

Article

Complete Chloroplast Genome and Phylogenomic Analysis of *Davallia trichomanoides* (Polypodiaceae)

Yingying Wang ¹, Ziqi Xiang ¹, Keqin Liu ¹, Yuan Lin ^{2,*} and Siyuan Dong ^{3,*}

¹ College of Liangshan, Lishui University, Lishui 323000, China; wangyy@lsu.edu.cn (Y.W.); xzq010507@lsu.edu.cn (Z.X.); lkq@lsu.edu.cn (K.L.)

² Institute of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

³ China National Research Institute of Food and Fermentation Industries Co., Ltd., Beijing 100015, China

* Correspondence: yuan.lin@siat.ac.cn (Y.L.); f471449781@163.com (S.D.)

Abstract

Background/Objectives: Chloroplast genomes (plastomes) are valuable for fern systematics, yet the epiphytic lineages have remained underexplored. **Methods:** The *Davallia trichomanoides* plastome was de novo assembled from Illumina data and annotated. **Results:** The plastome measures 154,217 bp with a GC content of 40.82% and contains 115 genes. Comparative analysis reveals two inverted repeat (IR) size classes (~24.0–24.6 kb vs. ~27.4–27.5 kb) and lineage-specific shifts at the IR junctions. For instance, the *ndhF* gene remains in the small single copy (SSC) region in *D. trichomanoides* and *Drynaria acuminata*, but it crosses into the IRb region in other species. We observed nucleotide diversity hotspots in the large single copy (LSC) and SSC regions. The IR regions are highly conserved. The ratios of nonsynonymous to synonymous substitutions (Ka/Ks) are mostly less than 1, indicating purifying selection. Phylogenetic analysis places *D. trichomanoides* as the sister to *D. acuminata*. **Conclusions:** This study highlights the stable plastome structure of *D. trichomanoides* and identifies candidate loci for barcoding. It also supports the stable placement of *Davallia* within the epiphytic Polypodiaceae.

Keywords: Polypodiaceae; *Davallia trichomanoides*; chloroplast genome; comparative plastome; genetic diversity; phylogenetic analysis



Academic Editor: Zhiqiang Wu

Received: 14 October 2025

Revised: 28 October 2025

Accepted: 28 October 2025

Published: 1 November 2025

Citation: Wang, Y.; Xiang, Z.; Liu, K.; Lin, Y.; Dong, S. Complete Chloroplast Genome and Phylogenomic Analysis of *Davallia trichomanoides* (Polypodiaceae). *Genes* **2025**, *16*, 1310. <https://doi.org/10.3390/genes16111310>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In land plants, chloroplast genomes (plastomes) are usually circular and divided into a large single-copy (LSC), a small single-copy (SSC), and two inverted repeats (IRs). This stable structure underlies the broad application of plastid DNA in systematics and comparative genomics, as well as in various applied contexts ranging from phylogenetic studies [1] to population-level analyses [2]. Although structurally conservative, most changes in length or gene order reflect small IR expansions or contractions, or local inversions, and the IRs are usually the most conserved parts [3]. This IR-linked stability and lower substitution rate are well documented across land-plant plastomes [4].

Ferns provide a complementary macroevolutionary view to seed plants. Broad surveys show that fern plastomes keep a conservative gene set but display lineage-specific structural changes—notably large inversions, shifts at IR boundaries, and occasional gene losses. These features help resolve deep nodes and assess rate heterogeneity in fern radiations [5]. Recent syntheses underscore that such structural variation is often clade-specific [6]. Within Polypodiales, the most species-rich fern order, plastid phylogenomics

has clarified relationships in Polypodiaceae and revealed distinct dynamics in epiphytic lineages [7]. Comparative work in Polypodiaceae shows overall collinearity alongside local structural shifts [8].

At finer scales, junction architecture and repeat landscapes carry much of the among-lineage signal. Specifically, border-proximal genes (e.g., *ycf1*, *ndhF*, *rps19/rpl2*) are repeatedly linked to IR expansion or contraction. These structural changes directly alter plastome size and collinearity. Non-coding spacers in the LSC/SSC usually diverge more than coding regions or the IRs, yielding candidate barcodes and phylogenetically informative windows [9,10]. Repeats add another axis of variation: dispersed repeats (forward, reverse, complement, palindromic) and mobile open reading frames (ORFs) can interrupt otherwise conserved synteny, whereas simple sequence repeats (SSRs) provide high-copy, population-level markers [4]. These patterns also appear in recent fern and angiosperm plastome comparisons [6,8].

Recent complete plastomes for *Drynaria acuminata* [11], *Lemmaphyllum intermedium* [12], *Leptochilus/Lepisorus* spp. [13], and *Neocheiropteris* [14] outline two IR size classes and local SSC/IR boundary shifts. Nevertheless, taxonomic coverage remains patchy. Key questions also persist regarding the extent and pace of junction movement, repeat-mediated micro-rearrangements, and the distribution of divergence hotspots. To our knowledge, a complete plastome for *D. trichomanoides* has not yet been reported.

D. trichomanoides is a widespread Asian epiphyte in the *Davalliaceae* complex and is often analyzed alongside Polypodiaceae in plastid phylogenomic matrices. A high-quality plastome for *D. trichomanoides* fills a taxon gap and enables side-by-side comparisons of IR junctions and repeat profiles across epiphytic clades. It also tests whether features reported in *Drynaria*, *Lemmaphyllum*, *Lepisorus/Leptochilus*, and *Neocheiropteris* extend to *Davallia* or reflect lineage-specific histories [7,11–14]. Adding *D. trichomanoides* lets us assign its IR size class and examine SSC/IR boundary behavior within a consistent comparative framework.

In this study, we present the complete plastome of *D. trichomanoides* and compare it with related epiphytic ferns. We report genome size and GC content and assess collinearity. We map the four junctions (JLB, JSB, JSA, JLA) to evaluate IR expansion/contraction and border-gene shifts. We catalog dispersed repeats and simple SSRs. We profile nucleotide diversity (Pi values) to flag hypervariable windows. We infer a plastome phylogeny and compare it with recent Polypodiaceae reconstructions. Placing *D. trichomanoides* in this literature-based framework lets us test whether Polypodiaceae patterns of junction dynamics and repeat signatures extend to *Davalliaceae* s.l., and it highlights candidate loci for barcoding and evolutionary analyses.

2. Materials and Methods

2.1. Sampling and Sequencing

We collected fresh leaves of *D. trichomanoides* near Huaihua, Hunan, China (26°14'54" N, 109°51'40" E). We deposited a voucher (LSU_20250611) in the Herbarium of Lishui University. We extracted total DNA from leaves using the Rapid Plant Genomic DNA Isolation Kit (Sangon Biotech, Shanghai, China). The extract contained both nuclear and plastid DNA; in downstream analyses, we specifically targeted the plastome for assembly. We prepared 150-bp paired-end (PE) libraries and sequenced them on an Illumina HiSeq 2500 (San Diego, CA, USA), yielding 93,481,172 reads (13.93 GB) for subsequent assembly.

2.2. Assembly and Annotation of the Plastome

We generated 13.93 GB of Illumina 150-bp PE reads for plastome assembly. We trimmed adapters and low-quality bases with Trimmomatic v0.39 [15]. We assembled the plastome with GetOrganelle v1.5 [16] and retained the circular contig as the final plastome.

We annotated genes with GeSeq v2.03 [17] and CPGAVAS2 v2.1.0 [18], identifying protein-coding genes (PCGs), rRNAs, tRNAs, IR boundaries, and genome orientation. We then checked and manually refined annotations in CPGView (<http://www.1kmpg.cn/cpgview>; accessed on 28 October 2025) [19], confirming exon–intron structure and cis-/trans-spliced genes. The finalized plastome of *D. trichomanoides* is deposited in GenBank (PX170254).

2.3. Plastome Characteristics

We calculated overall GC content from the finalized plastome in EditSeq v7.1.0 [20]. We generated the circular chloroplast map with OGDRAW v1.3.1 [21] using the curated annotation, showing the LSC, SSC, and IR partitions with standard gene-feature tracks.

2.4. SSR Analyses

We screened interspersed repeats with REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer/>, accessed on 28 October 2025) [22], searching for forward, reverse, complement, and palindromic matches ≥ 30 bp and ≤ 100 bp, and allowing up to three mismatches (Hamming distance ≤ 3). We then identified SSRs using MISA v2.1 [23] with motif thresholds of 10, 5, 4, 3, 3, and 3 for mono-, di-, tri-, tetra-, penta-, and hexanucleotides, respectively. We treated loci with interruptions no longer than 100 bp as compound SSRs.

2.5. Analysis of Codon Preference

We calculated codon-usage statistics and relative synonymous codon usage (RSCU) in CodonW v1.4.4 [24] using all annotated PCGs, the plant plastid genetic code (translation table 11), and excluding stop codons. We then visualized the RSCU profiles for comparative interpretation.

