Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Indian Trithuria (Hydatellaceae): complete plastid genome of Trithuria konkanensis, comparative plastome evolution and phylogenomics of the ANA grade angiosperms

Manoj M. Lekhak

mml_botany@unishivaji.ac.in

Shivaji University

Rupali N. Chougule

Shivaji University

Asif S. Tamboli

Independent Researcher

Siddharthan Surveswaran

University of Hyderabad

Research Article

Keywords: Early-diverging angiosperms. Nymphaeales. Plastid genome. Phylogenetic relationships

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11418920/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations:
No competing interests reported.

Tables 1 and 2 are available in the supplementary files section

The Indian *Trithuria* (Hydatellaceae): complete plastid genome of *Trithuria konkanensis*, comparative plastome evolution and phylogenomics of the ANA grade angiosperms

Manoj M. Lekhak[1*], Rupali N. Chougule[1], Asif S. Tamboli[2], Siddharthan Surveswaran[3]

[1]Angiosperm Taxonomy Laboratory, Department of Botany, Shivaji University, Kolhapur, 416004, Maharashtra, India

[2]Independent Researcher, Dighanchi, Sangli, Maharashtra, India

[3]Systematics, Ecology and Evolution Laboratory, Department of Plant Sciences, School of Life Sciences, University of Hyderabad, Hyderabad, 500046, Telangana, India

E-mail address: mml_botany@unishivaji.ac.in

Abstract

Background Hydatellaceae (Nymphaeales), represented by the single genus *Trithuria*, is an early-diverging angiosperm lineage within the Amborellales–Nymphaeales–Austrobaileyales (ANA) grade. *Trithuria konkanensis* is the only species of the genus reported from India; however, its plastid genome remains unexplored. Given its strategic phylogenetic position and importance for elucidating the early evolution and diversification of flowering plants, characterization of the plastid genome of *T. konkanensis* is essential. Such genomic information will provide valuable insights into plastid genome variation within Hydatellaceae and across ANA grade angiosperms, while also contributing to a better understanding of their phylogenetic relationships.

Results The plastid genome of *Trithuria konkanensis* exhibited a typical quadripartite structure comprising 114 unique genes, including 80 protein-coding genes, 30 tRNA genes, and four rRNA genes. The *clpP1* gene lacked introns. The plastid genome was 179,088 bp in length, with large single-copy (LSC) and small single-copy (SSC) regions of 68,412 and 6,052 bp, respectively, and an inverted repeat (IR) region of 52,312 bp. A total of 33 protein-coding genes were duplicated within the IR region. Gene-wise nucleotide diversity (π = 0.00000–0.08753) revealed substantial variation across the plastome. The genes *clpP1* (0.08753), *ndhK* (0.07136), *rpl32* (0.05338), *ccsA* (0.05278), *ndhF* (0.04962), *accD* (0.04444), and *ycf1* (0.03850) exhibited relatively high π values, identifying them as potential molecular markers for phylogenetic and comparative analyses. Among the intergenic spacers, *rps12*/*ycf1* showed the highest sequence variability, followed by *ndhC*/*atpE*, *ycf4*/*cemA*, *psbE*/*petL*, *ndhF*/*rpl32*, and *rpl32*/*ccsA*. Phylogenetic analyses based on protein-coding genes, the complete plastome, and nrITS sequences yielded congruent topologies, with *T. konkanensis* recovered as sister to *T. lanterna*, while *T. filmentosa* was recovered as sister to *T. inconspicua*.

Conclusions Among the ANA grade angiosperms, Hydatellaceae possess the largest plastomes (~180 kb). This study expands the plastome resources available for Hydatellaceae, providing valuable data for a better understanding of plastome evolution and the evolutionary history of the family. Broader taxon sampling will be essential to determine the extent of plastome diversity and structural variation within Hydatellaceae.

Keywords Early-diverging angiosperms. Nymphaeales. Plastid genome. Phylogenetic relationships.

Introduction

Chloroplasts are endosymbiotic organelles and have their own genetic material [1]. The development of next-generation sequencing technologies has revolutionized the study of chloroplast genetics and genomics. Chloroplast genomics improve our understanding of plant evolution and also has translational applications, such as conferring protection against biotic or abiotic stress and the development of vaccines and biopharmaceuticals in edible crop plants [2]. Most studied, angiosperm chloroplast genomes (often referred as plastomes) are circular, with a quadripartite structure comprising two inverted repeats (IRs) separating a large single-copy from a small single-copy (SSC) region, usually spanning 120–160 kb and encoding 110–130 unique genes [3]. The plastome size ranges from 11 kb (*Pilostyles aethiopica* Welw., Apodanthaceae) [4] to 242 kb (*Pelargonium transvaalense* R.Knuth., Geraniaceae) [5].

The early-diverging angiosperms, Amborellales, Nymphaeales, and Austrobaileyales constitute the ANA grade [6]. Amborellales includes a single family, Amborellaceae whereas Nymphaeales and Austrobaileyales consist of three families each. Nymphaeales occupies a central position in understanding the evolutionary history of angiosperms [7]. Testing hypotheses regarding gene order and inverted repeat size in the plastid genome of the ancestor of angiosperms requires the investigation of a broader taxonomic sampling within the order [7]. Nymphaeales is currently recognized as comprising three families: Cabombaceae, Hydatellaceae and Nymphaeaceae [6].

Hydatellaceae is a monogeneric family with *Trithuria* Hook.f., its only genus [8, 9], traditionally placed in Centrolepidaceae of the monocot order Poales. Hamann [10] segregated it as a distinct family based on embryology, seed and vegetative anatomy, floral and pollen morphology. Saarela et al. [11] based on a multigenic molecular phylogenetic analysis assigned

Hydatellaceae to the order Nymphaeales. The genus comprises 12 species, of which eight are endemic to mainland Australia, one each to Tasmania (*T. filamentosa* Rodway), New Zealand (*T. inconspicua* Cheeseman) and India (*T. konkanensis* S.RYadav & Janarth.); one species (*T. submersa* Hook.f.) occurs in both Tasmania and mainland Australia [8, 9, 12]. Molecular phylogeny of Hydatellaceae recognize four sections, *Altofinia, Hamannia, Hydatella* and *Trithuria*) based on sequences from the plastid regions (*atpB, matK, ndhF,* and *rbcL*) as well as the nuclear ribosomal internal transcribed spacer (nrITS) [13].

Trithuria konkanensis, the sole Indian species is placed in the section *Hammania* alongside *T. lanterna* D.A.Cooke and *T. polybracteata* D.A.Cooke ex D.D.Sokoloff, Remizowa, T.D.Macfarl. & Rudall [13]. *Hammania* and *Altofinia* (*T. cookeana* D.D.Sokoloff, Remizowa, T.D.Macfarl. & Rudall and *T. cowieana* D.D.Sokoloff, Remizowa, T.D.Macfarl. & Rudall) form the tropical clade of the genus. *T. lanterna* is sister to *T. konkanensis* and both are autogamous [8, 14]. The latter represents a recent long-distance dispersal event from Northern Australia [13]. Plastome based phylogeny of Hydatellaceae is lacking and so far plastome data are available for three species: *T. filamentosa* [15], *T. inconspicua* [16] and *T. lanterna* (GenBank accession MN365661, unpublished). Of these, *T. filamentosa* has the largest (180,562 bp) plastid genome among the Nymphaeales [7].

In this paper, we report complete plastid genome of *Trithuria konkanensis*. The plastid genome sequence was used to: (1) elucidate the plastome structure of *T. konkanensis*; (2) to investigate the variation of plastomes within Hydatellaceae, among ANA grade angiosperms, and in comparison with representative gymnosperms; (3) evaluate nucleotide diversity and identify highly variable plastid loci with potential phylogenetic utility; (4) assess phylogenetic relationships among ANA grade members, with two gymnosperm species included as outgroups; and (5)

compare the plastome-based phylogenetic relationships with those inferred from nrITS sequences. The addition of *T. konkanensis* plastome adds to the sampling of early-diverging angiosperms.

Materials and methods

Plant sampling

Individuals of *Trithuria konkanensis* were collected from Achirne, Vaibhavwadi taluka, Sindhudurga District, Maharashtra, India (16°26'20.6"N 73°45'43.6"E). The plants were identified by one of us (Manoj Madhwanand Lekhak). Vouchers (MML-802) were deposited in the Shivaji University Herbarium (SUK), Kolhapur, India. Herbarium acronym follows *Index Herbariorum* (https://sweetgum.nybg.org/science/ih/). The complete plastid genome sequence of *T. konkanensis* generated in this study has been deposited in the National Center for Biotechnology Information (NCBI) GenBank database under the accession number QB136381.

