ccmC*, *ccmFn*, *ccmFn*), exhibit high conservation. On the contrary, the *sdh3* and *sdh4* genes, components of mitochondrial complex II, have undergone substantial loss (Figure 2). Differentiation in mitochondrial gene content among land plants predominantly arises from variations in ribosomal protein-related genes(*rps2*, *rps10*, *rps11*, *rpl2*, *rpl10*, *rpl14*).

Moreover, we applied a selection process based on core mitochondrial genes to the aforementioned 26 land plants,

resulting in the identification of 13 conserved genes(*atp1*, 6, 9, *ccmB*, *ccmC*, *ccmFn*, *cob*, *cox1*, 2, *nad4*, 6, 9, *rps3*). These genes were then grouped together to construct a phylogenetic tree using the Neighbor-Joining method. The topology of the gene tree aligns with the known species tree, exhibiting a support rate exceeding 75%. Through phylogenetic analysis, we established the evolutionary relationships among the mitochondrial genomes of the 26 plants, representing the major taxa within land plants. Bryophytes and ferns were chosen as outgroups for the analysis. Based on the phylogenetic tree results, it distinctly indicates a parallel relationship between gymnosperms and angiosperms. Monocots and eudicots of angiosperms form a cluster, within which our

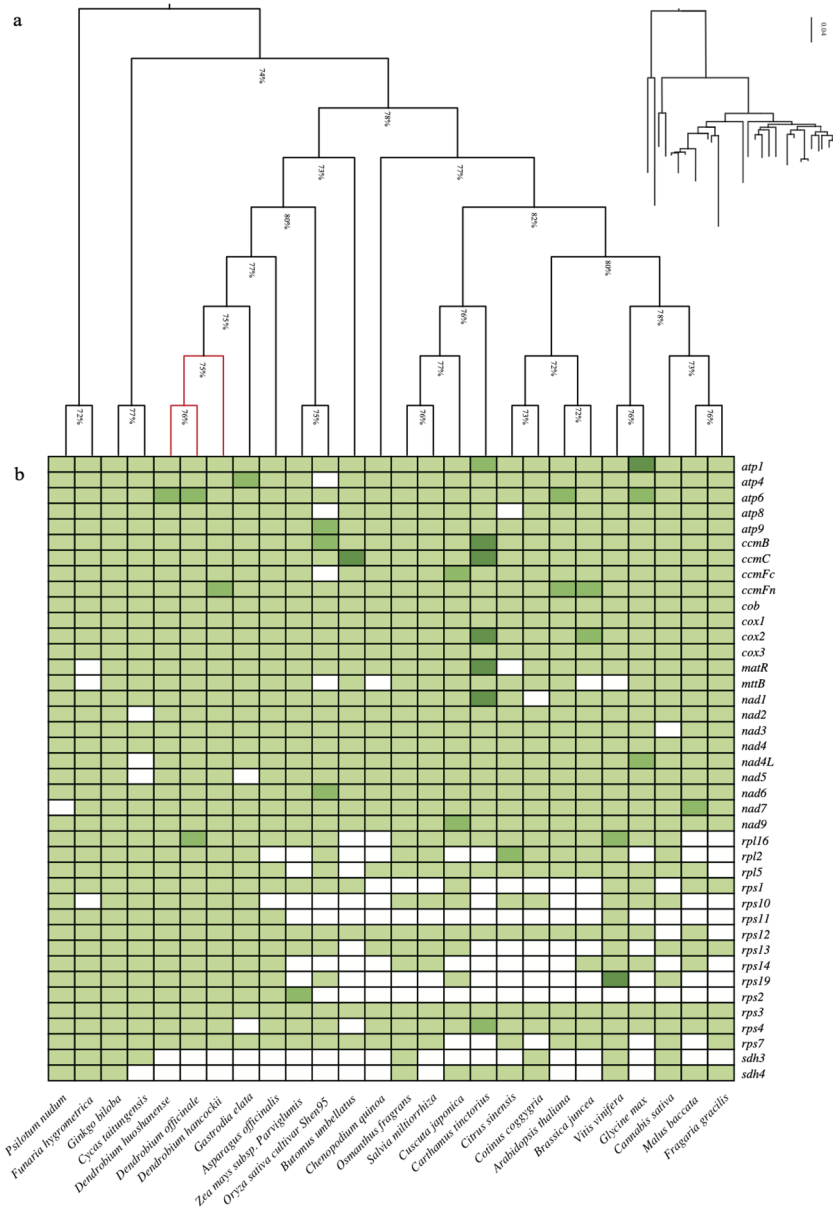
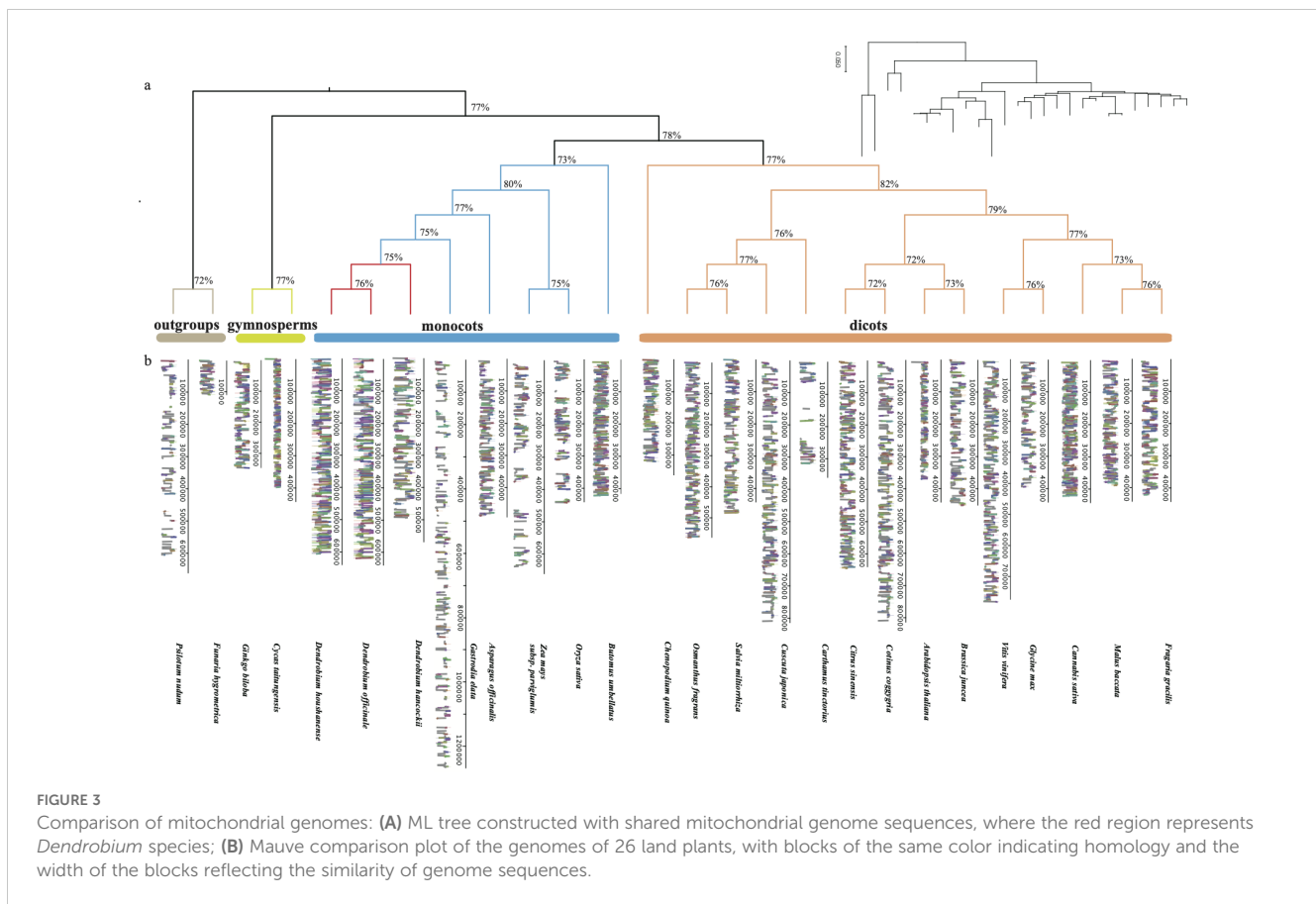


FIGURE 2 Comparison of the mitochondrial genes. **(A)** The NJ tree is constructed using the selected conserved mitochondrial genes set. **(B)** Distribution heatmap of mitochondrial genes across species. Colors indicate single-copy (light green), double-copy (green), triple-copy (dark green), and lost (white) genes.



primary research focus, *Dendrobium* spp., is grouped together with *Gastrodia elata* in the Orchidaceae branch. In addition to the Neighbor-Joining method (Figure 2A), we utilized the Maximum Likelihood method for phylogenetic tree analysis (Figure 3A). This method also classified *D. hancockii* as closely related to *D. officinale* and *D. huoshanense* within the Orchidaceae family of the Asparagales order. Moreover, phylogenetic trees constructed from tandem sequences of conserved genes robustly support the evolutionary relationships of plants, as established by the systematic taxonomy of the Angiosperm Phylogeny Group (APG IV).