2.6. Analysis of IR Boundary

We used IRscope v3.1 [25] to visualize the four plastome junctions (JLB, JSB, JSA, JLA) in *D. trichomanoides* and four related taxa: *D. acuminata*, *Neocheiropteris ovata*, *L. intermedium*, and *Lepisorus microphyllus*. We chose *D. acuminata*, *N. ovata*, *L. intermedium*, and a representative of *Lepisorus* because these epiphytic Polypodiaceae lineages are the closest and best-sampled plastid comparators relevant to *Davallia*'s placement; complete, publicly available plastomes allow for like-for-like analyses; and together, they encompass the two predominant IR size classes and the characteristic SSC/IR junction shifts documented for allied Polypodiaceae. We loaded curated GenBank annotations into IRscope, generated junction diagrams, and recorded junction coordinates, gene overlaps, and partial-gene incursions across the LSC, SSC, IRa, and IRb regions.

2.7. Analysis of Selection Pressure

We used the *D. trichomanoides* plastome as the reference and four related taxa as queries. We extracted CDS and proteins in PhyloSuite v1.2.1 [26]. We identified putative homologs with BLASTP (BLAST+ v2.10.1) against the *D. trichomanoides* proteome and kept the top hit per gene. We aligned proteins with MAFFT v7.526 [27] and back-translated to codon-aware nucleotide alignments (PAL2NAL-style). Before rate estimation, we removed problematic ORFs (frameshifts, ambiguous bases, and internal stops) and masked low-confidence columns with Gblocks (default). We estimated pairwise nonsynonymous (K_a) and synonymous (K_s) with KaKs_Calculator v3.0 [28] and summarized K_a/K_s (ω) per gene. To reduce artifacts, we reported ω as NA when $K_s = 0$, did not interpret genes with saturated divergence ($K_s > 2$), and did not claim positive selection from $\omega > 1$ without additional site-model support.

2.8. Comparative Analysis of Plastomes

We aligned four related plastomes to the *D. trichomanoides* reference in mVISTA (Shuffle-LAGAN) (<http://genome.lbl.gov/vista/mvista/submit.shtml>, accessed on 28 October 2025) [29]. We then computed Pi values in DnaSP v6 [30] using a 600-bp window and 200-bp step (gaps/missing treated as missing). We flagged windows above the 95th percentile of the Pi value as hypervariable for downstream reporting.

2.9. Phylogenetic Relationship Analysis

We analyzed complete plastomes from 18 fern taxa and set *Lepidomicrosorium hederaceum* as the outgroup. We aligned sequences with MUSCLE v3.8.31 [31]. We inferred a maximum-likelihood tree in RAxML v8.2.12 [32] under GTR+G+I with 1000 rapid bootstrap replicates and a best-scoring maximum-likelihood search. We formatted and annotated the topology and node supports in R using ggtree v3.2.1 [33].

2.10. Comparative Synteny Analysis

We aligned the circularized plastomes with MUMmer v4 [34] after re-orienting each genome at JLB (LSC–IRb). To avoid duplicate hits, we retained only one IR copy per genome. We filtered alignments to identity $\geq 80\%$ and length ≥ 500 bp, and used strand to classify forward vs. reverse matches. We visualized the filtered blocks in Circos v0.69 [35], highlighting the LSC/SSC/IR partitions, and used these plots to assess collinearity and detect short, localized inversions.

3. Results

3.1. Plastome Assembly and Genome Features

Illumina PE sequencing generated 93,481,172 reads (13.93 GB) with high base accuracy (Q20 = 98.84%, Q30 = 95.94%; Table S1). The plastome of *D. trichomanoides* was de novo assembled and circularized with GetOrganelle. Mapping all clean reads back to the assembly showed uniform coverage (mean depth 535.65 $\times$) with 100% (Table S2; Figure S1). The plastome of *D. trichomanoides* is circular and quadripartite (Figure 1), comprising a LSC region of 84,713 bp, an SSC region of 21,436 bp, and two inverted repeats (IRa/IRb) of 24,034 bp each. The overall GC content is 40.82% (IRs 44.38%, LSC 39.86%, SSC 36.62%).

The plastome annotation recovered 115 genes in total, comprising 75 PCGs, 32 tRNA genes, and 8 rRNA genes (Table 1). Sixteen genes were intron-bearing in this genome: *trnG-UCC*, *atpF*, *rpoC1*, *ycf3*, *trnL-UAA*, *trnV-UAC*, *rps12*, *clpP*, *rpl2*, *trnT-UGU*, *trnA-UGC*, *trnI-GAU*, *ndhA*, and the IR-duplicated copies *trnT-UGU-2*, *trnA-UGC-2*, *trnI-GAU-2* (Table 2). Among these, 10 genes contained a single intron (*trnG-UCC*, *atpF*, *rpoC1*, *trnL-UAA*, *trnV-UAC*, *rpl2*, *trnT-UGU*, *trnA-UGC*, *trnI-GAU*, and *ndhA*), whereas *clpP*, *rps12*, and *ycf3* each harbored two introns. Consistent with typical fern plastomes, *rps12* was trans-spliced. The remaining intron-bearing loci represented cis-splicing genes. These cis-splicing genes included *atpF*, *rpoC1*, *clpP*, *ycf3*, *rpl2*, *ndhA*, and six intron-containing tRNA genes (*trnG-UCC*, *trnL-UAA*, *trnV-UAC*, *trnT-UGU*, *trnA-UGC*, and *trnI-GAU*) (Table 2).

Table 1. Gene composition in the plastome of *Davallia trichomanoides*.

Category	Subcategory	Genes (Examples)	No. of Genes
Photosynthesis	Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>	5
	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>psbT</i> , <i>psbZ</i>	15
	NADH dehydrogenase	<i>ndhA</i> *, <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>	10

Table 1. Cont.

Category	Subcategory	Genes (Examples)	No. of Genes
	Cytochrome b6f	<i>petA, petG, petL, petN</i>	4
	ATP synthase	<i>atpA, atpB, atpE, atpF *</i> , <i>atpH, atpI</i>	6
	Rubisco	<i>rbcL</i>	1
	Protochlorophyllide reductase	<i>chlB, chlN</i>	2
Genetic apparatus	Ribosomal proteins (LSU)	<i>rpl14, rpl16, rpl2 *</i> , <i>rpl20, rpl21, rpl22, rpl23, rpl33, rpl36</i>	9
	Ribosomal proteins (SSU)	<i>rps11, rps12 **</i> , <i>rps14, rps18, rps19, rps2, rps3, rps4, rps7</i>	10
	RNA polymerase	<i>rpoA, rpoC1 *</i> , <i>rpoC2</i>	3
	rRNAs	<i>rrn16, rrn23, rrn4.5, rrn5</i>	8
	tRNAs	<i>trnA-UUC *</i> , <i>trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-GCC, trnG-UCC *</i> , <i>trnH-GUG, trnI-CAU, trnI-GAU *</i> , <i>trnL-UAA *</i> , <i>trnL-UAG, trnM-CAU, trnP-GGG, trnP-UGG, trnQ-UUG, trnR-ACG, trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU *</i> , <i>trnV-UAC *</i> , <i>trnW-CCA, trnY-GUA</i>	32
Other/housekeeping	Misc.	<i>accD, cemA, clpP **</i> , <i>ccsA, matK, infA</i>	6
	Conserved hypothetical chloroplast ORF	<i>ycf12, ycf2, ycf3 **</i> , <i>ycf4</i>	4
Total			115

* Gene with one intron; ** gene with two introns.

Table 2. Intron-bearing genes in the plastome of *Davallia trichomanoides*, with exon and intron lengths.