DNA isolation and next-generation sequencing

Fresh leaves of *Trithuria konkanensis* were used to isolate DNA by using the DNeasy Plant Kit (Qiagen, Hilden, Germany). The quantity and purity of DNA was checked using the Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA). DNA was quantified using the Qubit™ 4 Fluorometer (Thermo Fisher Scientific, USA) and further evaluated for its integrity by using agarose gel (Lonza, Belgium) electrophoresis (1% gel). Next-generation sequencing (NGS) library preparation was performed using KAPA HyperPlus Kit (Roche Sequencing Solutions, Inc Pleasanton, California, USA) and the library purification was done using the KAPA Pure Beads (Roche Sequencing Solutions, Inc USA). Library quantification was done using the Qubit™ 4 and the library length assessment was done using a DNA high-sensitivity chip using a bioanalyzer (Agilent Technologies, Inc. Santa Clara, California, USA). Sequencing

was performed on NovaSeq 6000 System (Illumina, Inc., San Diego, California, USA). The NGS was outsourced to Nucleome Informatics Private Limited (Hyderabad, Telangana, India).

Plastome assembly and annotation

A total of 6,66,75,606 paired-end reads (151 bp) were generated for *Trithuria konkanensis*. These reads were assembled *de novo* using GetOrganelle v1.7.7.1 [17] via the GALAXY platform [18]. Assembled plastomes were annotated by using PlastidHub platform [19]. PGA 2.0 (Plastid Genome Annotator) was used for plastome annotation [20]. Annotations were further edited in Geneious R10 (Biomatters, New Zealand) [21] and then submitted to GenBank. A circular map of the annotated plastomes was drawn using the OGDRAW program v1.2 [22].

Repeat sequence analysis

Simple sequence repeats (SSRs) were detected using MISA-web v2.1 [23] with minimal repeat numbers of 10, 5, 4, 3, 3, and 3 for mono-, di-, tri-, tetra-, penta-, and hexa-nucleotides, respectively. Tandem repeats were identified using Tandem Repeats Finder (TRF) [24], available at https://tandem.bu.edu/trf/trf.html, and inverted repeats were detected using Inverted Repeats Finder (IRF) [25], available at https://tandem.bu.edu/irf/home. The alignment parameters were set with match, mismatch and indel scores of 2, 7, and 7, respectively. The minimum score required to report a repeat was set to 50. The maximum period size was limited to 500, and the maximum tandem repeat array size was set to 2 million base pairs. For inverted repeats, the maximum stem length and maximum loop length were set to 10,000 bp and 1,000 bp, respectively.

Variation in plastome structure and size

A total of 83 plastomes were analysed. Of these, 81 belonged to ANA grade and two Pinophytina (*Cycas revoluta* Thunb. and *Ginkgo biloba* L.). All the plastomes were retrieved from the NCBI

database except for *Trithuria konkanensis* which was generated in the present study. The retrieved plastomes were reannotated. Inverted repeat (IR) boundaries were identified based on repeat finder plugin in Geneious R10 (Biomatters, New Zealand) [21]. Plastome boundary regions (IRs) for representative taxa of Nymphaeales were visualized using IRplus [26]. Minor discrepancies of 1–3 bp in IR lengths were observed in some taxa. *Trithuria lanterna* showed a larger discrepancy (~100 bp) (in LSC/IR/SSC) between GenBank annotation and IRplus output; GenBank annotation values were retained in the text, with IRplus-reported lengths provided for reference. *Trithuria filamentosa* was excluded from this analysis due to platform anomalies that prevented successful loading.

Differences in plastome length distributions between Hydatellaceae members were evaluated using two-rank-sum tests implemented in R version 2026.01.1+403. using the ggpubr, dplyr, and ggplot2 packages.

Nucleotide diversity analysis and identification of hypervariable regions

The genomes of four *Trithuria* species were aligned using MAFFT v7 [27] Nucleotide diversity (π) across the plastome was estimated using a sliding window analysis, and gene-wise π values were calculated using the DNA polymorphism function implemented in DnaSP v6.12.03 [28]. Sliding window analysis was performed with a window length of 600 bp and a step size of 200 bp. One copy of the inverted repeat (IR) region was excluded prior to analysis. Sequence divergence and gene-wise nucleotide diversity (π) were calculated using the 79 shared protein-coding genes; the putative gene *ycf15* was excluded because of its pseudogenized status in four *Trithuria* species (Fig. S1).

Phylogenetic analysis

Phylogenetic analyses were conducted using four independent datasets: (1) protein-coding genes (PCG), comprising all or a defined subset of plastid protein-coding loci; (2) the complete plastome (FP), complete plastome with one IR removed; and (3) the nuclear internal transcribed spacer region (ITS), encompassing the ITS1, 5.8S rRNA gene, and ITS2 regions. The ITS-based phylogenetic analyses included 120 taxa, whereas the plastome dataset contained only 83 taxa, as sequences for the remaining taxa were unavailable. The best-fit substitution model for each dataset was determined using ModelFinder as implemented in IQ-TREE 3 [29], applying the Bayesian Information Criterion (BIC) for model selection. Bayesian phylogenetic analyses were performed using MrBayes v3.2.7a [30] on the CIPRES Science Gateway [31]. For each of the plastid datasets (PCG and FP), two independent runs of Markov chain Monte Carlo (MCMC) were conducted for 10 million generations, each with four chains (one cold and three heated) sampled every 1,000 generations. For the ITS dataset, analyses were run for 5 million generations under the same sampling scheme. The substitution model applied was GTR+G or GTR+I+G depending on the best-fit model identified by ModelFinder for each respective dataset. The first 10% of sampled trees were discarded as burn-in (burninfrac = 0.1), and convergence was assessed by examining the average standard deviation of split frequencies, which was required to fall below 0.01.

Maximum likelihood (ML) analyses were carried out in IQ-TREE 3 [29] using the best-fit substitution model identified by ModelFinder [32] for each dataset. Node support was assessed using the ultrafast bootstrap approximation (UFBoot) [33] with 1,000 replicates. All resulting phylogenetic trees were visualized and edited using FigTree v1.4.4 [34].

Results

Plastome structure of *Trithuria konkanensis*

The assembled plastome of *Trithuria konkanensis* was 179,088 bp in length (Fig. 1). The assembled plastome showed an average sequencing depth of 407×. The large single-copy (LSC) and small single-copy (SSC) regions were 68,412 and 6,052 bp long, respectively. The inverted repeat (IR) region was 52,312 bp long (Table 1). The plastome contained a total of 114 unique genes, including 80 protein-coding genes (CDS), 30 transfer RNA (tRNA) genes, and four ribosomal RNA (rRNA) genes, with an overall GC content of 39.4%. A total of 33 CDS were duplicated in the IR regions (Table S1). Complete loss of the *clpP1* introns was observed across all four *Trithuria* species examined (Table S1).

Variation in plastome size and structure in ANA grade and Pinophytina

Comparative plastome characteristics of ANA grade, Hydatellaceae and Pinophytina is presented in Table 1. The genome size across 83 taxa ranged from 142,724 bp (*Illicium anisatum* L.) to 180,562 bp (*Trithuria filamentosa*) (Fig. S2). Pinophytina members, i.e. *Cycas revoluta* and *Ginkgo biloba* had 162,489 and 156,990 bp long genomes, respectively. The LSC region exhibited substantial variation, spanning from 68,412 in *T. konkanensis* to 101,578 bp in *I. dunnianum* Tutcher. The IR region was the longest in *T. filamentosa* (52,715 bp) and the shortest in *I. micranthum* Dunn (10,814 bp). The length of SSC region exhibited almost four-fold variation with a minimum in *T. lanterna* (6,015 bp) and maximum in *C. revoluta* (23,377 bp). The GC% varied from 38.1% (*Cabomba aquatica* and *C. furcata*) to 39.8% (*Kadsura longipedunculata* Finet & Gagnep. and *Schisandra macrocarpa* Q.Lin & Y.M.Shui).

Pairwise comparisons between the ANA grade and Hydatellaceae revealed significant differences in plastome size and all genomic partitions. Total plastome length differed significantly (p = 0.00082), with Hydatellaceae exhibiting substantially larger plastomes (169–180 kb) compared to the smaller and more variable ANA grade (142–164 kb). This difference is driven by

contrasting structural patterns across regions. The LSC region was significantly reduced in Hydatellaceae (p = 0.000959; 68–84 kb) relative to ANA grade (81–101 kb). In contrast, the IR region was significantly expanded in Hydatellaceae (p = 0.000818; 37–52 kb) compared to ANA grade (10–31 kb). Similarly, the SSC region was significantly smaller in Hydatellaceae (p = 0.00082; 6 kb) than in ANA grade (16–23 kb). Plastome size comparisons showed that Hydatellaceae occupy a distinct range relative to ANA grade and Pinophytina (Fig. 2).

IR/LSC/SSC junction analysis of Nymphaeales

The *IR/LSC/SSC* junctions are depicted in Fig. 3. The LSC/IRb boundary was consistently located between *rps19* and *rpl*2 across the analyzed ANA grade members except for *Trithuria filamentosa*, *T. konkanensis* and *T. lanterna* where the boundary was between *rps18* and *rpl20*. The SSC/IRb junction was entirely positioned within the *ndhF* region, while the SSC/IRa boundary occurred within the *ycf1* gene in all taxa examined except for Hydatellaceae where it was located within *ndhD* region. The LSC/IRa junction was found between *rpl2* and *psbA* except for *T. filamentosa, T. konkanensis* and *T. lanterna* where it was between *rpl20* and *psbA*.