In the field of mitochondrial genomics, as shown in Figure 3B, it is evident that the size of terrestrial plant mitochondria ranges from 100 kb to 1400 kb. Although significant variations in length have been observed in the orchid *Gastrodia elata*, MAUVE analysis indicates that most functional genes in the mitochondrial genomes of terrestrial plants are highly homologous, represented by similarly colored blocks in homologous regions. This results in a stable average genome size between 400 kb and 500 kb. However, mitochondrial genomes outside this range show significant differences in intergenic regions. Focusing on the eight orchid species (Supplementary Figure S2; Supplementary Table S7), both the number and length of homologous mitochondrial sequences increase, suggesting closer genetic relationships among these species. However, the connections between homologous regions reveal that the arrangement of these sequences differs, indicating extensive rearrangement events in orchid mitochondrial genomes.

Notably, *D. officinale*, *D. hancockii*, and *D. huoshanense* show the closest relationships, with species from the *Dendrobium* genus exhibiting high levels of homology. Some homologous segments cluster into longer homologous gene clusters. We identified five larger gene clusters: *ccmFn-atp1*, *atp4-nad4L-atp8*, *rpl16-ccmFn-rps3-rps19*, *mttB-rps13*, and *nad3-rps12*. The specific locations of these fragments have been marked within the corresponding genomic regions (Supplementary Figure S2).

3.5 The substitution rates of mitochondrial PCGs

Utilizing a set of 13 conserved protein-coding genes, in conjunction with *D. hancockii* as a reference, a total of 26 terrestrial plant species were examined for the nonsynonymous to synonymous substitution ratio (Ka/Ks). After excluding cases where Ka/Ks was unavailable (0 or NA), it is evident from Figure 4 and Supplementary Table S8 that, in the majority of protein-coding genes (PCGs), the Ka/Ks values are significantly less than 1. This observation indicates that these genes have been predominantly subjected to purifying selection during the course of evolution. In contrast, within *ccmB*, when compared to *D. hancockii*, *G. elata*, *F.gracilis*, *C.taitungensis*, *V.vinifera*, and *C.sativa* exhibit Ka/Ks values exceeding 1, suggesting positive selection acting on this gene. Additionally, positive selection is observed in the *ccFmn* of *F.gracilis*, *M.baccata*, *V.vinifera*, *C.coggygia*, *C.quinoa*, *C.japonica*,

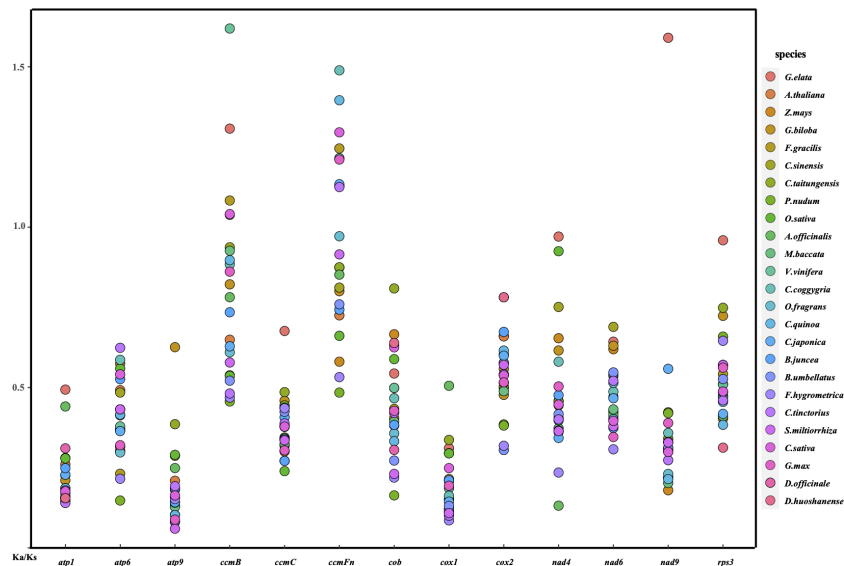


FIGURE 4

A dot plot of the Ka/Ks values of 13 protein-coding genes in mitogenomes of *D. hancockii* versus 25 species.

C. taitungensis, *C. sativa*, *G. max*, as well as the *nad9* of *G. elata*. The Ka/Ks values for *nad4* and *rps3* in *G. elata* are close to 1, implying a trend towards neutral selection. On the other hand, certain genes such as *atp9*, *ccmC*, and *cox1* exhibit relatively small Ka/Ks values, indicating their conserved nature. Within the *Dendrobium* genus, Ka/Ks values exhibit lower values, with 12 out of 13 PCGs showing strong purifying selection. This result suggests a lower background nucleotide mutation rate among *Dendrobium* species.

3.6 Sequence variation and genetic diversity

The sequence composed of high-quality SNP (Single Nucleotide Polymorphism) sites of the entire mitochondrial genome was utilized to assess nucleotide diversity. SNP calling was performed on resequenced data obtained from 103 samples of *D. hancockii*, collected from 23 distinct wild populations (Supplementary Table S9). This analysis resulted in the identification of a total of 19,101 high-quality SNP sites. According to Table 2, the nucleotide polymorphism range for the 23 wild populations varied from 0.1933 (PB population) to 0.09579 (WC population). Based on geographical distribution and morphological characteristics, they were divided into two groups. Group1, consisting of 15 populations, exhibited a higher P_i value (0.05660) compared to Group2, composed of 8 populations, with a P_i value of 0.02387. To further validate the reliability of this data, a segment comprising a total of 12,927 bp, derived from 8 nuclear single-copy fragments, confirmed these results. The nucleotide diversity within this nuclear data ranged from 0.01035 (GJ population) to 0.01678 (BH population). Notably, the P_i value for Group1 (0.01622) was higher than that of Group2 (0.01223).

In order to detect the genetic differences within the population, between the groups, within and between groups, the AMOVA analysis was performed using Arlequin software, and the results

showed that the variation within the population (64.56%) was greater than the variation between the total population (35.44%), and the genetic differentiation value was $\Phi_{ST} = 0.35440$ (Table 3). According to the geographical distribution of *D. hancockii* was divided into two groups, the AMOVA analysis of the two groups was as follows: the among group genetic variation was (53.32%) greater than the within-populations genetic variation (45.57%) and the within-group genetic variation (1.11%), and the corresponding genetic differentiation values were $\Phi_{CT} = 0.53320$, $\Phi_{ST} = 0.54425$ and $\Phi_{SC} = 0.02368$. Based on the mitochondrial SNPs, the variation of *D. hancockii* came from within the populations, and the differences among groups had a greater impact on the variation of *D. hancockii*, and its influence had a greater impact on the variation of *D. hancockii* than the differences among populations and within the group. The same analysis was also conducted on nuclear data, revealing that whether considering all 23 populations (70.73%) or the two groups (56.72%), the results indicate that a greater amount of genetic variation occurs inter-population. In summary, the variation of *D. hancockii* mainly came from among-group and inter-population.

Through the calculation of genetic differentiation values and gene flow between populations, we gain insights into the extent of genetic variation among groups. The genetic differentiation value (F_{ST}) was determined to be 0.42346, indicating a substantial degree of genetic differentiation between populations. This can also be observed from Figure 5 and Supplementary Table S10. When comparing two populations, each falling within the two groups delineated based on geographical distribution and morphological differences, there is a noticeable difference in F_{ST} among populations. The maximum F_{ST} is observed between the HX population and TL population ($F_{ST}=0.76487$). The gene flow ($N_m=0.0610$) between the total populations was found to be less than 1, suggesting limited gene flow between these populations. Permut analysis utilizing Permut software revealed G_{ST} and N_{ST}

TABLE 2 P_i by population based on mitochondrial snp and nuclear genetic data.