Gene	Location	Exon1	Exon2	Exon3	Intron1	Intron2
<i>trnG-UCC</i>	LSC	31	42		894	
<i>atpF</i>	LSC	144	411		1275	
<i>rpoC1</i>	LSC	177	831		1743	
<i>trnL-UAA</i>	LSC	34	51		580	
<i>trnV-UAC</i>	LSC	25	38		705	
<i>rps12</i>	IRa	232	26	114	844	71,702
<i>clpP</i>	LSC	71	290	242	1064	1084
<i>rpl2</i>	LSC	284	64		1039	
<i>trnI-GAU</i>	IRb	35	37		1080	
<i>trnA-UGC</i>	IRb	37	36		877	
<i>ndhA</i>	SSC	558	558		2044	
<i>trnA-UGC-2</i>	IRa	37	36		806	
<i>trnI-GAU-2</i>	IRa	35	37		1010	
<i>trnT-UGU</i>	IRb	34	38		539	
<i>trnT-UGU-2</i>	IRa	34	38		609	
<i>ycf3</i>	LSC	124	188	162	1095	985

Davallia trichomanoides

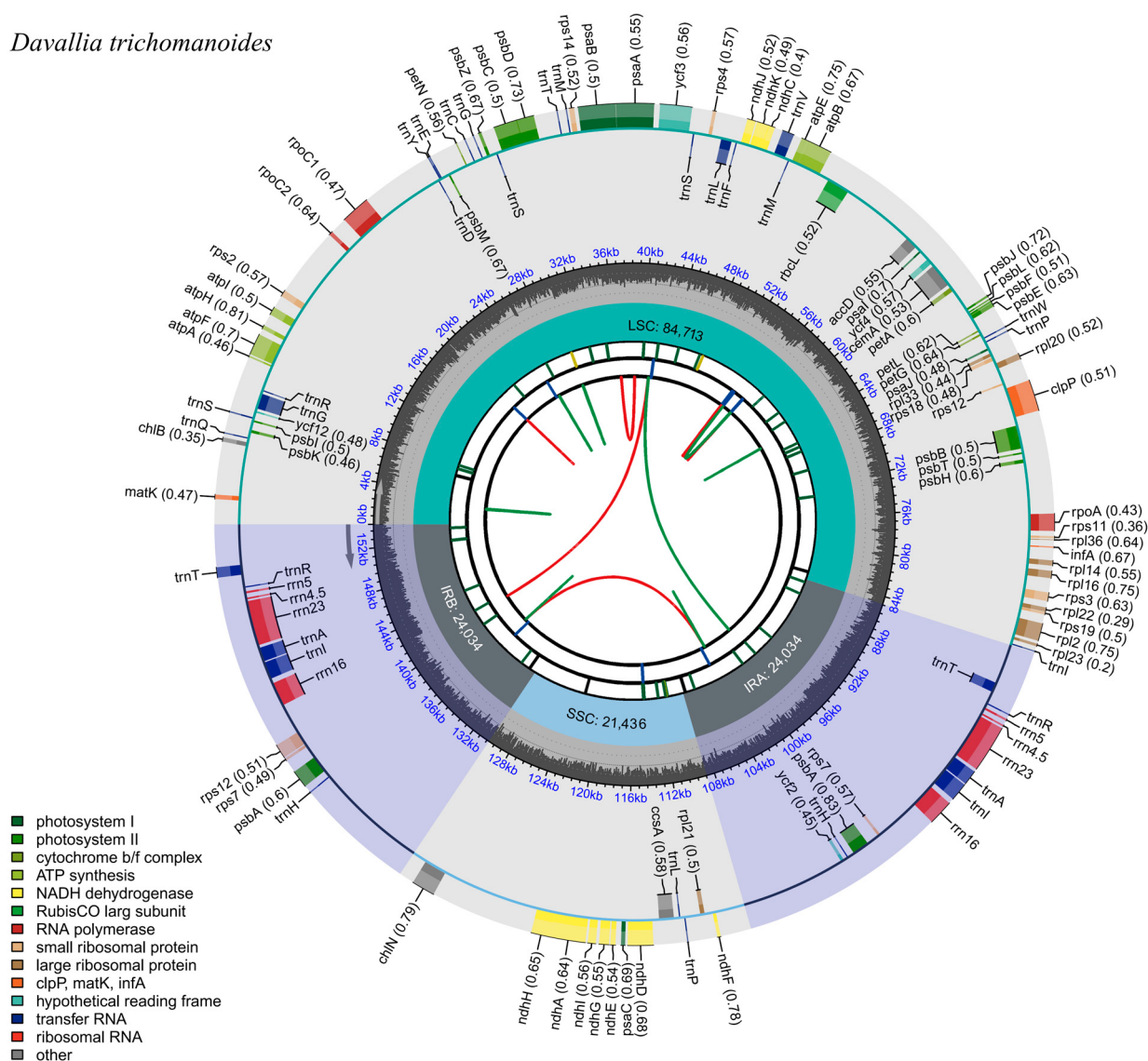


Figure 1. Circular plastome map of *Davallia trichomanoides* (154,217 bp). The plastome is quadripartite, comprising the large single-copy (LSC), small single-copy (SSC), and two inverted repeats (IRa/IRb; lengths indicated). Genes are colored by functional category; the inner ring shows GC content.

3.2. Repeat Sequences and SSR Analysis

Repeats detected with REPuter showed a strongly right-skewed length distribution: most fell in the 20–29 bp bin, counts declined sharply with increasing length, and only sporadic cases were ≥ 50 bp. By type, palindromic and forward repeats dominated, reverse repeats were secondary, and complementary repeats were rare or absent (Figure 2A). The SSR set comprised mono-, di-, tri-, and tetranucleotide motifs; penta- and hexanucleotide SSRs were not detected. Mononucleotide SSRs were most abundant, followed by dinucleotides; trinucleotides and tetranucleotides were less common. Among mononucleotides, A/T motifs predominated; among dinucleotides, AT/TA prevailed; trinucleotides were enriched for AAT/ATT and AAG/CTT types (Figure 2B).

3.3. Codon Usage of Plastome

RSCU profiling of plastid PCGs in *D. trichomanoides* revealed a third-position bias toward A/U. In two- and four-fold degenerate families, A/U-ending codons were generally over-represented ($\text{RSCU} > 1$), whereas most G/C-ending counterparts were under-represented ($\text{RSCU} < 1$). The same pattern held for the six-fold families (Leu, Ser, Arg), in

which A/U-ending codons showed higher RSCU than G/C-ending codons. Single-codon families—Met (AUG) and Trp (UGG)—had RSCU = 1.0 (Figure 3).

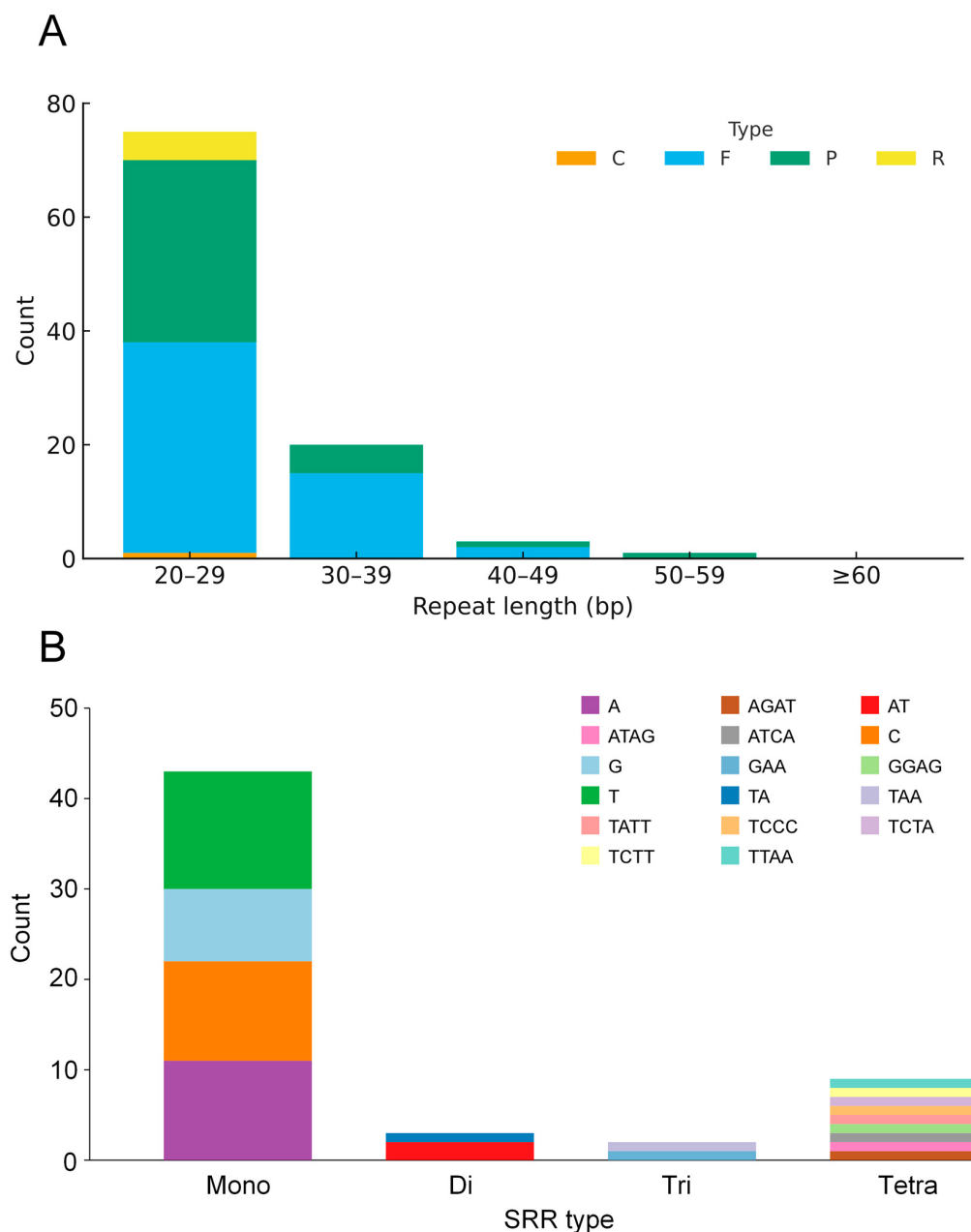


Figure 2. Analysis of repetitive elements in the *Davallia trichomanoides* plastome. **(A)** Counts of four types of interspersed repeats (forward, reverse, complement, and palindromic), identified with a size threshold of ≥ 30 bp and allowing ≤ 3 mismatches. **(B)** Counts of simple sequence repeats (SSRs) categorized by the length of their repeating motif (mono- to hexanucleotide).