Distribution of repeats

A total of 181 unique SSR motifs were identified (Fig. 4 and Table S2) across 83 taxa. Mononucleotide repeats (A/T) were the most abundant in all the species. The total number of A and T repeats was 1212 and 1167, respectively (Table S2). The total number of SSRs ranged from 20 (*Nymphaea amazonum* Mart. & Zucc., *N. glandulifera* Rodschied, *N. jamesoniana* Planch. and *N. lasiophylla* Mart. & Zucc.) to 113 (*Illicium floridanum* J.Ellis). For all the *Illicium* species investigated here, the SSR count was more than 90.

In *Trithuria*, SSRs ranged from 41 (*T. filamentosa* and *T. inconspicua*) to 68 (*T. konkanensis*) (Table 2). Hexanucleotide repeats were absent in *T. konkanensis* and *T. lanterna* whereas pentanucleotide was only present in *T. konkanensis* (Fig. 5).

The number of tandem repeats ranged from 19 (*Barclaya longifolia* Wall.) to 88 (*Cabomba furcata* Schult. & Schult.f.) and the inverted repeats 1 to 6 (*Kadsura interior* A.C.Sm.). Inverted repeats were absent in *Amborella trichopoda* Baill., *Brasenia schreberi* J.F.Gmel., *C. furcata* and *G.biloba*. A single inverted repeat was found in all the *Trithuria* species. Tandem repeats ranged from 36 (*T. inconspicua*) to 61 (*T. lanterna*) (Fig. S3 and Table S3).

Nucleotide diversity analysis and identification of hypervariable regions

Nucleotide diversity within Hydatellaceae was assessed using the aligned plastome dataset (one IR region excluded), comprising 132,242 sites, of which 123,870 ungapped sites were retained for analysis. A total of 5,320 polymorphic (segregating) sites were identified across the alignment. The mean nucleotide diversity (π) across the analyzed noncoding plastome region was 0.02801. Sliding window analysis revealed the hotspots of sequence variability in the LSC, IR and SSC portions of the chloroplast genome. Coding sequences exhibited substantial conservation across species, whereas non-coding regions displayed markedly higher levels of sequence divergence. Among all regions, the intergenic spacer *rps12/ycf1* exhibited the highest level of variability. Five other highly polymorphic intergenic spacers were also identified, viz. *ndhC/atpE, ycf4/cemA, psbE/petL, ndhF/rpl32* and *rpl32/ccsA* (Fig. 6). Two coding genes *clpP1* and *ycf1* also showed elevated sequence variation (Fig. 6).

Gene-wise nucleotide diversity (0.0–0.08753) revealed substantial heterogeneity across the plastome, with genes *clpP1* (0.08753), *rpl32* (0.05338), *ccsA* (0.05278), *ndhK* (0.07136), *ndhF* (0.04962), *accD* (0.04444) and *ycf1* (0.03850), exhibiting elevated π values, indicative of mutation

hotspots. Six genes (*petN, psbI, psbM, psbN, rpl14* and *rpl36*) displayed zero π and θ_W values (Fig. 7 and Table S4). Haplotype diversity (Hd) ranged from 0.0 to 1.0. The highest density of genes was concentrated at Hd ≈ 0.65–0.85 and π ≈ 0.005–0.015, as indicated by the innermost contour levels. A secondary distribution extended toward higher diversity values, with several genes exhibiting Hd ≈ 1.0 and elevated π, including *clpP1*, *ndhK*, *rpl32*, *ccsA*, *ndhF*, *accD* and *ycf1*. A small number of genes were located at Hd = 0 and π = 0, representing invariant loci. The contour distribution was unimodal with a pronounced density peak at moderate Hd and low π, and a trailing extension toward higher π values (Fig. 8 and Table S4).

CDS based phylogeny

The protein-coding gene (PCG) dataset comprised 83 taxa and a total aligned length of 84,678 characters, of which 54,992 (64.9%) were constant, 7,792 (9.2%) were variable but parsimony-uninformative, and 22,094 characters (26.1%) were parsimony-informative. The best-fit substitution model selected by ModelFinder under the BIC was GTR+F+I+G4. ModelFinder selected GTR+F+I+G4 as the best-fit model under the BIC. *Amborella* was the earliest divergent lineage among the flowering plants (Fig. 9). The Schisandraceae was sister to a clade containing Hydatellaceae, Cabombaceae and Nymphaeaceae. Within Schisandraceae, *Illicium* L. was monophyletic whereas *Kadsura* Kaempf. ex Juss. and *Schisandra* Michx. were polyphyletic. In Hydatellaceae, *Trithuria konkanensis* was sister to *T. lanterna* and *T. filmentosa* was sister to *T. inconspicua*. The other lineage contained Cabombaceae and Nymphaeaceae which were sister to each other. Cabombaceae contained *Cabomba* Aubl. and *Brasenia* Schreb. Within Nymphaeaceae, *Nuphar* Sm. was the early-divergent lineage followed by *Barclaya* Wall. and *Nymphaea* L. However, *Euryale ferox* Salisb. and *Victoria cruziana* A.D.Orb. were nested in *Nymphaea* making

them polyphyletic. All branches leading to the families were fully supported by Bayesian posterior probabilities (BPP) and Maximum Likelihood bootstrap (MLB).

Plastome based phylogeny

The full plastome (FP) dataset had 83 taxa and a total aligned length of 151,151 characters, comprising 80,766 (53.4%) constant sites, 15,944 (10.5%) variable parsimony-uninformative sites, and 54,944 (36.4%) parsimony-informative characters. The best-fit model selected by ModelFinder under the BIC was TVM+F+R7. The FP phylogeny (Fig. 10) was exactly of the same topology as the PCG dataset except for varied branch lengths. The families in the complete plastome tree were also fully supported by BPP and MLBS. In both the PCG and complete plastome analysis, Hydatellaceae had the longest branch lengths than other families.

nrITS based phylogeny

The nuclear ITS dataset comprised a total aligned length of 1,626 characters, of which 750 (46.1%) were constant, 173 (10.6%) were variable but parsimony-uninformative, and 703 (43.2%) were parsimony-informative. The best-fit substitution model identified by ModelFinder under the BIC was GTR+F+R4. The ITS dataset had more number of taxa than the plastome dataset, 116 taxa and 1626 aligned characters including gaps of which 750 characters were constant (46.1 %), 173 (10.6 %) were variable but parsimony-uninformative and 703 (43.3 %) were parsimony-informative. The best-fit substitution model identified by ModelFinder under the BIC was GTR+F+R3. The topology of the ITS tree (Fig. 11) was highly similar to the plastome datasets. Even in the ITS dataset, *Kadsura* and *Schisandra* were polyphyletic with each other. In Hydatellaceae, *Trithuria konkanensis* was in the same clade as *T. lanterna* though the representation was more in this dataset. Within Nymphaeaceae, *Barclaya* was the early diverging lineage followed by *Nuphar*, which was opposite in the case of the plastome dataset. Also,

313 *Nymphaea* was polyphyletic with *Euryale ferox* and *Victoria amazonica* similar to the plastome
314 data results. However, it was noted that the branch length of Hydatellaceae was not
315 disproportionately longer than the other families which was the case in the plastome dataset.

316 # Discussion

317 **Plastome evolution in Hydatellaceae**

318 Plastome structure was highly conserved across Hydatellaceae, with only negligible structural
319 variation observed among species. Plastome sizes range from 165,389 bp in *Trithuria inconspicua*
320 to 180,562 bp in *T. filamentosa*. The comparatively smaller plastome of *T. inconspicua* is primarily
321 attributable to contraction of the IR regions. Specifically, each IR in *T. inconspicua* measured
322 37,277 bp, whereas the IRs of the remaining species ranged from 52,257 to 52,715 bp.
323 Consequently, *T. inconspicua* had 14 CDS duplicated in the IR whereas in other taxa the number
324 was 33 (Table S1). IR boundary shifts, expansion and contractions are responsible for variation of
325 plastome architecture [3]. Similar expansions have been observed in *Nicotiana acuminata*
326 (Graham) Hook. [35], *Cistanche* Hoffmanns. & Link species [36], *Paphiopedilum* Pfitzer [37]. IR
327 expansion increases overall plastome size and affects gene dosage and regulation [38].

328 Both introns of *clpP1* were absent in *T. konkanensis*, *T. filamentosa*, and *T. lanterna*, whereas *T.*
329 *inconspicua*, which possesses only a single *clpP1* copy, also lacked intron. Intron losses for the
330 *clpP1* genes have been identified in other angiosperm lineages as well. For example, both of the
331 *clpP1* gene introns have been lost in Poaceae, Onagraceae, Oleaceae, and *Pinus* [39] demonstrating
332 that complete loss has evolved repeatedly. Nevertheless, both introns of *clpP1* are retained in the
333 sampled plastomes of other *Nymphaeales*, including representatives of Nymphaeaceae [40] and
334 Cabombaceae (*Brasenia*) [41]. Thus, intron loss appears to be a consistent feature of the *Trithuria*
335 plastomes examined here. The complete loss of both *clpP1* introns represents a notable structural

modification of the plastid genome and adds to the evidence for lineage-specific plastome remodeling in Hydatellaceae. Since *clpP1* introns have been retained in many angiosperm plastomes but have been independently lost in several lineages, their absence in *Trithuria* may reflect lineage-specific intron-loss events during plastome evolution. However, whether the loss occurred once in the common ancestor of the sampled *Trithuria* species or independently in different lineages cannot be resolved without broader taxon sampling.