Population	Nucleotide diversity of mt-data	Nucleotide diversity of nul-data
Group1	0.05660	0.01622
BH	0.04734	0.01678
BT	0.08070	0.01191
DDH	0.04503	0.01550
HX	0.01983	0.01521
KD	0.04931	0.01417
LP	0.03789	0.01552
NS	0.06591	0.01389
SNJ	0.07005	0.01524
SY	0.05308	0.01380
TB	0.02298	0.01492
WC	0.09579	0.01048
WD	0.05359	0.01492
WDS	0.04936	0.01368
WLS	0.04960	0.01365
YA	0.07594	0.01169
Group2	0.02387	0.01223
GJ	0.01518	0.01035
JS	0.03116	0.01276
PB	0.01933	0.01111
TC	0.02359	0.01092
TL	0.01439	0.01121
WM	0.02327	0.01173
XY	0.03900	0.01095
ZF	0.01997	0.01178

P_i : Nucleotide diversity; Group1: BH (BaiHe), BT (BuTu), DDH (DaDuHe), HX (HuiXian), KD (KangDing), LP (LiangPing), NS (NingShan), SNJ (ShenNongJia), SY (ShanYang), TB (TianBao), WC (WenChuang), WD (WuDuo), WDS (WuDangShan), WLS (WuLingShan), YA (YaAn); Group2: GJ (GeJiu), JS (JianShui), PB (PingBian), TC (TengChong), TL (TianLin), WM (WangMo), XY (XingYi), ZF (ZhenFeng).

values of 0.00068 and 0.43327, respectively, for the total population. The greater N_{ST} value compared to G_{ST} indicates the presence of a related geographical structure within the *D. hancockii* population, based on mitochondrial SNPs.

3.7 Genetic structure analysis

The analysis of gene structure involved the fusion of mitochondrial genome SNPs and nuclear gene fragments data. To elucidate the genetic relationships among various populations and individuals, we employed several analytical approaches. According to the mitochondrial data in Figure 6, although the samples from

various populations did not cluster into a single branch, they were effectively divided into two branches with high support. One branch exclusively consisted of samples from populations in the Group1 region, while the other branch comprised samples from the Group2 region. Similar outcomes were confirmed in nuclear data.

Despite some topological differences in the phylogenetic relationships between the two data types, both successfully separated the 103 samples into two branches, with relationships within each branch appearing more intricate, possibly due to higher similarity within each branch. In terms of PCA analysis, the mitochondrial data distinctly differentiated 32 samples from Group1 and 101 samples from Group2. The first two principal components, PC1 and PC2, accounted for 20.02% of the total genome covariance. The nuclear data analysis corroborated this result, where PC1 and PC2, representing 16.64% of the total genome covariance, also supported the separation of the entire sample set into two clusters, with 32 samples from Group1 and 101 samples from Group2 forming distinct clusters.

Furthermore, utilizing fastSTRUCTURE for structure analysis, the optimal K value was determined to be 2. When K=2 (Figure 6), the results effectively grouped all individuals into two distinct subgroups, aligning with the patterns observed in PCA and neighbor-joining trees. Specifically, minimal genetic admixture was observed in only two populations, WM and TL, in the Group2 region, indicating limited gene flow between the two subgroups. Similar results were found in the nuclear gene data's structure analysis, where K=2 was the optimal value, dividing *Dendrobium* into two subgroups and detecting minimal gene flow among populations JC, TC, WC, and BH.

4 Discussion

4.1 Genomic architectural variation of *Dendrobium* mitochondria

In this study, we conducted the assembly of the mitochondrial genome of *D. hancockii*, spanning a length of 523,952 base pairs (bp). The mitogenome size is remarkably smaller than that of *D. officinale*, but larger than *D. gratiosissimum* (Wang et al., 2022). The assembly process utilized PacBio and Illumina sequence data, resulting in the identification of a monocyclic mitochondrial structure. The assembled genome consisted of a total of 40 coding genes and 41 non-coding RNAs, which were distributed across the genome. The GC content of the *D. hancockii* mitochondrial genome was found to be 44.28%, closely resembling that of related orchid species (42.9%-46.7%) (Ke et al., 2023). This observation suggests that the size and GC content of *Dendrobium* mitochondrial genomes has remained conserved throughout evolutionary processes. The acquisition of this result significantly contributes to the limited availability of mitochondrial genomic data in approximately 1500 *Dendrobium* species, thereby providing valuable molecular insights for future investigations.

In contrast to the conserved nature of mitochondrial sequence length and GC content in the *Dendrobium* genus, *D. hancockii* exhibits structural differences in its mitochondrial genome,

TABLE 3 AMOVA analysis of population based on nuclear genes and mitochondrial genome of *D. hancockii*.

Data Type	Source of variation	df	Sum of squares	Variance components	Percentage variation	Fixation indices
Mito-SNP	All regions					
	Among populations	22	30760.77	222.18667 Va	35.44%	
	Within populations	80	32380.017	404.75021 Vb	64.56%	$\Phi_{ST} = 0.35440$
	Total	102	63140.786	626.93688		
	Two groups					
	Among groups	1	21338.973	473.52963 Va	53.32%	$\Phi_{CT} = 0.53320$
	Among populations within groups	21	9421.797	9.81807 Vb	1.11%	$\Phi_{SC} = 0.02368$
	Within populations	80	32380.017	404.75021 Vc	45.57%	$\Phi_{ST} = 0.54425$
Total	102	63140.786	888.09791			
Nul-data	All regions					
	Among populations	22	5412.31	35.71594 Va	29.27%	
	Within populations	80	6905.35	86.31687 Vb	70.73%	$\Phi_{ST} = 0.29267$
	Total	102	12317.660	122.03282		
	Two groups					
	Among groups	1	2548.736	54.68597Va	35.93%	$\Phi_{CT} = 0.35932$
	Among populations within groups	21	2863.574	11.19037Vb	7.35%	$\Phi_{SC} = 0.11476$
	Within populations	80	6905.35	86.31688Vc	56.72%	$\Phi_{ST} = 0.43285$
Total	102	12317.66	152.19322			

df, degree of freedom; Φ_{CT} , genetic variation among groups; Φ_{SC} , genetic variation among populations within groups; Φ_{ST} , genetic variation within populations.

featuring a singular main ring structure. This deviates from the previously reported multi-conformational patterns observed in two other *Dendrobium* species. Considering the potential involvement of extensive repetitive segments, possibly facilitating frequent and reversible homologous recombination events within the species, our findings contribute to the understanding of the emergence of multi-conformational phenomena in mitogenomes (Sloan et al., 2012). We identified the presence of large fragments (>500 bp) in five dispersed repeats, constituting 1.70% of the total genome length, capable of inducing conformational changes. Furthermore, our investigation of mitochondrial genomes in the *Dendrobium* genus, specifically *D. officinale* and *D. huoshanense*, utilizing PacBio and Illumina sequence data, disclosed a comparable number and proportion of large fragment repeats within the main ring structure of *D. hancockii* (Wang et al., 2022). This outcome suggests a potential role of substantial repetitive segments in contributing to conformational variations in the mitogenomic structure of the *Dendrobium* genus.

Reflecting on the study of structure of mitochondrial genomes, it is evident that research on plant mitochondrial genomes emerged later than studies on animal and bacterial mitochondria (Unselid et al., 1997). Consequently, scholars investigating plant mitochondrial genes have largely followed the established model. Initially, mitochondria predominantly exhibited a monocyclic structure, as

observed in species such as *Camellia sinensis* (Li et al., 2023) and *Eucalyptus grandis* (Pinard et al., 2019). However, as research on the mitochondrial genomes of plants advanced, the concept of multichromosomal structures within mitochondrial genomes gained traction. This was first observed in the mitochondrial genome of *Cucumis sativus*, which was found to consist of three independent cyclic isomers (Alverson et al., 2011). Subsequently, the recognition of the diverse structural variability of mitochondria has gradually garnered acceptance among scholars. As demonstrated by our findings, the three species within the *Dendrobium* genus exhibit mitochondrial genomes with multiple structures. Variation in size and conformation within the mitochondrial genomes of related species is a common occurrence in plants (Sun et al., 2022). Consequently, we discourage the use of a specific structure to generalize the mitochondrial structure of *Dendrobium* plants, as both monocyclic and polycyclic structures possess inherent limitations. A monocyclic genomic map does not necessarily represent the absolute conformation of the mitochondrial genome within organisms, as alternative arrangements resulting from intramolecular or intermolecular recombination are plausible (Sloan, 2013). Such phenomena have been observed in *Photinia serratifolia* (Wang Y. et al., 2023). Similarly, polycyclic structures may arise due to changes in the mitochondrial genome prompted by environmental conditions. For instance, under stressful