3.4. IR Boundary Comparison

Comparative mapping of the four junctions (JLB, JSB, JSA, and JLA) across five Polypodiaceae plastomes shows a generally conserved quadripartite organization with modest size variation (Figure 4). Total plastome sizes span 151,591–158,029 bp. The LSC ranges from 80,661–84,713 bp, the SSC from 21,436–21,797 bp, and each IR from 24,034–27,494 bp. Two IR size classes are apparent: ~ 24.0 – 24.6 kb in *D. trichomanoides* (24,034 bp), *D. acuminata* (24,621 bp) and *N. ovata* (24,609 bp), versus ~ 27.4 – 27.5 kb in *L. intermedium* (27,421 bp) and *L. microphyllus* (27,494 bp), indicating lineage-specific IR expansion in the latter pair.

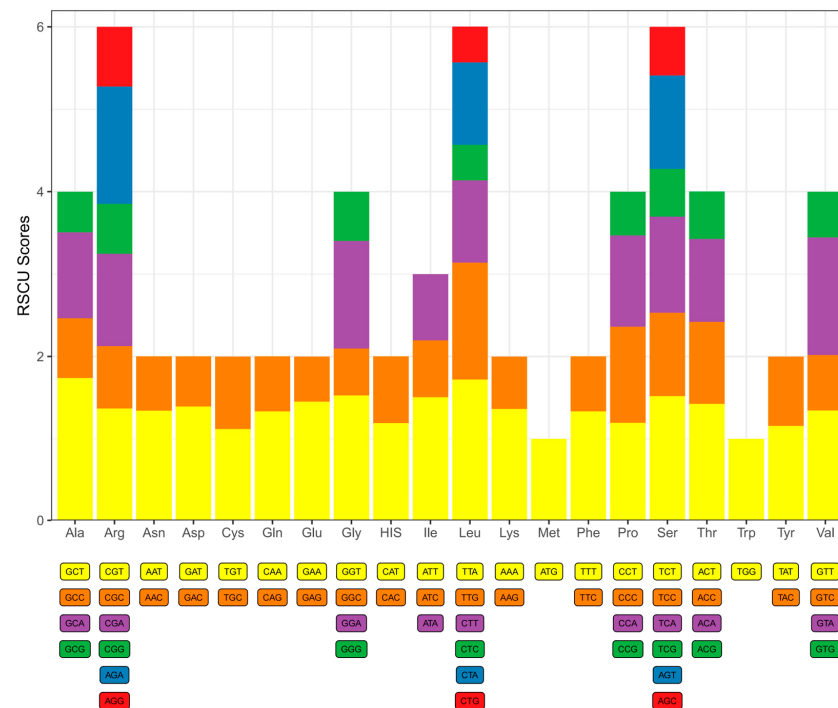


Figure 3. Relative synonymous codon usage (RSCU) in the plastome protein-coding genes of *Davallia trichomanoides*. Bars are grouped by amino acid. Within each group, stacked segments represent the RSCU value of each synonymous codon (labels shown below). Segment heights > 1.0 indicate over-represented codons and < 1.0 indicate under-represented codons. Single-codon families (Met/AUG and Trp/UGG) show RSCU = 1.0.

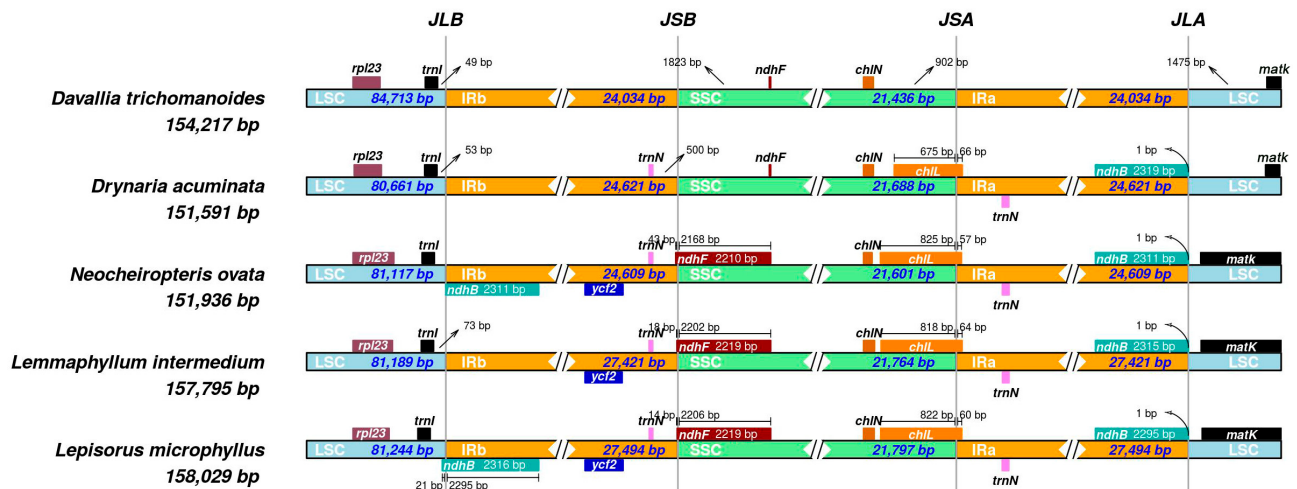


Figure 4. Comparison of the four junction sites (JLB, JSB, JSA, and JLA) between the inverted repeats (IRa, IRb) and single-copy regions (LSC and SSC) in five Polypodiaceae plastomes. Genes near the junctions are shown. Numbers indicate the distance in base pairs from the junction to the start or end of the adjacent gene.

At JLB (LSC–IRb), *trnI* lies consistently close to the border (~49–73 bp across taxa), indicating minimal LSC→IRb drift. At JSB (IRb–SSC), *ndhF* shows lineage-specific shifts: in *D. trichomanoides* and *D. acuminata* it remains in the SSC at similar distances (~1.8 kb) from JSB, whereas in *N. ovata*, *L. intermedium*, and *L. microphyllus*, the 3' end of *ndhF* crosses the junction into IRb by ~43 bp, 18 bp, and 14 bp, respectively. In contrast, JSA (SSC–IRa) is highly conserved, with *chlN* situated ~900 bp from the junction in all species. At JLA (IRa–LSC), *matK* in *D. trichomanoides* is 1475 bp from the border; *D. acuminata* appears

similar, whereas in *N. ovata*, *L. intermedium*, and *L. microphyllus*, *matK* lies closer to JLA. The duplicated *ndhB* copies about JLA (~1 bp) in the compared taxa.

3.5. Comparative Analysis of Plastomes

The mVISTA program (reference *D. trichomanoides*) shows high overall similarity among the five plastomes (Figure 5). Consistent with plastome organization, regions corresponding to the IRs appear more conserved than the LSC/SSC, and non-coding spacers vary more than coding sequences. Conserved peaks cluster around photosynthesis- and ribosome-related gene blocks (e.g., *psb/psa/atp/rpl/rps*), whereas most divergence is localized to intergenic spacers across the LSC and SSC.