Structural variation in plastomes across ANA Grade and Pinophytina

Hydatellaceae possesses the largest plastomes among the sampled ANA grade angiosperms. Within the ANA grade families overall size of the plastome across the families is highly conserved. Nevertheless, unusually large differences in IR length (approximately 10.8 kb to 52.7 kb) were observed. Most of the *Illicium* species represented an IR of 10-11 kb. IR reduction in *I. oligandrum* Merr. & Chun is similar to *Cuscuta reflexa* Roxb. at the LSC/IRa junction [42]. Due to contraction, the number of duplicated genes in IR is substantially reduced. *Illicium anisatum* L. has only 10 genes in the IR [43]. Large IR contractions (>10,000 bp) are less frequent than IR expansions in autotrophic vascular plants [44]. Despite such remarkable variation in IR size, GC content remains highly conserved (38.1–39.8%), suggesting that plastome size evolution among these early-diverging angiosperm lineages is predominantly associated with structural modifications rather than shifts in nucleotide composition. The gradient (10.8–25.2 kb) of IR sizes across Schisandraceae suggests IR expansion (*Kadsura* and *Schisandra*) and contraction (*Illicium*) events during the diversification of the family, making it one of the most structurally dynamic plastome groups among basal angiosperms. In contrast, plastomes of Nymphaeaceae were highly conserved with IR lengths remarkably stable at approximately 25–26 kb. This structural conservation is

358 consistent with the comparative analyses of [40] who reported IR regions of approximately 25 kb
359 containing 16 duplicated genes across twelve *Nymphaea* plastomes.

360 Overall, these results demonstrate that IR expansion and contraction constitute the
361 principal mechanisms underlying plastome size evolution among ANA grade angiosperms and
362 their relatives. While Nymphaeaceae and Amborellaceae have retained highly conserved plastome
363 architectures, Schisandraceae, particularly *Illicium*, is characterized by pronounced IR contraction,
364 whereas Hydatellaceae exhibits exceptional IR expansion.

365 **Evolution of IR Boundaries in Nymphaeales**

366 The gene positioning at the *IR/LSC/SSC* junctions was largely conserved across the Nymphaeales,
367 indicating a high degree of structural stability in plastome organization. Similar observations have
368 been made by [45] and [40] in the case of Nymphaeaceae. Hydatellaceae, nevertheless exhibited
369 lineage-specific shifts, particularly within the LSC/IRb and SSC/IRa junctions. This shift is seen
370 in *Trithuria filamentosa*, *T. konkanensis*, and *T. lanterna* that have expanded IRs. A 10 kb IR
371 contraction in Schisandraceae, a member of ANA grade has also led to change in LSC/IRb (*trnL-*
372 *CAA* and *ndhB*) and LSC/IRa (*ndhB* and *trnH-GUG*) [46]. Similarly, the junctions of *IR/LSC* and
373 *IR/SSC* were highly variable is *Cistanche* plastomes due to expansion/contraction of IRs [36]. IR
374 expansion and contraction often results in genome size variation among plant lineages and is
375 important for understanding evolution [47].

376 **Repetitive sequences and plastome evolution**

377 Genome organization, gene activities and DNA recombination, replication and repair are
378 influenced by repetitive sequences [48]. The exceptionally high abundance of SSRs in *Illicium*
379 (92–113 SSRs), nearly two- to five-fold greater than that observed in most species of

Nymphaeaceae indicates that SSR accumulation is largely independent of overall plastome size or IR length and is more likely associated with lineage-specific mutational processes, such as replication slippage and the expansion of non-coding AT-rich regions. Lineage-specific mutations and structural variation is recorded in two closely related species of *Linum* L., *L. usitatissimum* L. and *L. altaicum* Ledeb. ex Juz. suggesting that taxonomic proximity does not necessarily predict microsatellite conservation and phylogenetic proximity [49]. SSRs are widely used to assess genetic polymorphism and investigate genetic diversity and population structure within and among populations [50]. SSRs in *Illicium* can serve as marker for population genetics and species delimitation. The predominance of mononucleotide SSRs suggests that the increase is largely driven by the accumulation of A/T-rich microsatellites, a common feature of plastid genomes. Of the total SSRs measured in Nymphaeacae, 50% were mononucleotides and 90% of these were A/T repeats [40]. Nymphaeales had an average tandem repeat of 43.08 ± 1.72 whereas for Austrobaileyales it was 30.25 ± 1.10. Among the Nymphaeales, high numbers observed in *Cabomba furcata* and *C. aquatica* Aubl. may reflect lineage-specific differences in the gain, expansion, or retention of plastid repeats. Such repeat-rich regions may represent useful targets for developing highly variable plastid markers.

Dispersed inverted repeats were less (0–6) abundant than tandem repeats. This pattern suggests that plastome evolution in ANA grade angiosperms is driven primarily by small-scale sequence variation rather than extensive recombination mediated by dispersed inverted repeats. Whole genome alignment using Mauve also reveal that Nymphaeaceae have a well-conserved chloroplast genomes lacking major inversions [45]. The relatively higher numbers observed in *Kadsura* and *Schisandra* may indicate lineage-specific increases in recombination potential, but

the overall scarcity of dispersed inverted repeats reflects the strong structural constraints acting on plastid genomes.

Hydatellaceae exhibited moderate tandem-repeat abundance despite possessing exceptionally expanded IR regions, indicating that repeat accumulation and IR expansion are not necessarily correlated.

Sequence divergence and hypervariable regions in Hydatellaceae

Nucleotide diversity analyses revealed distinct patterns of sequence variation among the examined lineages. In Hydatellaceae, the highly variable *rps12/ycf1* intergenic spacer was located within the IR region, whereas other polymorphic intergenic spacers, including *ndhC/atpE, ycf4/cemA, and psbE/petL*, occurred in the LSC region, and *ndhF/rpl32* and *rpl32/ccsA* were located in the SSC region. Among coding regions, *clpP1* (LSC) and *ycf1* (IR) exhibited elevated nucleotide diversity. These findings indicate that sequence variation in Hydatellaceae is not restricted to the single-copy regions but extends into the IR, a genomic compartment that is generally considered more conserved because of copy-correction through gene conversion.

In contrast, previous studies in Nymphaeaceae have shown that highly variable loci are predominantly associated with coding genes. For example, Zhang et al. [51] identified *psaI, rps19, ndhF, rps15*, and *ycf1* as the most variable genes in *Nymphaea nouchali*, highlighting their potential as plastid molecular markers. Similarly, comparative plastome analyses of *Illicium* species demonstrated that sequence divergence is concentrated primarily in intergenic spacers, with *ndhB–trnH, atpH–atpI, rpoB–trnC,* and *ndhG–ndhI* identified as the most variable regions, of which *ndhB–trnH* exhibited the highest nucleotide diversity [52].

Gene-wise nucleotide diversity revealed marked heterogeneity across the plastome, indicating unequal evolutionary rates among plastid genes. The elevated π values observed in *clpP1, ndhK, rpl32, ccsA, ndhF, accD,* and *ycf1* identify these loci as mutation hotspots, consistent with previous plastome studies that have recognized these genes as rapidly evolving and informative for phylogenetic inference and DNA barcoding [40, 53, 54]. In contrast, the complete conservation of *petN, psbI, psbM, psbN, rpl14*, and *rpl36* reflects strong purifying selection acting on genes with essential photosynthetic and ribosomal functions.

The Hd–π distribution further demonstrated that most plastid genes possess moderate haplotype diversity but low nucleotide diversity, indicating limited sequence divergence despite the presence of multiple haplotypes. Only a small subset of genes exhibited both high haplotype and nucleotide diversity, suggesting that these loci contribute disproportionately to plastome evolution.

Phylogeny of Nymphaeales

Song et al. [40] performed the most comprehensive phylogenomic analysis of Nymphaeales with 43 taxa based on 79 CDS and the complete chloroplast genomes. Our results were comparable to their study except that our sampling of the order Nymphaeales was expanded to 83 taxa. Our sampling of the ANA grade is the most comprehensive to our knowledge. Our phylogeny recovered a strongly supported monophyletic Nymphaeaceae that is clearly separated from Cabombaceae and Hydatellaceae consistent with recent analysis [40, 55]. In our PCG analysis, *Amborella* was the earliest divergent lineage, however, it was sister to the *Illicium-Schisandra-Kadsura* lineage of the order Austrobaileyes. Such minor descrepancies arise due to varying gene markers and methods of phylogenetic analysis. For example, in Goremykin et al. [16] *Amborella*

Baill. was found sister to *Nuphar* and closely related to Nymphaeales and *Trithuria*. Nuclear genome analyses [56] support maintaining *Victoria* R.H.Schomb. and *Euryale* Salisb. as separate genera. However, chloroplast phylogeny nests them within the *Nymphaea* clade [40]. Our nuclear ITS analysis also places *Euryale* ferox and *Victoria amazonica* inside *Nymphaea* sister to the clade containing *Nymphaea odorata*. However, earlier reports using ITS analysis and plastid *trnT-trnF,* placed *Euryale* and *Victoria* outside the *Nymphaea* [57].