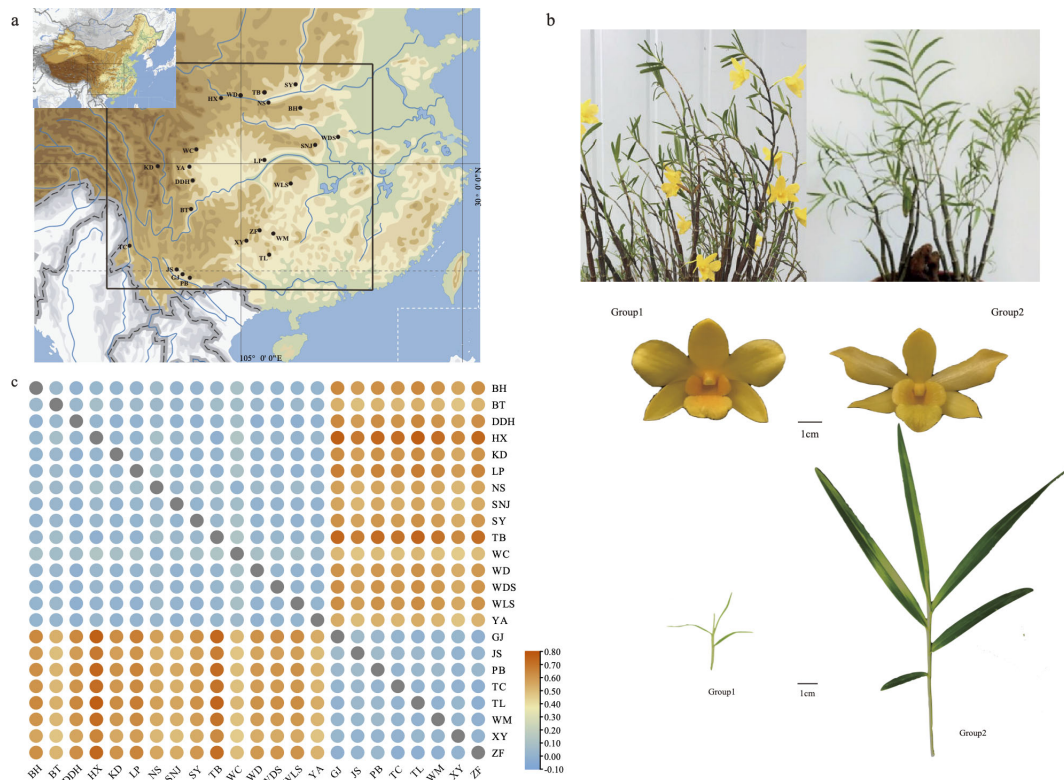


FIGURE 5

Phenotype, sampling locations, the genetic differentiation value of *D. hancockii* (A) Distribution of 23 wild populations; (B) Morphological phenotype information of *D. hancockii* in two groups; (C) The genetic differentiation value of 23 wild populations.

environments, plants may downregulate genes that typically inhibit recombination between short replicates of the mitochondrial genome. This mechanism may explain the extensive structural rearrangements and increased frequency of supercoiled DNA observed when plant tissues are cultured *in vitro* (Arrieta-Montiel and Mackenzie, 2011).

Consequently, we suggest the importance of understanding the diversity of mitochondrial genome conformations. Leveraging third-generation long-read DNA sequencing technology facilitates the high-quality assembly of plant mitochondrial genomes, facilitating more precise and sensitive detection of homologous recombination. This approach minimizes false positives introduced by PCR artifacts or limitations in read length inherent in next-generation sequencing methods (Schönegger et al., 2022). Through this approach, assembly discrepancies rather than true structural differences caused by sequencing issues are avoided. Hence, under the conditions of high sequencing depth, the assembly of mitochondrial genomes is poised to become a novel trend. Therefore, the utilization of high-depth sequencing techniques and next-generation sequencing technologies for the assembly of mitochondrial genomes is anticipated to emerge as a new trend. This approach aims to mitigate structural disparities caused by sequencing issues, offering rich and accurate data support for subsequent studies on mitochondrial genomes.

4.2 The mitochondrial genome mediating evolution through DNA fragment transfer.

The original mitochondrial genome is recognized for harboring a more comprehensive set of genes, reflecting functional evolution during its early evolution (Roger et al., 2017). However, owing to gene exchange between the nucleus and mitochondria, an increasing number of mitochondrial genes have translocated to the nuclear genome, leading to the depletion of genes from the mitochondrial genome itself (Mower, 2020). Similar phenomena are observed in organelle genomes as well (Kleine et al., 2009; Liu et al., 2013). In the mitochondrial genome of *D. hancockii*, this event occurred, with a total of 31,333 base pairs derived from the plastid, comprising approximately 5.98% of the total mitochondrial gene content. This proportion is consistent in plastid-originating mitochondria, ranging from 2 kb (0.5%) in *Vigna* genus plants to 113 kb (11.5%) in *Cucurbita* genus plants (Alverson et al., 2011). In the case of *D. hancockii*, the proportion closely aligns with that of *Apostasia shenzhenica* at 34 kb (5.12%) (Ke et al., 2023), in accordance with our expectations. The occurrence of similar evolutionary events in closely related species may potentially manifest in horizontal gene transfer at the genetic level. However, species-specific horizontal gene transfer events may occur. For example, the plastid-derived *rpl14* has been identified in the mitochondrial genome of *D. hancockii*, which is not observed in

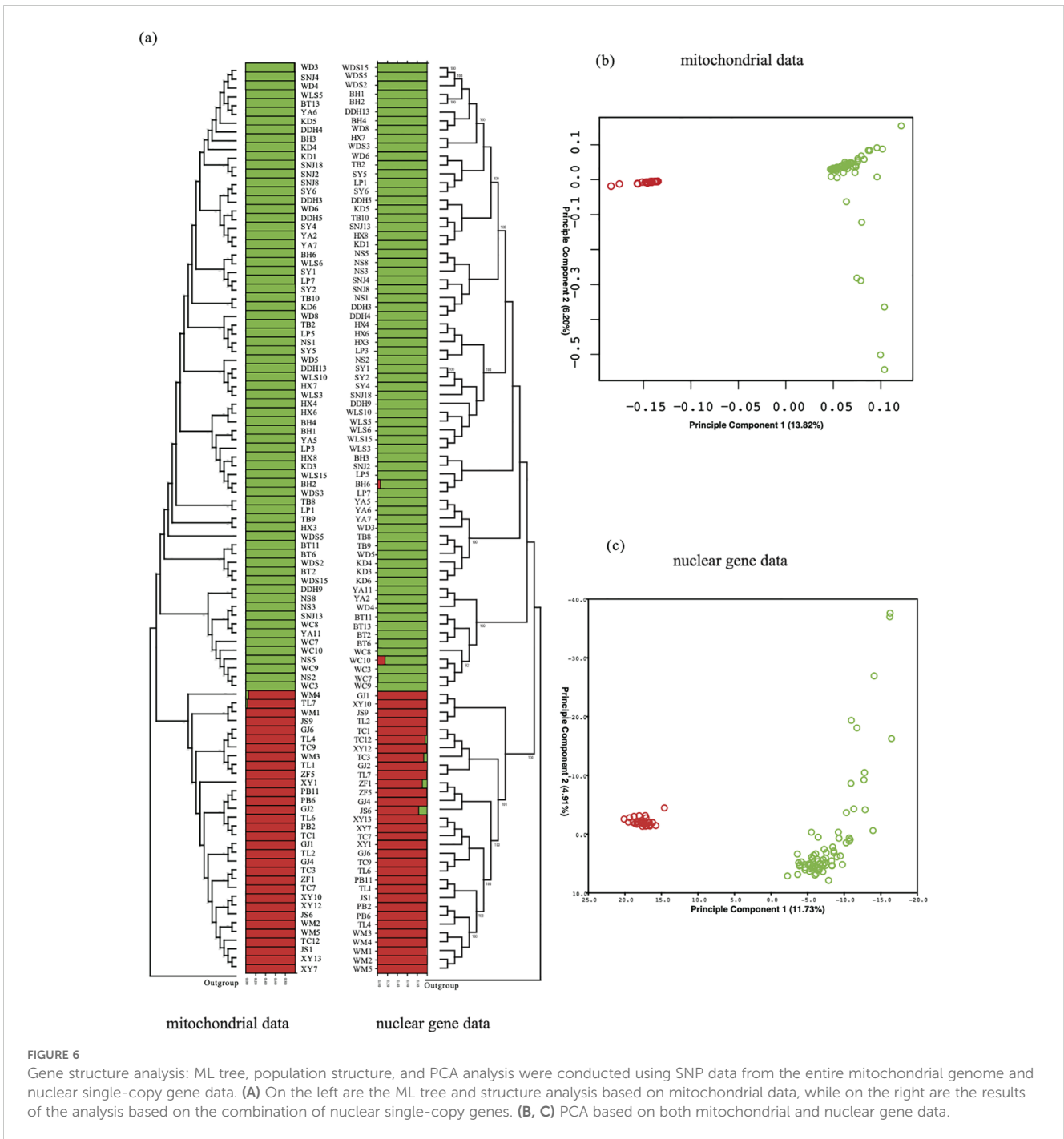


FIGURE 6 Gene structure analysis: ML tree, population structure, and PCA analysis were conducted using SNP data from the entire mitochondrial genome and nuclear single-copy gene data. (A) On the left are the ML tree and structure analysis based on mitochondrial data, while on the right are the results of the analysis based on the combination of nuclear single-copy genes. (B, C) PCA based on both mitochondrial and nuclear gene data.

other species within the same genus (Wang et al., 2022). The horizontal gene transfer of *rpl14* from plastids to mitochondria may have taken place after the divergence of *D. hancockii* species. Furthermore, the occurrence of a similar event in the mitochondrial genome of *Vitis vinifera* suggests a multi-origin nature of horizontal gene transfer of *rpl14*.