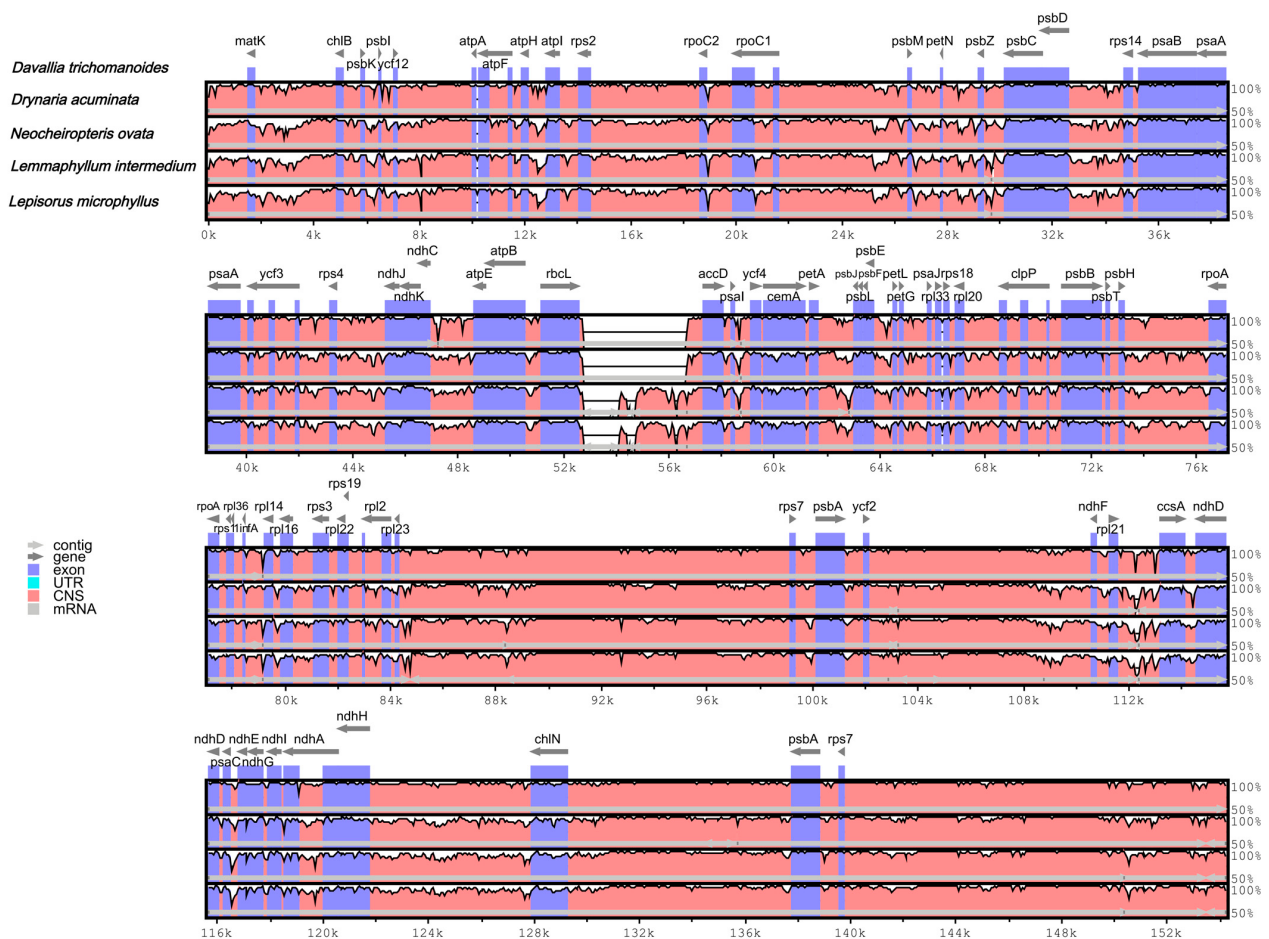


Figure 5. Whole-plastome alignment (mVISTA; reference *D. trichomanoides*). Gray arrows denote gene orientation, and thick black bars indicate the positions of the inverted repeats (IRs). The y-axis represents percent identity ranging from 50% to 100%. The color coding of the genomic region is protein coding (exon, purple), ribosomal RNA (rRNA, cyan), and conserved non-coding sequences (CNS, pink).

Using DnaSP sliding windows, nucleotide diversity across five plastomes shows low overall variability in the IRs and elevated peaks in the LSC/SSC (Figure 6). Pi values span ~0–0.059, with the highest signal immediately upstream of *psaI-ycf4* (58382–59402 bp; Pi = 0.059). Recurrent hypervariable segments (Pi ≥ 0.046) occur around *trnS-GCU*, *trnC-GCA*, and *trnG-GCC* block, the *psaI-ycf4* region, and within the SSC near the IRb boundary (e.g., 104.6–106.2 kb, 108.7–110.0 kb, the latter at *trnP-GGG*). In contrast, IR-embedded loci such as *rrn23/rrn16* and *psbA* remain highly conserved (Pi ≈ 0–0.01).

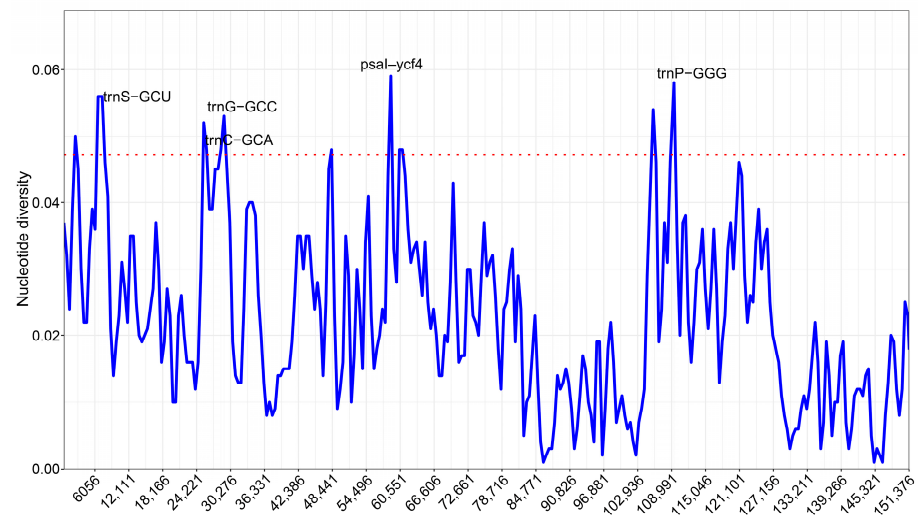


Figure 6. Sliding-window nucleotide diversity (Pi values) across five plastomes (window 600 bp; step 200 bp). The dashed line marks the 95th percentile. Labeled peaks denote hypervariable windows.

3.6. Selection Pressure Analysis

Ks and Ka substitution rates were estimated for plastome PCGs. Gene-wise Ka/Ks ratios ranged from 0.00 to 0.88 (mean 0.15). No gene exhibited Ka/Ks ≥ 1.0 , indicating uniformly low ratios across loci (Figure 7).

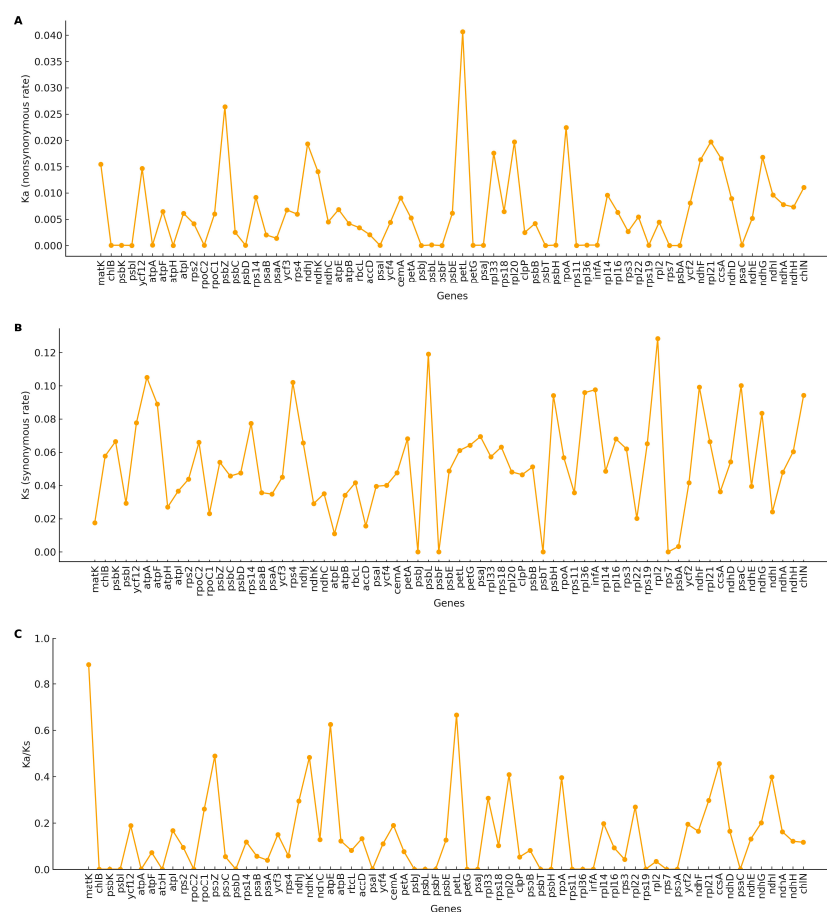


Figure 7. Selective pressure on protein-coding genes across five plastomes. (A) Nonsynonymous substitution rate (Ka) for each gene. (B) Synonymous substitution rate (Ks) for each gene. (C) Ratio of nonsynonymous to synonymous substitutions (Ka/Ks or ω). Each point represents a single gene. Genes with Ks = 0 were treated as NA; genes with saturated Ks (e.g., >2) were not interpreted.

3.7. Comparative Synteny Analysis

The circos synteny plot indicates extensive forward collinearity between the focal plastome and closely related genomes, with paired signals across the IRs and only short, localized reverse matches in the single-copy regions. These patterns support a largely conserved plastome architecture with minor micro-rearrangements (Figure 8).