Similary, several other studies also show polyphyly of *Schisandra* and *Kadsura* [58, 59]. Historical biogeography study among the *Trithuria* showed a close relationship between the Australian *A. lanterna* and *T. konkanensis* and also between *T. filamentosa* and *T. inconspicua*. The split between *T. konkanesis* and *T. lanterna* is quite recent, during the Pleistocene and therefore it is probably by a long-distance dispersal event from a north Australian lineage of *Trithuria* [13].

A conspicuous difference between the whole-plastome (Fig. 10) and nuclear ribosomal ITS (Fig. 11) phylogenies was observed in the branch lengths associated with *Trithuria*. The plastome phylogeny exhibited a markedly elongated stem branch leading to *Trithuria*, indicating substantial plastome sequence divergence relative to other lineages of Nymphaeales. In contrast, the ITS phylogeny did not exhibit a comparably pronounced elongation of the *Trithuria* stem branch. This discrepancy suggests that the relative accumulation of nucleotide substitutions along the Hydatellaceae lineage differs between the plastid and nuclear ribosomal datasets. The contrasting branch-length patterns may reflect differences in evolutionary rates and substitution dynamics between the two genomic compartments. The long branch in the plastome phylogeny could indicate lineage-specific acceleration of plastid sequence evolution, although a prolonged period of independent evolution may also contribute to the observed divergence. The comparatively less pronounced branch in the ITS phylogeny suggests that the same pattern of relative divergence is

not evident in the nuclear ribosomal region. Similar heterogeneity in molecular evolutionary rates and branch-length patterns has been documented in other angiosperm lineages, such as Annonaceae and *Populus* L. [60, 61]. However, the contrasting patterns observed here may also be influenced by differences in taxon sampling, alignment characteristics, and evolutionary models, and therefore require further investigation.

Conclusion

Our study expands the plastome resources available for early-diverging angiosperms by providing a comprehensive plastome analysis of *Trithuria konkanensis* and comparative insights into Hydatellaceae. Among the currently sampled members of ANA grade members, Hydatellaceae possess the largest plastomes, reaching approximately 180 kb. Despite variation in plastome size, plastome architecture remains highly conserved across the family, indicating remarkable structural stability during evolution. Nevertheless, only four of the twelve recognized Hydatellaceae species have been investigated to date. Future plastome sequencing of the remaining species, together with the generation of complete mitochondrial and nuclear genome resources, will provide a more comprehensive understanding of organellar and nuclear genome evolution, genome structural dynamics, phylogenetic relationships, and the evolutionary history of this ancient angiosperm lineage.

Acknowledgements MML and RNC are thankful to the Head, Department of Botany, Shivaji University, Kolhapur (SUK) for providing the necessary research facilities. SS thanks Institution of Eminence grant No. UoH-IoE-RC5-22-042 to the University of Hyderabad, Hyderabad for funding.

Authors' contributions Manoj M. Lekhak: Conceptualisation, plant sampling, methodology, investigation, writing-original draft, Rupali N. Chougule: Plant sampling, writing–review and editing. Asif. S. Tamboli: Methodology, investigation, writing–review and editing. Siddharthan Surveswaran: Conceptualisation, methodology, investigation, writing–review and editing and funding acquisition.

Funding Not applicable.

Data availability The chloroplast genome sequences used in this study have been deposited in the NCBI database.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. Jarvis P, López-Juez E. Biogenesis and homeostasis of chloroplasts and other plastids. Nat Rev Mol Cell Biol. 2013; 14:787–802. https://doi.org/10.1038/nrm3702
2. Daniell H, Lin CS, Yu M, Chang WJ. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol. 2016;17:134. doi: 10.1186/s13059-016-1004-2.
3. Cauz-Santos LA. Beyond conservation: the landscape of chloroplast genome rearrangements in angiosperms. New Phytol. 2025;247: 2571–80. https://doi.org/10.1111/nph.70364
4. Bellot S, Renner SS. The plastomes of two species in the endoparasite genus *Pilostyles* (Apodanthaceae) each retain just five or six possibly functional genes. Genome Biol Evol. 2015;8:189–201.
5. Weng ML, Ruhlman TA, Jansen RK. Expansion of inverted repeat does not decrease substitution rates in *Pelargonium* plastid genomes. New Phytol. 2017;214:842–51.
6. The Angiosperm Phylogeny Group. Large-scale nuclear and plastid phylogenomic analyses inform an updated Angiosperm Phylogeny Group classification: APG V. J Syst Evol. 2026;64:845–74. https://doi.org/10.1111/jse.70096
7. Gruenstaeudl M, Nauheimer L, Borsch T. Plastid genome structure and phylogenomics of Nymphaeales: Conserved gene order and new insights into relationships. Plant Syst Evol. 2017;303:1251–70.
8. Sokoloff DD, Remizowa MV, Macfarlane TD, Rudall PJ. Classification of the early-divergent angiosperm family Hydatellaceae: one genus instead of two, four new species and sexual dimorphism in dioecious taxa. Taxon. 2008;57:179–200.
9. Sokoloff DD, Remizowa MV, Macfarlane TD, Yadav SR, Rudall PJ. Hydatellaceae: a historical review of systematics and ecology. Rheedea. 2011;21:115–38.
10. Hamann U. Hydatellaceae — a new family of monocotyledoneae. New Zeal J Bot. 1976;14:193–6.
11. Saarela JM, Rai HS, Doyle JA, Endress PK, Mathews S, Marchant AD et al. Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. Nature. 2007;446:312–15.
12. Yadav SR, Janarthanam MK. Hydatellaceae: a new family to Indian flora with a new species. Rheedea. 1994;4:17–20.
13. Iles WJD, Lee C, Sokoloff DD, Remizowa MV, Yadav SR, Barrett MD, et al. Reconstructing the age and historical biogeography of the ancient flowering-plant family Hydatellaceae (Nymphaeales). BMC Evol Biol. 2014;14:102.
14. Sokoloff DD, Remizowa MV, Yadav SR, Rudall PJ. Development of reproductive structures in the sole Indian species of Hydatellaceae, *Trithuria konkanensis*, and its morphological differences from Australian taxa. Aust Syst Bot. 2010;23:217–28.
15. Drew BT, Ruhfel BR, Smith SA, Moore MJ, Briggs BG, Gitzendanner MA, et al. Another look at the root of the angiosperms reveals a familiar tale. Syst Biol. 2014;63:368–82.
16. Goremykin VV, Nikiforova SV, Biggs PJ, Zhong B, Delange P, Martin W, et al. The evolutionary root of flowering plants. Syst Biol. 2013;62:50–61.
17. Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, et al. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. Genome Biol. 2020;21:241.

18. The Galaxy Community. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2024 update. Nucleic Acids Res. 52:W83–W94. https://doi.org/10.1093/nar/gkae410
19. Zhang NN, Stull GW, Zhang XJ, Fan SJ, Yi TS, Qu XJ. PlastidHub: An integrated analysis platform for plastid phylogenomics and comparative genomics. Plant Divers. 2025;47:544–60.
20. Qu XJ, Moore MJ, Li DZ, Yi TS. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods. 2019;15:50.
21. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–49.
22. 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;52:267–74.
23. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33:2583–85. https://doi.org/10.1093/bioinformatics/btx198
24. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res.1999;27:573–80.
25. Warburton PE, Giordano J, Cheung F, Gelfand Y, Benson G. Inverted Repeat Structure of the Human Genome: The X-chromosome contains a preponderance of large, highly homologous inverted repeats that contain testes genes.Genome Res. 2004;14:1861–69.
26. Díez Menéndez C, Poczai P, Williams B, Myllys L, Amiryousefi A. IRplus: An augmented tool to detect inverted repeats in plastid genomes. Genome Biol Evol. 2023;15:evad177.
27. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
28. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large datasets. Mol Biol Evol. 2017;34:3299–302.
29. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New models and methods for phylogenetic inference. Mol Biol Evol. 2020;37:1530–34.
30. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model selection across a large model space. Syst Biol. 2012;61:539–42.
31. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE). 2010. p. 1–8.
32. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14:587–589.
33. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35:518–522.
34. Rambaut A. FigTree version 1.4.4. 2018.
35. Goulding SE, Olmstead RG, Morden CW, Wolfe KH. Ebb and flow of the chloroplast inverted repeat. Mol Gen Genet.1996;252:195–206.