It is crucial to note that a considerable proportion of migratory sequences have experienced degradation in their integrity over the course of evolution. Consequently, a limited number of genes have been discovered in the mitochondrial genome, the majority of which are tRNA genes. The findings indicate that tRNA genes exhibit a higher degree of conservation within the mitochondrial

genome compared to the migratory patterns observed in protein-coding genes, emphasizing their crucial role within mitochondria. According to the analysis results, *D. hancockii* harbors nine tRNA genes derived from plastids, among which *trnN* and *trnM* exhibit relative conservation in terrestrial plants (Liao et al., 2018). This implies that these two tRNA genes may have undergone differentiation before the Orchidaceae divergence. Additionally, research indicates the loss of *trnV* in angiosperms, although it is present in certain dicotyledonous plants (Cheng et al., 2021). However, our study identifies the presence of *trnV*, suggesting frequent exchanges between mitochondrial and plastid genomes. Furthermore, the identification of *trnV* implies instances of gene

loss and subsequent regain within the mitochondrial and plastid genomes, indicating dynamic evolutionary processes at play.

As widely acknowledged, a distinctive feature of plant mitochondrial genome evolution is the frequent insertion of foreign DNA or the horizontal transfer of mitochondrial DNA, including the transfer of its own plastid and nuclear DNA (Mower et al., 2012; Rose, 2021). Due to the unpublished status of the complete genome of *D. hancockii*, we are currently unable to determine whether there is any transfer of nuclear genes in the mitochondrial genome. However, based on the findings from published genome studies of *G. elata* (Yuan et al., 2018) and *D. officinale* (Niu et al., 2021), it is evident that transfer events between nuclear, mitochondrial, and chloroplast genomes occur within the orchid family.

4.3 Comparative analysis of mitochondrial genomes in 26 land plant species

To unravel the evolutionary history encoded in the mitochondrial genome, a comparative analysis of mitochondrial genomes was conducted among 26 green plants, including *D. hancockii*. Throughout the course of mitochondrial evolution in land plants, numerous instances of replication, loss, and rearrangement events have been identified. Despite these evolutionary dynamics, a set of highly conserved genes essential for maintaining mitochondrial functionality persists throughout (Karakaidos and Rampias, 2020). Thirteen conserved protein-coding genes (*atp1*, 6, 9, *ccmB*, *ccmC*, *ccmFn*, *cob*, *cox1*, 2, *nad4*, 6, 9, *rps3*) were identified from mitochondrial genome datasets of mosses (*Funaria hygrometrica*), ferns (*Psilotum nudum*), gymnosperms (*Ginkgo biloba*, *Cycas revoluta*), and angiosperms (including parasitic plants). According to Sloan DB (Sloan et al., 2008; Pfeifer et al., 2019), the majority of these PCGs are associated with the mitochondrial proton-translocating ATP synthase complex, cytochrome c biogenesis, mature enzyme functions, and transport membrane proteins. These results indicate that conserved mitochondrial genes serve indispensable and irreplaceable functions. Therefore, in the non-synonymous/synonymous substitution rate of conservative PCGs, the majority of Ka/Ks ratios are less than 1. However, genes *ccmB* and *ccmFn*, associated with cytochrome c, exhibit a significantly higher frequency of positive selection, implying that mutations in these genes may bring about adaptive functional changes. Moreover, these changes are more advantageous for survival. Thus, this phenomenon has been observed in various mitochondrial genes across many plant species, such as in *Suaeda glauca* (Cheng et al., 2021) and *Vitex rotundifolia* (Yu et al., 2022). Additionally, these PCGs can be leveraged to infer evolutionary relationships among species through phylogenetic analysis, with the resulting topology aligning with the mainstream APG IV system (APG group, 2016), underscoring the significant potential of plant mitochondrial genes in studying species evolution.

Concerning gene loss, intriguing results suggest that mosses (37 PCGs), ferns (39 PCGs), and even gymnosperms (40/37 PCGs) exhibit a greater diversity of mitochondrial genes, hinting at more

frequent mitochondrial recombination events occurring after the evolution of angiosperms (Kan et al., 2021). At the species level within the family, higher levels of polymorphism are observed in the *rps* and *sdh* genes, as depicted by the gene cluster in Figure 3. This variation can be attributed to the deletion of mitochondrial genes resulting from the transfer of nuclear DNA and plastid DNA (Kleine et al., 2009). Further research into the comparative mitochondrial genomics of orchids has revealed varying degrees of gene rearrangement or loss (Supplementary Figure S2; Supplementary Table S7), which may be related to their adaptability to different environments. For example, among the eight orchid species, *Gastrodia* not only shows significant differences in mitochondrial structure but has also lost the relatively conserved genes *rps4* and *nad5*, possibly due to its mycoheterotrophic lifestyle. In orchids, closer phylogenetic relationships correspond to an increased number and length of homologous segments, which aligns with the principles of genetic evolution. Species within the same genus are more likely to retain a greater number of homologous segments during frequent structural changes. Through comparative analysis of the sequenced mitochondrial genomes of three species within the *Dendrobium* genus, we have not only revealed abundant homologous segments but also identified five conserved gene cluster fragments. This suggests that gene clusters may exhibit higher structural and sequence conservation. For instance, the identified “*rpl16-ccmFn-rps3-rps19*” gene cluster in this study displays significant similarity to the highly conserved “*rps19-rps3-rpl16*” gene cluster found in the fern *Adiantum capillus-veneris* (Bonavita and Regina, 2016). Therefore, we suggest that within the *Dendrobium* genus, mitochondrial diversity primarily exists in structural differences, manifested through variations such as repeat sequences, rearrangements, inversions, and other conformational changes. This pattern of rapid structural changes alongside slow sequence variation in the nucleotide background aligns with the mitochondrial evolution pattern observed in *Fragaria* species (Fan et al., 2022), and is consistent with research on *Paphiopedilum micranthum* within the Orchidaceae family (Yang et al., 2023).

4.4 The potential application of mitochondria in the analysis of genetic differentiation patterns

Firstly, the phylogeographic structure of *D. hancockii* and the major variations originating from among populations were confirmed. Analysis based on nuclear single-copy gene fragment data indicates that the N_{st} value, reflecting the inter-population genetic differentiation coefficient, surpasses the G_{st} value ($G_{st} = 0.00068$, $N_{st} = 0.43327$). This suggests the presence of a phylogeographic structure within *D. hancockii*, wherein populations that are closer geographically exhibit a closer genetic relationship. The majority of the genetic variation (64.56%) was found within populations, surpassing the among populations (35.44%). This finding aligns with results from related species, such as *D. officinale* (Hou et al., 2017) and *D. moniliforme* (Ye M.

et al., 2017), where differentiation among wild populations is lower than within populations. The high differentiation observed among wild populations of *D. hancockii* suggests that gene flow between populations is limited compared to other *Dendrobium* species. Although there are no significant differences in seed and pollen dispersal among species within the same genus, the wide distribution of *D. hancockii* populations likely contributes to this phenomenon. Geographic isolation and the discontinuous distribution of populations also hinder gene flow (Pfeifer et al., 2019). Our analysis of gene flow confirms this assumption, as indicated by the value of $nDNA: 0 < N_m (0.06103) < 1$, suggesting that overall, only some subpopulations of *D. hancockii* exhibit strong gene flow, which aligns with the results obtained from the phylogenetic analysis.