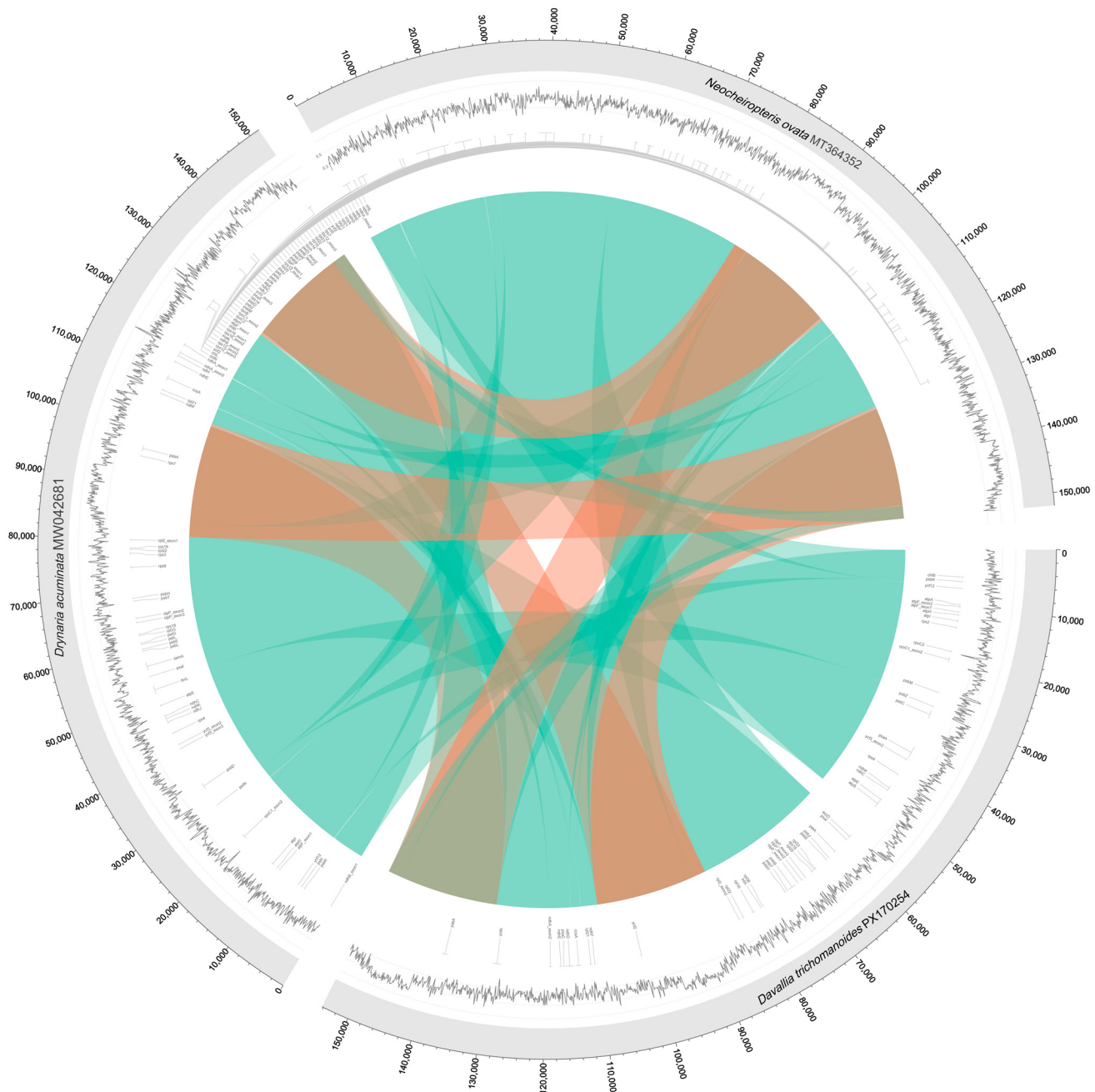


Figure 8. Circos plot visualizing synteny and alignment between three plastomes. The rings from outer to inner show: gene positions (outer ring), GC content (inner ring, black peaks), and aligned sequences between genomes (central ribbons). Ribbons are colored cyan for forward alignments and red for reverse alignments. Alignments were filtered for $\geq 80\%$ identity and ≥ 500 bp length.

3.8. Phylogenetic Analysis

A maximum-likelihood tree of 18 fern plastomes (*L. hederaceum* as outgroup) recovered a backbone with uniformly high support (most nodes bootstrap ≥ 99 ; many at 100). *D. trichomanoides* formed a strongly supported lineage sister to *D. acuminata*

(bootstrap = 100), clarifying its position relative to allied polypodiaceous genera. Monophyly was recovered for the sampled clades of *Pyrrosia* (multiple accessions clustering with bootstrap = 100) and *Lemmaphyllum* (*L. intermedium* + *L. carnosum*, bootstrap = 100), with *Neocheiropteris* accessions grouping together and distinct from *Microsorium*/*Lepisorus* lineages, all with maximal or near-maximal support. Overall, the topology is congruent with current taxonomy and indicates limited conflict among plastome histories, while highlighting the close affinity of *Davallia* and *Drynaria* within this sampling (Figure 9).

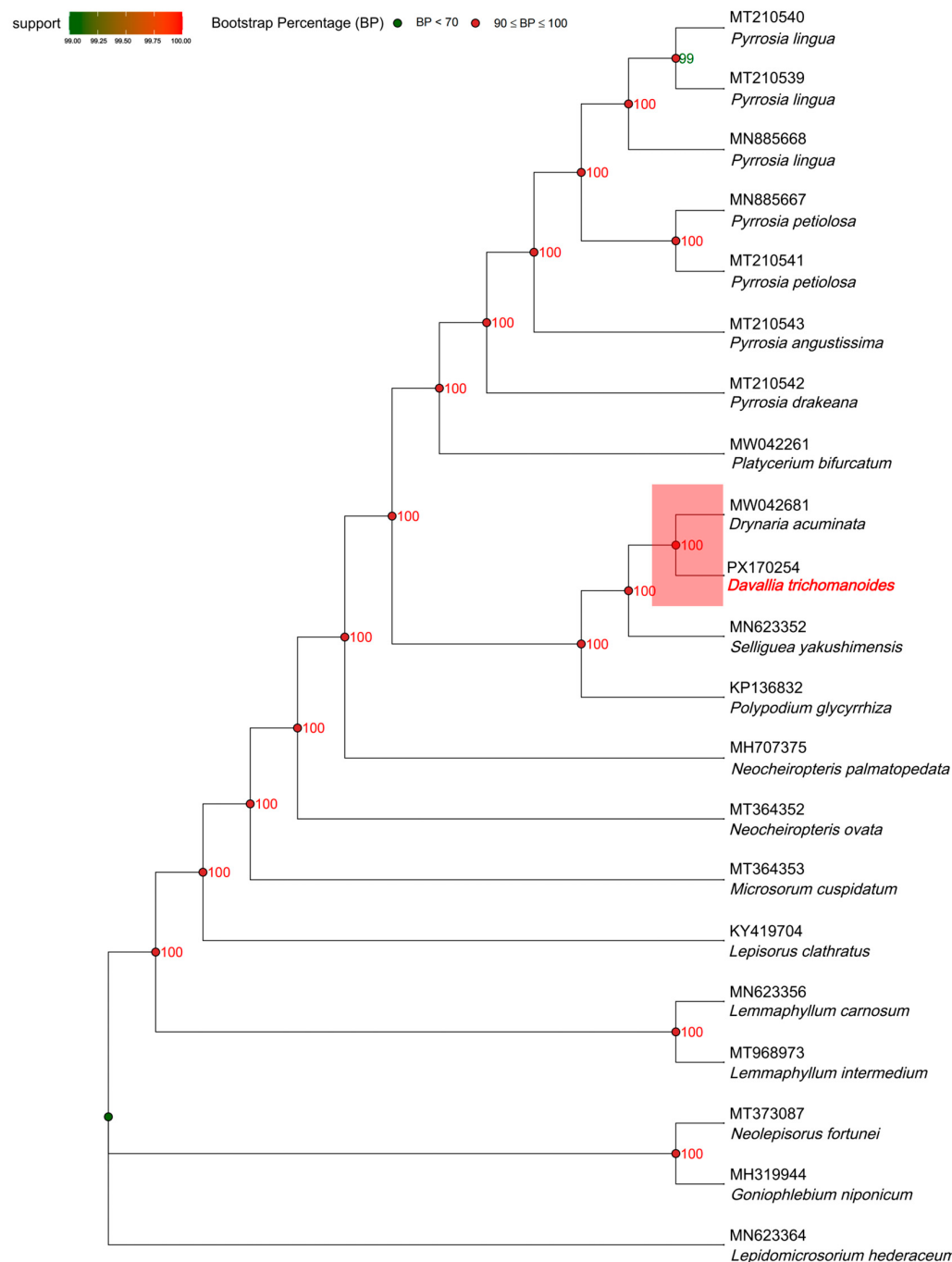


Figure 9. Maximum-likelihood plastome phylogeny of *Davallia trichomanoides* and related Polypodiaceae taxa. Node labels indicate bootstrap supports (%) (1000 replicates) under the best-fit partitioned model selected by ModelFinder. The tree is rooted with *Lepidomicrosorium hederaceum*; branch lengths are scaled to substitutions per site. For clarity, only nodes with support ≥ 70 are labeled; lower-support nodes are omitted.