36. Miao Y, Chen H, Xu W, Yang Q, Liu C, Huang L. Structural mutations of small single copy (SSC) region in the plastid genomes of five *Cistanche* species and inter-species identification. BMC Plant Biol.2022;22:412. https://doi.org/10.1186/s12870-022-03682-x
37. Guo YY, Yang JX, Bai MZ, Zhang GQ, Liu ZJ.The chloroplast genome evolution of Venus slipper (*Paphiopedilum*): IR expansion, SSC contraction, and highly rearranged SSC regions. BMC Plant Biol. 2021;21:248. https://doi.org/10.1186/s12870-021-03053-y
38. Krämer C, Boehm CR, Liu J, Ting MKY, Hertle AP, Forner J, et al. 2024. 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: 923–35.
39. Jansen RK, Cai Z, Raubeson LA, Daniell H, Depamphilis CW, Leebens-Mack J, et al. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA. 2007;104:19369–74. https://doi.org/10.1073/pnas.070912110404
40. Song W, Shi W, Wang H, Zhang Z, Tao R, Liu J, et al. Comparative analysis of 12 water lily plastid genomes reveals genomic divergence and evolutionary relationships in early flowering plants. Mar Life Sci Technol. 2024;6:425–41.
41. Sun, Y., Li, M., Ma, J., He, M., & Zheng, Y. Complete chloroplast genome sequence of a new variety of *Brasenia schreberi*: Genome characteristics, comparative analysis, and phylogenetic relationships. Agronomy.2022;12: 2972. https://doi.org/10.3390/agronomy12122972
42. Hansen DR, Dastidar SG, Cai Z, Penaflor C, Kuehl JV, Boore JL et al. Phylogenetic and evolutionary implications of complete chloroplast genome sequences of four early-diverging angiosperms: *Buxus* (Buxaceae), *Chloranthus* (Chloranthaceae), *Dioscorea* (Dioscoreaceae), and *Illicium* (Schisandraceae). Mol Phylogenet Evol. 2007;45: 547–63.
43. Park J, Kim Y, Xi H. The complete chloroplast genome of aniseed tree, *Illicium anisatum* L. (Schisandraceae). Mitochondrial DNA B.2019;4:1023–4.
44. Mower JP, Vickrey TL. Structural diversity among plastid genomes of land plants. In: Chaw SM, Jansen RK, editors. Advances in Botanical Research. Vol. 85. Academic Press; 2018. p. 263–292.
45. He DX, Gichira AW, Li ZZ, Nzei JM, Guo YH, Wang QF, et al. Intergeneric relationships within the early-diverging angiosperm family Nymphaeaceae based on chloroplast phylogenomics. Int J Mol Sci. 2018;19:9–11
46. Li B, ZhengY. Dynamic evolution and phylogenomic analysis of the chloroplast genome in Schisandraceae. Sci Rep. 2018; 8;9285. https://doi.org/10.1038/s41598-018-27453-7
47. Wang R-J, Cheng C-L,, Chang C-C, Wu C-L, Su T-M, Chaw S-M. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. BMC Evol Biol.2008; 8:36. https://doi.org/10.1186/1471-2148-8-36
48. Li Y-C, Korol AB, Fahima T, Beiles A, Nevo E. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol. 2002;11: 2453–65. https://doi.org/10.1046/j.1365-294X.2002.01643.x
49. Raut VK, Ngangkham U, Yadav A, Kaur V, Malav PK, Saroha A. et al. Cross-species transferability and genetic diversity analysis in *Linum* species using microsatellite markers. Sci Rep. 2026;16:13358. https://doi.org/10.1038/s41598-026-42744-0
50. Vieira ML, Santini L, Diniz AL, Munhoz Cde F. Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol. 2016;39:312–28. doi: 10.1590/1678-4685-GMB-2016-0027.

51. Zhang H, Si Y, Zhao R, Sheng Q, Zhu Z. Complete chloroplast genome and phylogenetic relationship of *Nymphaea nouchali* (Nymphaeaceae), a rare species of water lily in China. Gene. 2023;858:147139.
52. Cao Y, Lai Y, Li Z, Zhai S, Dai Y, Tao J, et al. The complete chloroplast genome of *Illicium verum* and comparative analysis with related species from Magnoliaceae and Illiciaceae. Front Genet.2024;15:1452680.
53. Thode VA, Lohmann LG. Comparative chloroplast genomics at low taxonomic levels: A case study using *Amphilophium* (Bignonieae, Bignoniaceae). Front Plant Sci. 2019;10:796. doi: 10.3389/fpls.2019.00796
54. Chen L, Song B-N, Yang L, Wang Y, Wang Y-Y, Aou X, et al. Phylogeny, adaptive evolution, and taxonomy of *Acronema* (Apiaceae): evidence from plastid phylogenomics and morphological data. Front Plant Sci. 2024;15:1425158. https://doi.org/10.3389/fpls.2024.1425158
55. Roestel JA, Wiersema JH, Jansen RK, Borsch T, Gruenstaeudl M. On the importance of sequence alignment inspections in plastid phylogenomics - an example from revisiting the relationships of the water-lilies. Cladistics. 2024;40:469–95. https://doi.org/10.1111/cla.12584
56. Zhang L, Chen F, Zhang X, Li Z, Zhao Y, Lohaus R, et al. The water lily genome and the early evolution of flowering plants. Nature. 2020;577:79–84.
57. Pellicer J, Kelly LJ, Magdalena C, Leitch IJ. Insights into the dynamics of genome size and chromosome evolution in the early diverging angiosperm lineage Nymphaeales (water lilies). Genome. 2013;56:437–49. https://doi.org/10.1139/gen-2013-0039
58. Wang YH, Zhang SZ, Gao JP, Chen DF. Phylogeny of the Schisandraceae based on cpDNA *matK* and *rpl16* intron data. Chem Biodivers. 2006;3:359–369.
59. Chen H, Bashir NH, Zhu L, Zhao TX, Xu T, Gao Y, et al. The complete chloroplast genome and phylogenetic analysis of a rare species *Schisandra macrocarpa* (Schisandraceae). Mitochondrial DNA B Resour.2024;9:1693–97.
60. Hoekstra PH, Wieringa JJ, Smets E, Brandao RD, de Carvalho Lopes J, Erkens RH, et al. Correlated evolutionary rates across genomic compartments in Annonaceae. Mol Phylogenetics Evol. 2017;114: 63–72.
61. Huang DI, Hefer CA, Kolosova N, Douglas CJ, Cronk QCB. (2014), Whole plastome sequencing reveals deep plastid divergence and cytonuclear discordance between closely related balsam poplars, *Populus balsamifera* and *P. trichocarpa* (Salicaceae). New Phytol.2014;204: 693–703. https://doi.org/10.1111/nph.12956

Table legends

Table 1. List of plastome sequences used for comparative genomic analyses, showing species, GenBank accession number, family, total plastome length, lengths of the large single-copy (LSC), inverted repeat (IR), and small single-copy (SSC) regions, and overall GC content.

Table 2. Distribution of simple sequence repeats (SSRs) in the plastid genomes of four *Trithuria* species.

Figure legends

Fig. 1. Circular maps of the complete plastid genomes of *Trithuria inconspicua* and *T. konkanensis*. Note expanded inverted repeats (IRs) in *T. konkanensis.*

Fig. 2. Comparative plastome size variation across major genome regions in ANA grade angiosperms, Hydatellaceae, and Pinophytina. Boxplots depict the distribution of sequence lengths (bp) for the total plastome, LSC, IR and SSC regions across sampled taxa (n = 83). Boxes represent the interquartile range (IQR), the horizontal line within each box indicates the median, and whiskers extend to 1.5× IQR. Individual points beyond this range are shown as outliers. Green diamonds denote mean values. Statistical significance between Hydatellaceae and ANA-grade angiosperms was assessed using the Wilcoxon rank-sum test, with corresponding p-values indicated within each panel. Colors differentiate taxonomic groups as shown.

Fig. 3. Comparison of the borders of large single-copy (LSC), small single-copy (SSC), and inverted repeat (IR) regions among representative members of Cabombaceae (*Brasenia* and *Cabomba*), Nymphaeaceae (*Barclaya, Euryale, Nuphar, Nymphaea* and *Victoria* and Hydatellaceae (*Trithuria*).

Fig. 4. Heatmap showing the distribution of unique Simple Sequence Repeat (SSR) repeat motifs across 83 species. The colored cells and light grey cells indicate presence and absence of motifs, respectively. The color intensity within each cell represents the count of unique SSRs within that species, ranging from 0 to 40.

Fig. 5. Distribution of simple sequence repeats (SSRs) according to repeat motif length in the plastomes of *Trithuria* species.

Fig. 6. Sliding-window analysis of plastomes of four species of *Trithuria* (excluding one IR region). X-axis represents the nucleotide position, while Y-axis indicates the nucleotide diversity (Pi). The eight most variable regions are highlighted.

Fig. 7. Gene-wise nucleotide diversity (π) across plastid protein-coding genes ranked in descending order. Bars represent π, while points and connecting lines indicate Watterson's θ (θ_W). Genes exhibiting higher π values correspond to more variable loci across sampled taxa.

Fig. 8. Relationship between haplotype diversity (Hd) and nucleotide diversity (π) across plastid protein-coding genes. Points represent individual genes; the eight genes with the highest π values are highlighted in black and labeled in italics, whereas remaining genes are shown in grey. Solid contour lines denote two-dimensional kernel density estimates, illustrating the distribution of genes in Hd–π space. Most genes cluster within regions of moderate Hd (≈0.6–0.85) and low π, indicating overall sequence conservation, while a subset of genes with elevated π fall outside the main density contours, suggesting accelerated sequence evolution.