The results of phylogenetic reconstruction indicate that both mitochondrial and nuclear data support the division of the 103 *D. hancockii* samples into two groups. The population genetic structure analysis and PCA analysis (Figure 6) also support the viewpoint of dividing this species into two major lineages. The restricted gene flow between these lineages also implies that *D. hancockii* may have undergone outward expansion from two geographically distant refugia. By considering the influence of geological activity and glaciation on species expansion (Qiu et al., 2011), we suggest that the main factor contributing to the current phylogeographic pattern in *D. hancockii* is an adaptive response to ancient geological and climatic changes. Based on existing distribution information, we find that the division of wild *D. hancockii* populations into two major groups aligns with the “Tanaka-Kaiyong Line” division in the plant district flora (Ye J. et al., 2017). This alignment lends support to the phylogeographic division of *D. hancockii* and the boundary between the Sino-Himalayan forest subkingdom and Sino-Japanese forest subkingdom. This phenomenon has also been observed in other species, including *Sophora davidii* (Fan et al., 2013), *Tacca chantrieri* (Zhao and Zhang, 2015), *Leucomeris decora*, and *Noelia insignis* (Zhao and Gong, 2015).

The integration of mitochondrial genome analysis with studies on phylogeography and genetic structures in animals is a well-established practice, as demonstrated in research on scorpions (Moreno-Carmona et al., 2023) and *Parasa sinica* (Zhang et al., 2023). This study marks the inaugural application of comprehensive mitochondrial genome SNP analysis for assessing genetic diversity and population structures in wild populations of *D. hancockii*. Analyses, encompassing genetic diversity, population genetic structure, and PCA, align closely with nuclear gene data, suggesting the potential utility of the mitochondrial genome in elucidating genetic differentiation patterns. Compared to nuclear gene data analysis, the results of mitochondrial analysis have confirmed its potential for application at the population level. However, it is undeniable that the use of mitochondrial genomes at the population level in species with narrow habitats and single distributions may present risks. Beyond the limitation of maternal inheritance in organellar genomes (Camus et al., 2022), the lower substitution rate of mitochondrial genomes (Drouin et al., 2008) may pose challenges when applied to endangered species with low diversity. However, due to the trans-regional distribution of *D.*

hancockii, the impeded gene flow between populations caused by the pollen structure of Orchidaceae, and the climatic and geographic differences between northern and southern populations, the ‘good performance’ of the mitochondrial genome in *D. hancockii* may be revised by species-specific advantages. Therefore, the performance of the mitochondrial genome in population structure and genetic diversity needs further supplementary application across more species.

At the same time, the mitochondrial genome also has its unique advantages. Considering that certain heterotrophic parasitic plants, such as *G. elata*, have undergone the loss of photosynthetic genes (Xu et al., 2021), the mitochondrial genome emerges as a preferable choice for specific analyses. Studies reveal discrepancies in the phylogenetic relationships between plastid and mitotic genomes in *Aldrovanda vesiculosa* (Štorchová et al., 2023) and *Dendrobium* plants (Wang M. et al., 2023), suggesting unique applications and directions for different organelle genomes. Even within the *Potentilla* genus, the mitochondrial genome exhibits superior capabilities in constructing phylogenetic relationships compared to the chloroplast genome (Xue et al., 2023).

Current studies focused on plant phylogeography, genetic structure, and phylogenetics predominantly employ research methods that combine nuclear genes with organelle genes (Morris and Shaw, 2018). This approach has become mainstream, providing a more accurate reflection of incomplete lineage sorting (Steenwyk et al., 2023) or hybridization events (Fleming et al., 2023) during species formation, based on inconsistencies between nuclear and cytoplasmic genetic materials. The advent of mitochondrial genome analysis will enhance our exploration of plant genetic diversity, population structure, phylogeography, and systematics from a distinct perspective.

5 Conclusions

In this study, a comprehensive investigation was conducted on the rare plant *D. hancockii* through the application of second and third-generation sequencing methods. This approach allowed for *de novo* splicing assembly annotation, resulting in the successful acquisition of a circular and complete mitochondrial genome. The assembled genome revealed a cyclic structure and provided crucial information on 40 coding proteins, along with 37 tRNA and 4 rRNA genes. These findings significantly contribute to the enrichment of mitochondrial data within the *Dendrobium* species. Moreover, the study included a comparative analysis of the mitochondrial genome, shedding light on its evolutionary position and potential evolutionary events. The study also conducted comparative analyses of the mitochondrial genome, revealing its evolutionary position and potential evolutionary events. During the evolution of the *Dendrobium* genus, the mitochondrial structure demonstrates a pattern of variability while the sequence genes remain relatively conserved. Subsequently, a comprehensive population genetic structure analysis was conducted, utilizing the single nucleotide polymorphism sites within the mitochondrial group of 103 wild individuals. This data, combined with genetic structure analysis of nuclear single-copy gene sequences, led to the classification of *D.*

hancockii from 23 distinct populations into two distinct subgroups. Furthermore, the study inferred that the genetic differentiation pattern of *D. hancockii* may be attributed to ancient geological and climatic factors. This elucidates the potential of the mitochondrial genome in genealogical geography and contributes to a comprehensive understanding of the species' evolutionary history.

Data availability statement

The datasets presented in this study can be found in online repositories. The data that support the findings of this study are openly available in the Science Data Bank at <https://doi.org/10.57760/sciencedb.15636> & <https://doi.org/10.57760/sciencedb.j00143.00099>. The assembled mitochondrial genome sequences were submitted to the DDBJ with accession numbers: LC771085.

Author contributions

ZH: Formal analysis, Investigation, Software, Visualization, Writing – original draft. MW: Investigation, Software, Validation, Visualization, Writing – review & editing. YJ: Formal analysis, Methodology, Writing – review & editing. QX: Project administration, Resources, Supervision, Writing – review & editing. WL: Methodology, Project administration, Writing – review & editing. ZN: Conceptualization, Methodology, Project administration, Resources, Supervision, Visualization, Writing – review & editing. XD: Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research

References

- Alverson, A. J., Rice, D. W., Dickinson, S., Barry, K., and Palmer, J. D. (2011). Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* 23, 2499–2513. doi: 10.1105/tpc.111.087189
- Arrieta-Montiel, M. P., and Mackenzie, S. A. (2011). "Plant mitochondrial genomes and recombination," in Ed. F. Kempken. *Plant Mitochondria*, vol. 1. (New York, NY: Springer). doi: 10.1007/978-0-387-89781-3_3
- 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
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573
- Bhardwaj, K., Bhargav, R., Patocka, J., Sharma, R., Navratilova, Z., Oleksak, P., et al. (2024). Dendrobine: A neuroprotective sesquiterpene alkaloid for the prevention and treatment of diseases: A review. *Mini Rev. Med. Chem.* 24, 1395–1408. doi: 10.2174/0113895575274314240125105120
- Bonavita, S., and Regina, T. M. (2016). The evolutionary conservation of *rps3* introns and *rps19-rps3-rpl16* gene cluster in *Adiantum capillusveneris* mitochondria. *Curr. Genet.* 62, 173–184. doi: 10.1007/s00294-015-0512-z
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10, 421. doi: 10.1186/1471-2105-10-421
- Camus, M. F., Alexander-Lawrie, B., Sharbrough, J., and Hurst, G. D. D. (2022). Inheritance through the cytoplasm. *Heredity* 129, 31–43. doi: 10.1038/s41437-022-00540-2
- Chen, J., Guan, R., Chang, S., Du, T., Zhang, H., and Xing, H. (2011). Substoichiometrically different mitotypes coexist in mitochondrial genomes of *Brassica napus* L. *PLoS One* 6, e17662. doi: 10.1371/journal.pone.0017662
- Cheng, Y., He, X., Priyadarshani, S. V. G. N., Wang, Y., Li, Y., Shi, C., et al. (2021). Assembly and comparative analysis of the complete mitochondrial genome of *Suaeda glauca*. *BMC Genomics* 22, 167. doi: 10.1186/s12864-021-07490-9
- Chevigny, N., Schatz-Daas, D., Lotfi, F., and Gualberto, J. M. (2020). DNA repair and the stability of the plant mitochondrial genome. *Int. J. Mol. Sci.* 21, 328. doi: 10.3390/ijms21010328
- Choi, I. S., Schwarz, E. N., Ruhlman, T. A., Khiyami, M. A., Sabir, J. S. M., Hajarrah, N. H., et al. (2019). Fluctuations in *Fabaceae* mitochondrial genome size and content are both ancient and recent. *BMC Plant Biol.* 19, 448. doi: 10.1186/s12870-019-2064-8
- Darling, A. C., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498. doi: 10.1038/ng.806

was partially supported by the Natural Science Foundation of China (#31670330, #32070353) and Foundation of Key Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization (OU202205).