4. Discussion

4.1. Plastome Organization in *D. trichomanoides*

The *D. trichomanoides* plastome matches the canonical quadripartite layout and size typical of leptosporangiate ferns, with GC content and gene repertoire within the range for Polypodiales [36–38]. The IRs are far less variable than the LSC/SSC regions, and gene order is largely conserved, underscoring the stabilizing role of the duplicated IRs [36,37]. Overall, these features agree with broad syntheses showing that most land-plant plastomes—including ferns—are structurally conservative, with lineage-specific departures concentrated at single-copy/IR junctions and in non-coding spacers [36–38].

4.2. IR Boundary Dynamics

Across the five plastomes, we recovered two IR size classes and small, taxon-specific junction shifts. At JLB, *trnI* sits at the border in all taxa, indicating little LSC→IRb drift. JSB is heterogeneous: in *D. trichomanoides* and *D. acuminata*, *ndhF* remains in the SSC at a similar offset, whereas in *N. ovata*, *L. intermedium*, and *L. microphyllus*, the 3' end extends a few tens of bp into IRb. Such minor, repeated incursions of *ndhF* or *ycf1* across JSB/JSA are a common sign of IR expansion/contraction [7,36]. These small junction shifts explain much of the length variation observed among otherwise collinear plastomes. At JLA, *matK* lies ~1.5 kb inside the LSC in *D. trichomanoides*, marking a species-specific terminal shift without broader disruption—again consistent with incremental boundary movement rather than large rearrangements [7,36].

4.3. Patterns of Sequence Divergence and Marker Potential

Sliding-window nucleotide diversity (Pi values) shows that divergence concentrates in LSC/SSC intergenic spacers and a few coding genes, whereas the IRs are conserved [36–38]. The top windows (e.g., *psaI-ycf4*, *trnS-trnG*, *trnC-trnG*, and SSC-proximal segments) are practical targets for marker development in *Davallia* and its relatives. Peaks near *ycf1* and *ndhF* match broader evidence that these loci are among the most variable plastid genes at low taxonomic levels and can outperform classical barcodes [39,40]. Including these hotspots in multi-locus panels should improve resolution for populations and species complexes in epiphytic ferns [39,40].

4.4. Repeats and SSRs

The repeat landscape is dominated by forward and palindromic repeats, as in many fern and seed-plant plastomes. The SSR set is A/T-rich, with mononucleotide runs most abundant—patterns widely reported for plastomes [36,41]. SSRs, especially A/T homopolymers in the LSC/SSC regions, are useful for population studies and for building genotyping panels, provided homopolymer length variation is validated across platforms [41]. This A/T bias in plastid SSRs recurs across diverse plant groups, underscoring its generality and value for intraspecific work [41].

4.5. Selection on PCGs

Ka/Ks ratios are mostly <1, indicating pervasive purifying selection on plastid PCGs, a pattern widely reported for land plants and ferns [38,41]. Occasional locus-specific elevations are more plausibly due to relaxed constraint, lineage-specific rate variation, or local alignment uncertainty than to broad positive selection; substantiating adaptive signals would require targeted site-model tests and denser taxon sampling [38,41].

4.6. Collinearity and Structural Stability

Whole-plastome collinearity is high, with only short, localized reverse matches in the single-copy regions and no large inversions like those reported in deeper fern lin-

eages [37,42]. This supports a “conservative-with-micro-edits” model for *D. trichomanoides*: genome-wide synteny persists, while small changes arise via repeat-mediated events and incremental junction shifts [37,42].

4.7. Phylogenetic Placement and Taxonomic Congruence

The plastome maximum-likelihood tree places *D. trichomanoides* as sister to *D. acuminata* with strong support, matching recent plastid phylogenomic studies in Polypodiaceae that stabilize epiphytic lineages [7]. The small boundary differences noted above do not disrupt the plastome-scale signal at this depth. The plastome is a single, largely non-recombining and typically uniparentally inherited locus. Consequently, it records the maternal history and can be affected by processes like introgression or incomplete lineage sorting. We therefore interpret this topology as representing the plastid history of the group. Future work will add nuclear target-capture data to test plastid–nuclear concordance and refine relationships.

4.8. Ecological Context of Epiphytic Ferns

Epiphytic canopies impose rapid light swings and intermittent water supply, and many epiphytic ferns are poikilohydric with quick post-drying recovery [43]. In this context, several genomic features of *D. trichomanoides* are consistent with stabilizing selection on core photosynthetic machinery [8,44]. These features include its high plastome collinearity, strong IR conservation (evidenced by relatively low Pi values compared to LSC/SSC), and gene-wise Ka/Ks ratios consistently below one. The IR likely enhances plastome stability and increases rRNA operon dosage, supporting translation under stress [45]. Retention of chloroplast *ndh* genes is also ecologically plausible, as the NDH complex protects PSI and sustains cyclic electron flow under fluctuating light and water deficit [46].

5. Conclusions

This study characterizes the plastome of *D. trichomanoides* and, in comparison with related epiphytic ferns, shows lineage-specific differences at inverted-repeat/single-copy junctions that explain modest size variation while preserving overall collinearity. Small IR-border shifts—often involving *ycf1* and *ndhF*—fit the common expansion–contraction model that drives plastome length change. The plastome tree places *D. trichomanoides* with *D. acuminata* with strong support, consistent with recent Polypodiaceae plastid phylogenies and clarifying relationships among epiphytic lineages. Dispersed repeats and A/T-rich cpSSRs, together with Pi values peaks concentrated in LSC/SSC, provide practical loci for barcoding and population studies; *ycf1* and adjacent spacers remain high-value targets. Going forward, we will use these hotspots (e.g., *ycf1*-, *ndhF*-, and *psaI*–*ycf4* regions) to build multi-locus barcodes and test whether SSR variation tracks geography, informing germplasm characterization and conservation in *Davallia*. These steps extend the present plastome framework from systematics to mechanism-focused and population-scale questions in epiphytic Polypodiaceae.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes16111310/s1>, Figure S1: The distribution of chloroplast genome sequencing depth for *Davallia trichomanoides* is depicted in the graph, with the horizontal axis representing genomic position and the vertical axis indicating sequencing depth; Table S1: All clean data information; Table S2: Coverage summary.

Author Contributions: Conceptualization, Y.L. and S.D.; data curation, Y.W. and Z.X.; formal analysis, Y.W., Z.X. and K.L.; methodology, Y.L.; project administration, S.D.; resources, Z.X., K.L. and Y.L.; software, Y.L. and S.D.; supervision, Y.L.; writing—original draft, Y.W.; writing—review and editing, Y.L. and S.D. All authors have read and agreed to the published version of the manuscript.

Funding: The funding for this study was provided by the National Natural Science Foundation of China Cultivation Project (Grant No. QDZK112022022).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in GenBank of NCBI at <https://www.ncbi.nlm.nih.gov> (accessed on 28 October 2025) under the accession number PX170254. The associated BioProject, SRA, and Biosample numbers are PRJNA1303833, SRR34948204, and SAMN50537047, respectively.

Conflicts of Interest: Author Siyuan Dong was employed by the company China National Research Institute of Food and Fermentation Industries Co., Ltd. The remaining 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.