Fig. 9. Phylogram of ANA grade and Pinophytina inferred from maximum likelihood (ML) and Bayesian inference (BI) based on protein coding genes (PCGs) of 83 plastomes. Numbers above The branches are Bayesian posterior probabilities (BPP) and maximum-likelihood bootstrap support (BS) values, respectively. Sequences generated for this study are indicated in bold face.

Fig. 10. Phylogram of ANA grade and Pinophytina inferred from maximum likelihood (ML) and Bayesian inference (BI) based on 83 complete plastomes, excluding one copy of the inverted repeat (IR). Numbers above branches are Bayesian posterior probabilities (BPP) and maximum-likelihood bootstrap support (BS) values, respectively. Sequences generated for this study are indicated in bold face.

Fig. 11. Phylogram of ANA grade and Pinophytina inferred from maximum likelihood (ML) and Bayesian inference (BI) based on ITS sequences of 120 taxa. Numbers above branches are Bayesian posterior probabilities (BPP) and maximum-likelihood bootstrap support (BS) values, respectively.

Supplementary material

Table S1. Gene content of the plastomes of *Trithuria* species.

Table S2. Distribution of simple sequence repeat (SSR) motif types across 83 taxa. Values represent the number of SSRs corresponding to each motif type identified in the plastome of each taxon.

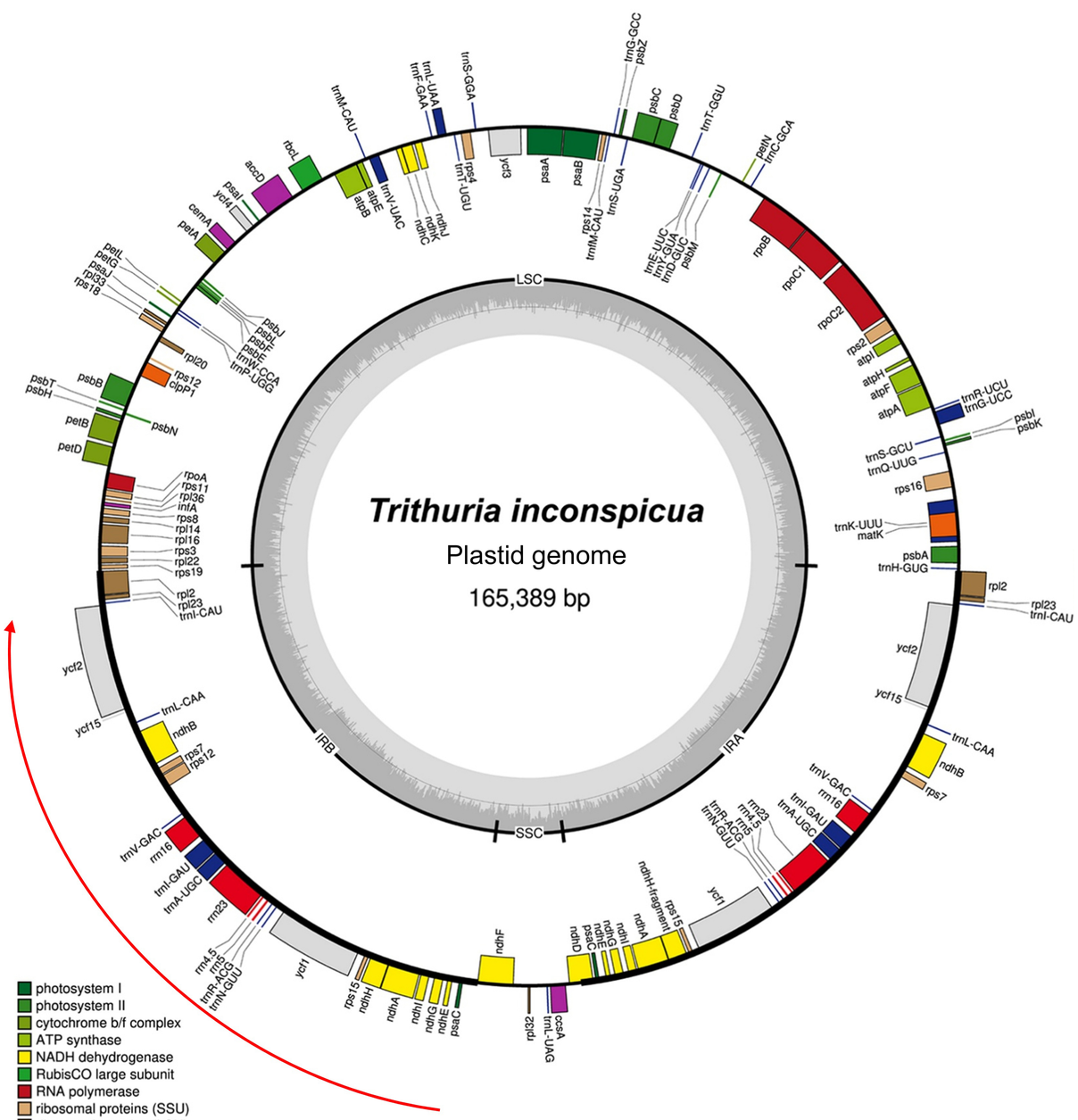
Table S3. Distribution of tandem and inverted repeats across taxa. Values represent the number of tandem and inverted repeats identified in the plastome of each taxon.

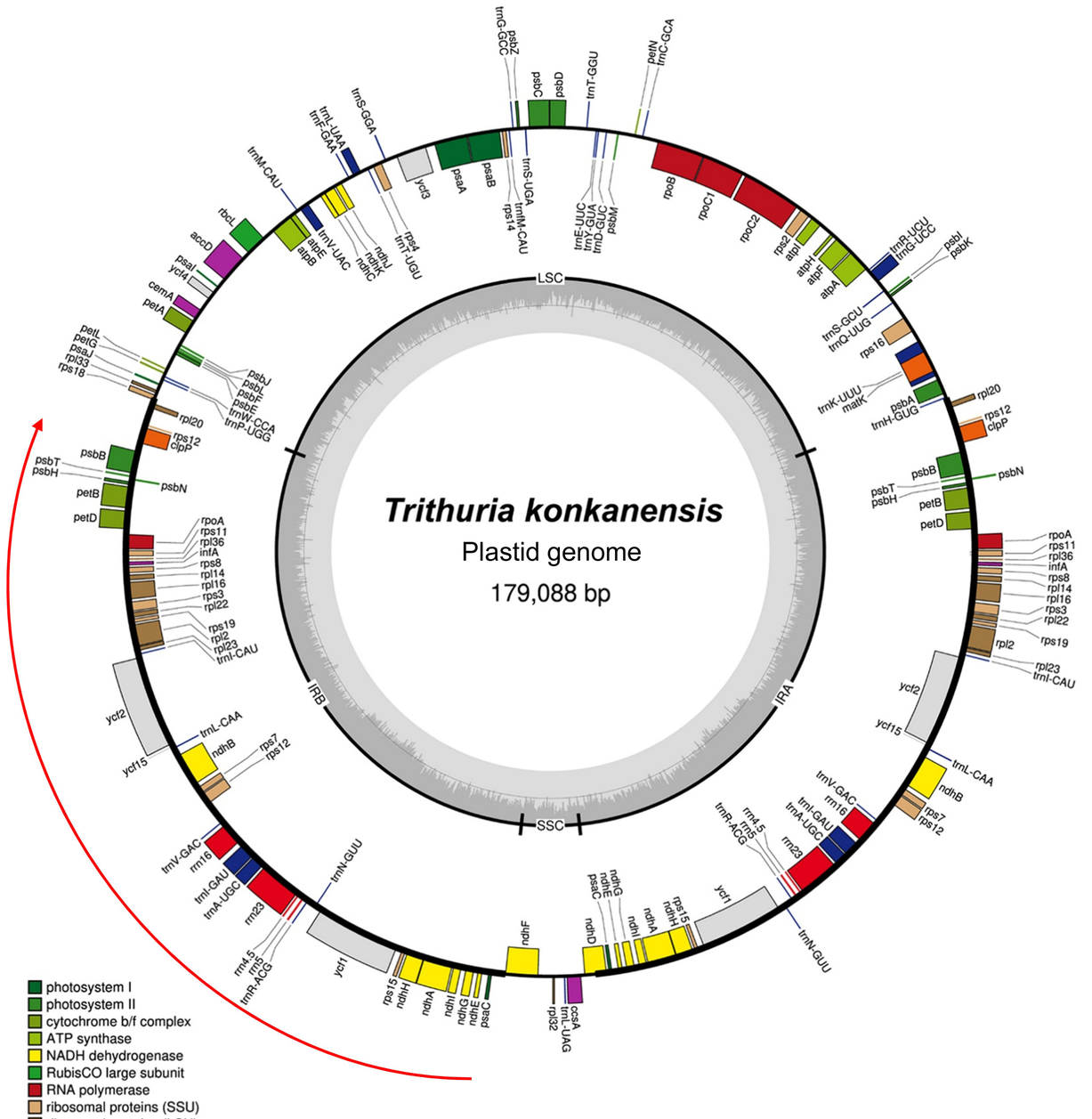
Table S4. Nucleotide diversity statistics for shared plastid protein-coding genes among the four *Trithuria* species. For each gene, the table shows gene length (bp), number of segregating sites (S), number of haplotypes (h), haplotype diversity (Hd), nucleotide diversity (π), Jukes–Cantor-corrected nucleotide diversity (πJC), Watterson's estimator (θW), and the average number of nucleotide differences (k).