Acknowledgments

We would like to thank the College of Life Sciences, Nanjing Normal University for supporting this work. We are also grateful to the Jiangsu Provincial Engineering Research Center for Technical Industrialization for *Dendrobiums* for technical support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.2024.1469267/full#supplementary-material>

- Drouin, G., Daoud, H., and Xia, J. (2008). Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol.* 49, 827–831. doi: 10.1016/j.ympev.2008.09.009
- Earl, D. A., and vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7
- Falush, D., Stephens, M., and Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587. doi: 10.1093/genetics/164.4.1567
- Fan, D. M., Yue, J. P., Nie, Z. L., Li, Z. M., Comes, H. P., and Sun, H. (2013). Phylogeography of *Sophora davidii* (Leguminosae) across the ‘Tanaka-Kaiyong Line’, an important phylogeographic boundary in Southwest China. *Mol. Ecol.* 22, 4270–4288. doi: 10.1111/mec.12388
- Fan, W., Liu, F., Jia, Q., Du, H., Chen, W., Ruan, J., et al. (2022). Fragaria mitogenomes evolve rapidly in structure but slowly in sequence and incur frequent multinucleotide mutations mediated by microinversions. *New Phytol.* 236, 745–759. doi: 10.1111/nph.18334
- Fleming, J. F., Valero-Gracia, A., and Struck, T. H. (2023). Identifying and addressing methodological incongruence in phylogenomics: A review. *Evol. Appl.* 16, 1087–1104. doi: 10.1111/eva.13565
- Gui, S., Wu, Z., Zhang, H., Zheng, Y., Zhu, Z., Liang, D., et al. (2016). The mitochondrial genome map of *Nelumbo nucifera* reveals ancient evolutionary features. *Sci. Rep.* 6, 30158. doi: 10.1038/srep30158
- Hou, B. W., Luo, J., Zhang, Y., Niu, Z. T., Xue, Q. Y., and Ding, X. Y. (2017). Iteration expansion and regional evolution: phylogeography of *Dendrobium officinale* and four related taxa in southern China. *Sci. Rep.* 7, 43525. doi: 10.1038/srep43525
- Hou, Z. Y., Jiang, Y., Li, C., Niu, Z. T., and Ding, X. Y. (2019). The complete chloroplast genome sequence of a Chinese endemic plant *Dendrobium hancockii* Rolfe (Orchidaceae). *Mitochondrial DNA B Resour.* 4, 3860–3861. doi: 10.1080/23820259.2019.1687036
- Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., et al. (2017). ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. *Genome Res.* 27, 768–777. doi: 10.1101/gr.214346.116
- Kan, S. L., Shen, T. T., Ran, J. H., and Wang, X. Q. (2021). Both Conifer II and Gnetales are characterized by a high frequency of ancient mitochondrial gene transfer to the nuclear genome. *BMC Biol.* 19, 146. doi: 10.1186/s12915-021-01096-z
- Karakaidos, P., and Rampias, T. (2020). Mitonuclear interactions in the maintenance of mitochondrial integrity. *Life-Basel* 10, 173. doi: 10.3390/life10090173
- 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
- Ke, S. J., Liu, D. K., Tu, X. D., He, X., Zhang, M. M., Zhu, M. J., et al. (2023). *Apostasia* Mitochondrial genome analysis and monocot mitochondria phylogenomics. *Int. J. Mol. Sci.* 24, 7837. doi: 10.3390/ijms24097837
- Kleine, T., Maier, U. G., and Leister, D. (2009). DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu. Rev. Plant Biol.* 60, 115–138. doi: 10.1146/annurev.arplant.043008.092119
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633
- Lai, C. H., Huo, C. Y., Xu, J., Han, Q. B., and Li, L. F. (2024). Critical review on the research of chemical structure, bioactivities, and mechanism of actions of *Dendrobium officinale* polysaccharide. *Int. J. Biol. Macromol.* 263, 130315. doi: 10.1016/j.jbiomac.2024.130315
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* 1, 3997. doi: 10.48550/arXiv.1303.3997
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, J., Tang, H., Luo, H., Tang, J., Zhong, N., and Xiao, L. (2023). Complete mitochondrial genome assembly and comparison of *Camellia sinensis* var. *Assamica* cv. *Duntisa*. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1117002
- Liao, X., Zhao, Y., Kong, X., Khan, A., Zhou, B. J., Liu, D. M., et al. (2018). Complete sequence of kenaf (*Hibiscus cannabinus*) mitochondrial genome and comparative analysis with the mitochondrial genomes of other plants. *Sci. Rep.* 8, 12714. doi: 10.1038/s41598-018-30297-w
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Liu, G., Cao, D., Li, S., Su, A., Geng, J., Grover, C. E., et al. (2013). The complete mitochondrial genome of *Gossypium hirsutum* and evolutionary analysis of higher plant mitochondrial genomes. *PLoS One* 8, e69476. doi: 10.1371/journal.pone.0069476
- Lloyd, E. D., Hlongwane, T. T., Joshi, S. V., and Riaño, P. D. M. (2019). The sugarcane mitochondrial genome: assembly, phylogenetics and transcriptomics. *PeerJ* 7, e7558. doi: 10.7717/peerj.7558
- Lohse, M., Bolger, A. M., Nagel, A., Fernie, A. R., Lunn, J. E., Stitt, M., et al. (2012). RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res.* 40, W622–W627. doi: 10.1093/nar/gks540
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* 1, 18. doi: 10.1186/2047-217X-1-18
- Ma, J., Wang, S., Zhu, X., Sun, G., Chang, G., Li, L., et al. (2022). Major episodes of horizontal gene transfer drove the evolution of land plants. *Mol. Plant* 15, 857–871. doi: 10.1016/j.molp.2022.02.001
- Meleshko, D., Mohimani, H., Tracanna, V., Hajirasouliha, I., Medema, M. H., Korobeynikov, A., et al. (2019). BiosyntheticSPades: reconstructing biosynthetic gene clusters from assembly graphs. *Genome Res.* 29, 1352–1362. doi: 10.1101/gr.243477.118
- Meng, Y., Liang, Z., Zhang, L., Hu, L., Fu, J., Wei, G., et al. (2023). Identification of three types of O-glycosylated flavonoids in *Dendrobium loddigesii*, *Dendrobium primulinum*, *Dendrobium crepidatum*, *Dendrobium porphyrochilum*, and *Dendrobium hancockii* using mass spectrometry. *Rapid Commun. Mass Spectrom.* 37, e9421. doi: 10.1002/rcm.9421
- Møller, I. M., Rasmusson, A. G., and Van, Aken, O. (2021). Plant mitochondria - past, present and future. *Plant J.* 108, 912–959. doi: 10.1111/tpj.15495
- Moreno-Carmona, M., Montaña-Lozano, P., Quiroga, C. F. P., and Baeza, J. A. (2023). Comparative analysis of mitochondrial genomes reveals family-specific architectures and molecular features in scorpions (Arthropoda: Arachnida: Scorpiones). *Gene* 859, 147189. doi: 10.1016/j.gene.2023.147189
- Morris, A. B., and Shaw, J. (2018). Markers in time and space: A review of the last decade of plant phylogeographic approaches. *Mol. Ecol.* 27, 2317–2333. doi: 10.1111/mec.14695
- Mower, J. P. (2009). The PREP suite: predictive RNA editors for plant mitochondrial genes, chloroplast genes and user-defined alignments. *Nucleic Acids Res.* 37, W253–W259. doi: 10.1093/nar/gkp337
- Mower, J. P. (2020). Variation in protein gene and intron content among land plant mitogenomes. *Mitochondrion* 53, 203–213. doi: 10.1016/j.mito.2020.06.002
- Mower, J. P., Jain, K., and Hepburn, N. J. (2012). “The role of horizontal transfer in shaping the plant mitochondrial genome,” in *Adv Bot Res*, vol. 63. Ed. L. Marechal-Drouard (New York, NY: Academic Press). doi: 10.1016/B978-0-12-394279-1.00003-X
- Nguyen, L. T., Schmidt, H. A., Haeseler, A. V., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Niu, Z., Zhu, F., Fan, Y., Li, C., Zhang, B., Zhu, S., et al. (2021). The chromosome-level reference genome assembly for *Dendrobium officinale* and its utility of functional genomics research and molecular breeding study. *Acta Pharm. Sin. B.* 11, 2080–2092. doi: 10.1016/j.apsb.2021.01.019
- Pan, C., Chen, S., Chen, Z., Li, Y., Liu, Y., Zhang, Z., et al. (2022). Assessing the geographical distribution of 76 *Dendrobium* species and impacts of climate change on their potential suitable distribution area in China. *Environ. Sci. Pollut. Res.* 29, 20571–20592. doi: 10.1007/s11356-021-15788-0
- Patterson, N., Price, A. L., and Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genet.* 2, e190. doi: 10.1371/journal.pgen.0020190
- Pfeifer, M., Schatz, B., Xavier, P. F., Nicodemo, G. P., Michael, F. F., Pete, D. C., et al. (2019). Phylogeography and genetic structure of the orchid *himantoglossum hircinum* (L.) spreng. across its european central–marginal gradient. *J. Biogeogr.* 36, 2353–2365. doi: 10.1111/j.1365-2699.2009.02168.x
- Pinard, D., Myburg, A. A., and Mizrahi, E. (2019). The plastid and mitochondrial genomes of *Eucalyptus grandis*. *BMC Genomics* 20, 132. doi: 10.1186/s12864-019-5444-4
- Preuten, T., Cincu, E., Fuchs, J., Zoschke, R., Liere, K., and Börner, T. (2010). Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. *Plant J.* 64, 948–959. doi: 10.1111/j.1365-3113X.2010.04389.x
- Qiu, Y. X., Fu, C. X., and Comes, H. P. (2011). Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world’s most diverse temperate flora. *Mol. Phylogenet. Evol.* 59, 225–244. doi: 10.1016/j.ympev.2011.01.012
- Raj, A., Stephens, M., and Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 197, 573–589. doi: 10.1534/genetics.114.164350
- Roger, A. J., Muñoz-Gómez, S. A., and Kamikawa, R. (2017). The origin and diversification of mitochondria. *Curr. Biol.* 27, R1177–R1192. doi: 10.1016/j.cub.2017.09.015
- Rose, R. J. (2021). Contribution of massive mitochondrial fusion and subsequent fission in the plant life cycle to the integrity of the mitochondrion and its genome. *Int. J. Mol. Sci.* 22, 5429. doi: 10.3390/ijms22115429
- Salmela, L., and Rivals, E. (2014). LoRDEC: accurate and efficient long read error correction. *Bioinformatics* 30, 3506–3514. doi: 10.1093/bioinformatics/btu538
- Schattner, P., Brooks, A. N., and Lowe, T. M. (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33, W686–W689. doi: 10.1093/nar/gki366
- Schönegger, E. S., Crisp, A., Müller, M., Fertl, J., Serdjukow, S., Croce, S., et al. (2022). Click chemistry enables rapid amplification of full-length reverse transcripts for long-read third generation sequencing. *Bioconjugate Chem.* 33, 1789–1795. doi: 10.1021/acs.bioconjchem.2c00353