References

1. Dong, Y.; Cao, Q.; Yu, K.; Wang, Z.; Chen, S.; Chen, F.; Song, A. Chloroplast phylogenomics reveals the maternal ancestry of cultivated chrysanthemums. *Genom. Commun.* **2025**, *2*, e019. [\[CrossRef\]](#)
2. Kan, J.; Nie, L.; Wang, M.; Tiwari, R.; Tembrock, L.R.; Wang, J. The Mendelian pea pan-plastome: Insights into genomic structure, evolutionary history, and genetic diversity of an essential food crop. *Genom. Commun.* **2024**, *1*, e004. [\[CrossRef\]](#)
3. Hu, S.; Sablok, G.; Wang, B.; Qu, D.; Barbaro, E.; Viola, R.; Li, M.; Varotto, C. Plastome organization and evolution of chloroplast genes in *Cardamine* species adapted to contrasting habitats. *BMC Genom.* **2015**, *16*, 306. [\[CrossRef\]](#)
4. Robison, T.A.; Grusz, A.L.; Wolf, P.G.; Mower, J.P.; Fauskee, B.D.; Sosa, K.; Schuettpelz, E. Mobile elements shape plastome evolution in ferns. *Genome Biol. Evol.* **2018**, *10*, 2558–2571. [\[CrossRef\]](#)
5. Kuo, L.Y.; Qi, X.; Ma, H.; Li, F.W. Order-level fern plastome phylogenomics: New insights from Hymenophyllales. *Am. J. Bot.* **2018**, *105*, 1545–1555. [\[CrossRef\]](#)
6. Du, X.-Y.; Kuo, L.-Y.; Zuo, Z.-Y.; Li, D.-Z.; Lu, J.-M. Structural variation of plastomes provides key insight into the deep phylogeny of ferns. *Front. Plant Sci.* **2022**, *13*, 862772. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Wei, R.; Yang, J.; He, L.J.; Liu, H.M.; Hu, J.Y.; Liang, S.Q.; Wei, X.P.; Zhao, C.F.; Zhang, X.C. Plastid phylogenomics provides novel insights into the infrafamilial relationship of *Polypodiaceae*. *Cladistics* **2021**, *37*, 717–727. [\[CrossRef\]](#)
8. Liu, S.; Wang, Z.; Su, Y.; Wang, T. Comparative genomic analysis of *Polypodiaceae* chloroplasts reveals fine structural features and dynamic insertion sequences. *BMC Plant Biol.* **2021**, *21*, 31. [\[CrossRef\]](#)
9. Ping, J.; Hao, J.; Wang, T.; Su, Y. Comparative analysis of plastid genomes reveals rearrangements, repetitive sequence features, and phylogeny in the *Annonaceae*. *Front. Plant Sci.* **2024**, *15*, 1351388. [\[CrossRef\]](#)
10. Niu, Z.; Pan, J.; Zhu, S.; Li, L.; Xue, Q.; Liu, W.; Ding, X. Comparative analysis of the complete plastomes of *Apostasia wallichii* and *Neuwiedia singapureana* (*Apostasioideae*) reveals different evolutionary dynamics of IR/SSC boundary among photosynthetic orchids. *Front. Plant Sci.* **2017**, *8*, 1713. [\[CrossRef\]](#)
11. Wang, K.; Liu, H. The first complete chloroplast genome of *Drynaria acuminata* (*Polypodiaceae*), a local rare fern species. *Mitochondrial DNA Part B Resour.* **2021**, *6*, 250–251. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Wang, Z.; Li, N.; Li, Z.; Liu, Y.; Yang, P.; Hong, Y.; He, Z.; Su, Y.; Wang, T. The complete chloroplast genome of *Lemmaphyllum intermedium*, a valuable medicinal fern. *Mitochondrial DNA Part B Resour.* **2021**, *6*, 423–424. [\[CrossRef\]](#)
13. Min, Y.; Guan, J.; Li, S.; Liu, S.; Hong, Y.; Wang, Z.; Wang, T.; Su, Y. The complete chloroplast genome of *Leptochilus hemionitideus*, a traditional Chinese medical fern. *Mitochondrial DNA Part B Resour.* **2018**, *3*, 784–785. [\[CrossRef\]](#)
14. Du, X.-Y.; Lu, J.-M.; Lu, S.-G.; Li, D.-Z. Complete plastome of an endemic fern species from China: *Neocheiropteris palmatopedata* (*Polypodiaceae*). *Mitochondrial DNA Part B Resour.* **2019**, *4*, 2394–2395. [\[CrossRef\]](#)
15. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [\[CrossRef\]](#)
16. Jin, J.-J.; Yu, W.-B.; Yang, J.-B.; Song, Y.; dePamphilis, C.W.; Yi, T.-S.; Li, D.-Z. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **2020**, *21*, 241. [\[CrossRef\]](#)
17. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq—Versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Shi, L.; Chen, H.; Jiang, M.; Wang, L.; Wu, X.; Huang, L.; Liu, C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* **2019**, *47*, W65–W73. [\[CrossRef\]](#)

19. Liu, S.; Ni, Y.; Li, J.; Zhang, X.; Yang, H.; Chen, H.; Liu, C. CPGView: A package for visualizing detailed chloroplast genome structures. *Mol. Ecol. Resour.* **2023**, *23*, 694–704. [[CrossRef](#)] [[PubMed](#)]
20. Burland, T.G. DNASTAR's Lasergene sequence analysis software. In *Bioinformatics Methods and Protocols*; Springer: New York, NY, USA, 2000; pp. 71–91.
21. Lohse, M.; Drechsel, O.; Bock, R. OrganellarGenomeDRAW (OGDRAW): A tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* **2007**, *52*, 267–274. [[CrossRef](#)]
22. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [[CrossRef](#)]
23. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)] [[PubMed](#)]
24. Sharp, P.M.; Li, W.-H. The codon adaptation index—A measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **1987**, *15*, 1281–1295. [[CrossRef](#)]
25. Amiryousefi, A.; Hyvönen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [[CrossRef](#)] [[PubMed](#)]
26. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [[CrossRef](#)] [[PubMed](#)]
27. Katoh, K.; Toh, H. Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics* **2010**, *26*, 1899–1900. [[CrossRef](#)]
28. Zhang, Z. KaKs_Calculator 3.0: Calculating selective pressure on coding and non-coding sequences. *Genom. Proteom. Bioinform.* **2022**, *20*, 536–540. [[CrossRef](#)]
29. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32*, W273–W279. [[CrossRef](#)]
30. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)]
31. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
32. Stamatakis, A. RAxML-VI-HP: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **2006**, *22*, 2688–2690. [[CrossRef](#)]
33. Xu, S.; Li, L.; Luo, X.; Chen, M.; Tang, W.; Zhan, L.; Dai, Z.; Lam, T.T.; Guan, Y.; Yu, G. Ggtree: A serialized data object for visualization of a phylogenetic tree and annotation data. *iMeta* **2022**, *1*, e56. [[CrossRef](#)] [[PubMed](#)]
34. Marçais, G.; Delcher, A.L.; Phillippy, A.M.; Coston, R.; Salzberg, S.L.; Zimin, A. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput. Biol.* **2018**, *14*, e1005944. [[CrossRef](#)] [[PubMed](#)]
35. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* **2009**, *19*, 1639–1645. [[CrossRef](#)] [[PubMed](#)]
36. Daniell, H.; Lin, C.-S.; Yu, M.; Chang, W.-J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [[CrossRef](#)]
37. Wicke, S.; Schneeweiss, G.M.; dePamphilis, C.W.; Müller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [[CrossRef](#)]
38. Wolf, P.G.; Der, J.P.; Duffy, A.M.; Davidson, J.B.; Grusz, A.L.; Pryer, K.M. The evolution of chloroplast genes and genomes in ferns. *Plant Mol. Biol.* **2011**, *76*, 251–261. [[CrossRef](#)]
39. Dong, W.; Xu, C.; Li, C.; Sun, J.; Zuo, Y.; Shi, S.; Cheng, T.; Guo, J.; Zhou, S. Ycf1, the most promising plastid DNA barcode of land plants. *Sci. Rep.* **2015**, *5*, 8348. [[CrossRef](#)]
40. Amar, H.M. ycf1-ndhF genes, the most promising plastid genomic barcode, sheds light on phylogeny at low taxonomic levels in *Prunus persica*. *J. Genet. Eng. Biotechnol.* **2020**, *18*, 42. [[CrossRef](#)]
41. Song, W.; Ji, C.; Chen, Z.; Cai, H.; Wu, X.; Shi, C.; Wang, S. Comparative analysis of the complete chloroplast genomes of nine *Musa* species: Genomic features, comparative analysis, and phylogenetic implications. *Front. Plant Sci.* **2022**, *13*, 832884. [[CrossRef](#)]
42. Wolf, P.G.; Roper, J.M.; Duffy, A.M. The evolution of chloroplast genome structure in ferns. *Genome* **2010**, *53*, 731–738. [[CrossRef](#)]
43. Zotz, G.; Hietz, P. The physiological ecology of vascular epiphytes: Current knowledge, open questions. *J. Exp. Bot.* **2001**, *52*, 2067–2078. [[CrossRef](#)]
44. Fan, R.; Ma, W.; Liu, S.; Huang, Q. Integrated analysis of three newly sequenced fern chloroplast genomes: Genome structure and comparative analysis. *Ecol. Evol.* **2021**, *11*, 4550–4563. [[CrossRef](#)] [[PubMed](#)]

45. Krämer, C.; Boehm, C.R.; Liu, J.; Ting, M.K.Y.; Hertle, A.P.; Forner, J.; Ruf, S.; Schöttler, M.A.; Zoschke, R.; Bock, R. Removal of the large inverted repeat from the plastid genome reveals gene dosage effects and leads to increased genome copy number. *Nat. Plants* **2024**, *10*, 923–935. [[CrossRef](#)] [[PubMed](#)]
46. Shikanai, T.; Yamamoto, H. Contribution of cyclic and pseudo-cyclic electron transport to the formation of proton motive force in chloroplasts. *Mol. Plant* **2017**, *10*, 20–29. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.