- Singh, V. K., Mangalam, A. K., Dwivedi, S., and Naik, S. (1998). Primer premier: program for design of degenerate primers from a protein sequence. *BioTechniques* 24, 318–319. doi: 10.2144/98242pf02
- Sloan, D. B. (2013). One ring to rule them all? Genome sequencing provides new insights into the 'master circle' model of plant mitochondrial DNA structure. *New Phytol.* 200, 978–985. doi: 10.1111/nph.12395
- Sloan, D. B., Alverson, A. J., Chackalovcak, J. P., Wu, M., McCauley, D. E., Palmer, J. D., et al. (2012). Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* 10, e1001241. doi: 10.1371/journal.pbio.1001241
- Sloan, D. B., Wu, Z., and Sharbrough, J. (2008). Correction of persistent errors in *Arabidopsis* reference mitochondrial genomes. *Plant Cell* 30, 525–527. doi: 10.1105/tpc.18.00024
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Steenwyk, J. L., Li, Y., Zhou, X., Shen, X. X., and Rokas, A. (2023). Incongruence in the phylogenomics era. *Nat. Rev. Genet.* 24, 834–850. doi: 10.1038/s41576-023-00620-x
- Štorchová, H., Krüger, M., Small, I., Adamec, L., Nevill, P. G., Howell, K. A., et al. (2023). Global mitochondrial and chloroplast genome diversity in the threatened aquatic carnivorous plant *Aldrovanda vesiculosa*. *Aquat. Bot.* 191, 103742. doi: 10.1016/j.aquabot.2023.103742
- 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
- Tamura, K., Nei, M., and Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11030–11035. doi: 10.1073/pnas.0404206101
- The angiosperm phylogeny group (2016). An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* 181, 1–20. doi: 10.1111/boj.12385
- Unsold, M., Marienfeld, J. R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.* 15, 57–61. doi: 10.1038/ng0197-57
- Waltz, F., and Giegé, P. (2020). Striking diversity of mitochondria-specific translation processes across eukaryotes. *Trends Biochem. Sci.* 45, 149–162. doi: 10.1016/j.tibs.2019.10.004
- Wang, Y., Chen, S., and Chen, J. (2023). Characterization and phylogenetic analysis of the complete mitochondrial genome sequence of *Photinia serratifolia*. *Sci. Rep.* 13, 770. doi: 10.1038/s41598-022-24327-x
- Wang, M. T., Hou, Z. Y., Li, C., Yang, J. P., Niu, Z., Xue, Q. Y., et al. (2022). Rapid structural evolution of *Dendrobium* mitogenomes and mito-nuclear phylogeny discordances in *Dendrobium* (Orchidaceae). *J. Syst. Evol.* 61, 790–805. doi: 10.1111/jse.12912
- Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49. doi: 10.1093/nar/gkr1293
- Wang, M., Yu, W., Yang, J., Hou, Z., Li, C., Niu, Z., et al. (2023). Mitochondrial genome comparison and phylogenetic analysis of *Dendrobium* (Orchidaceae) based on whole mitogenomes. *BMC Plant Biol.* 23, 586. doi: 10.1186/s12870-023-04618-9
- Xu, H. J., Hou, B. W., Zhang, J. Z., Tian, M., Yuan, Y. H., Niu, Z. T., et al. (2012). Detecting adulteration of *Dendrobium officinale* by real-time PCR coupled with ARMS. *Int. J. Food Sci. Tech.* 47, 1695–1700. doi: 10.1111/j.1365-2621.2012.03023.x
- 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
- Xue, T. T., Janssens, S. B., Liu, B. B., and Yu, S. X. (2023). Phylogenomic conflict analyses of the plastid and mitochondrial genomes via deep genome skimming highlight their independent evolutionary histories: A case study in the cinquefoil genus *Potentilla sensu lato* (Potentilleae, Rosaceae). *Mol. Phylogenet. Evol.* 190, 107956. doi: 10.1016/j.ympev.2023.107956
- 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
- Ye, M., Liu, W., Xue, Q., Hou, B., Luo, J., and Ding, X. (2017). Phylogeography of the endangered orchid *Dendrobium moniliforme* in East Asia inferred from chloroplast DNA sequences. *Mitochondrial DNA A.* 28, 880–891. doi: 10.1080/24701394.2016.1202942
- Ye, J. W., Zhang, Y., and Wang, X. J. (2017). Phylogeographic breaks and the mechanisms of their formation in the Sino-Japanese floristic region. *Chin. J. Plan. Ecol.* 41, 1003–1019. doi: 10.17521/cjpe.2016.0388
- You, L., Yang, L., Liu, F., Kang, L., Chen, H., Yang, B., et al. (2022). Assembly and marker analysis of mitochondrial genomes provide insights into origin, evolution and spread of *Brassica juncea* (L.) Czern. et Coss. *Crop J.* 10, 010–003. doi: 10.1016/j.cj.2021.10.004
- Yu, X., Duan, Z., Wang, Y., Zhang, Q., and Li, W. (2022). Sequence Analysis of the Complete Mitochondrial Genome of a Medicinal Plant, *Vitex rotundifolia* Linnaeus f. (Lamiales: Lamiaceae). *Genes* 13, 839. doi: 10.3390/genes13050839
- 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, S. P., Zhang, J., Xu, J., Wang, Q. H., Ye, Y., Wang, G., et al. (2023). Complete mitochondrial genome of *Parasa sinica*: New insights into the phylogeny of Limacodidae. *Heliyon* 9, e21375. doi: 10.1016/j.heliyon.2023.e21375
- 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. *Hortic. Res.* 9, uhac220. doi: 10.1093/hr/uhac220
- Zhao, Y. J., and Gong, X. (2015). Genetic divergence and phylogeographic history of two closely related species (*Leucomeris decora* and *Nouelia insignis*) across the 'Tanaka Line' in Southwest China. *BMC Evol. Biol.* 15, 134. doi: 10.1186/s12862-015-0374-5
- Zhao, Y., and Zhang, L. (2015). The phylogeographic history of the self-pollinated herb *Tacca chantrieri* (Dioscoreaceae) in the tropics of mainland Southeast Asia. *Biochem. Syst. Ecol.* 58, 139–148. doi: 10.1016/j.bse.2014.11.011