Fig. S1. Alignment of the *ycf15* coding sequence (CDS) of *Amborella trichopoda* (AJ506156) with those of *Trithuria* species. *A. trichopoda* possesses an intact *ycf15* CDS, whereas all four *Trithuria* species exhibit a large internal deletion and a disrupted C-terminal coding region with a premature stop codon. The substantial loss of the coding region and disruption of the open reading frame suggest that *ycf15* is likely pseudogenized in *Trithuria*.

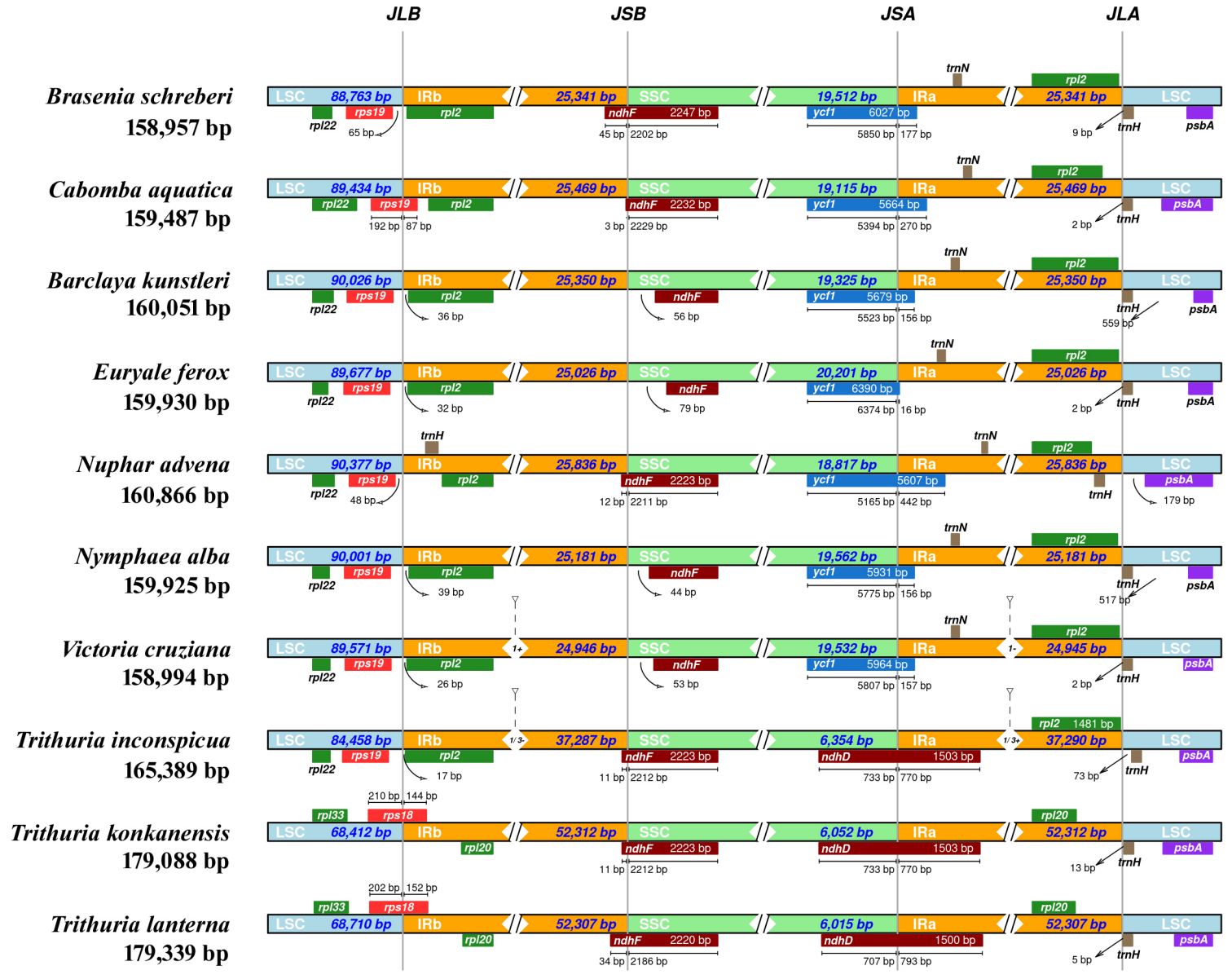
Fig. S2. Variation in plastome genome size across 83 taxa arranged in ascending order. Bars represent genome size (bp) for each taxon. *Trithuria konkanensis* is highlighted in vermilion, while all remaining taxa are shown in grey. Genome size values are indicated for all *Trithuria* taxa and for the three smallest genomes. The dashed blue line denotes the mean genome size across all sampled taxa.

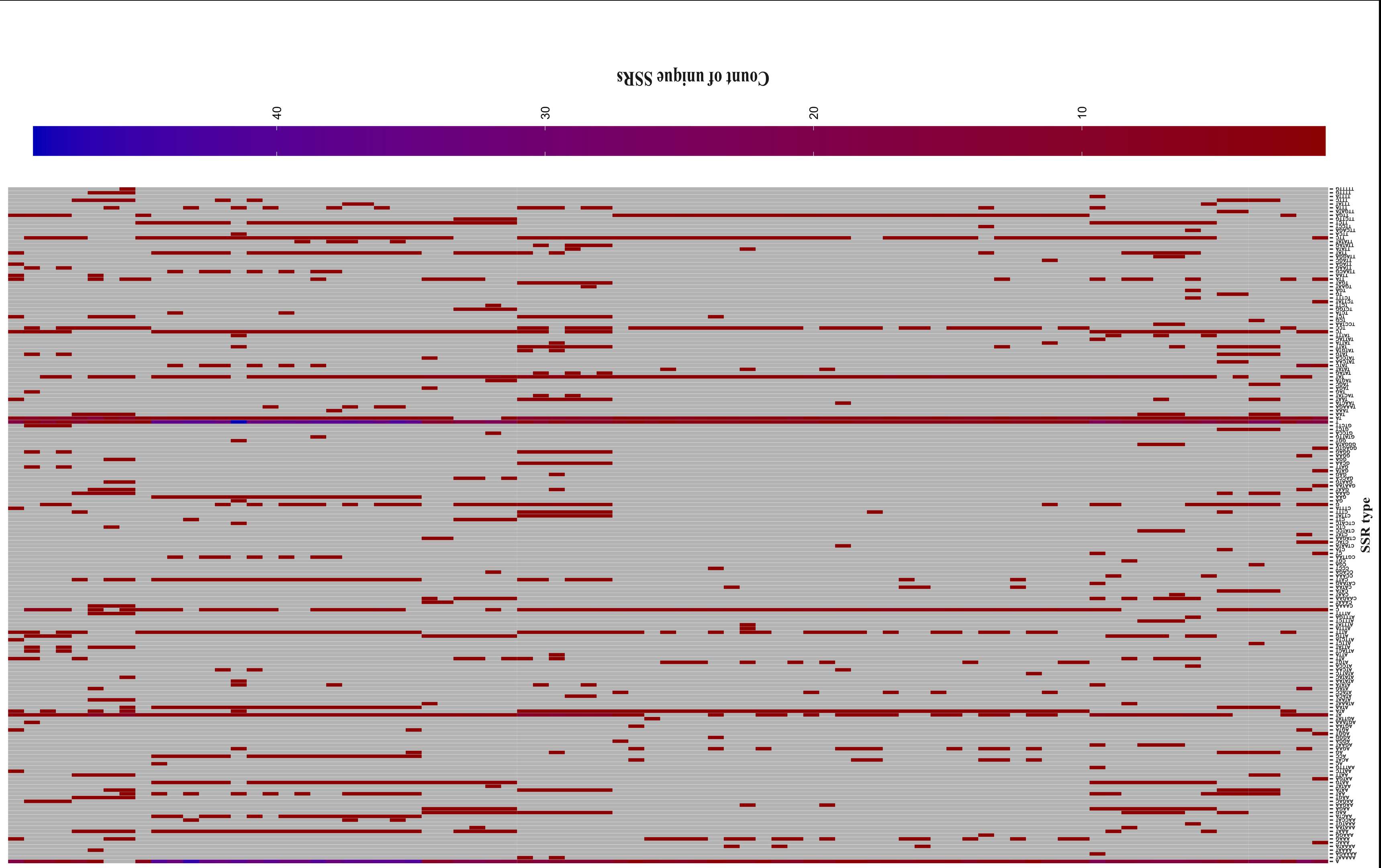
Fig. S3. Tandem and inverted repeat content across plastomes of ANA grade angiosperms, *Cycas revoluta* and *Ginkgo biloba*.

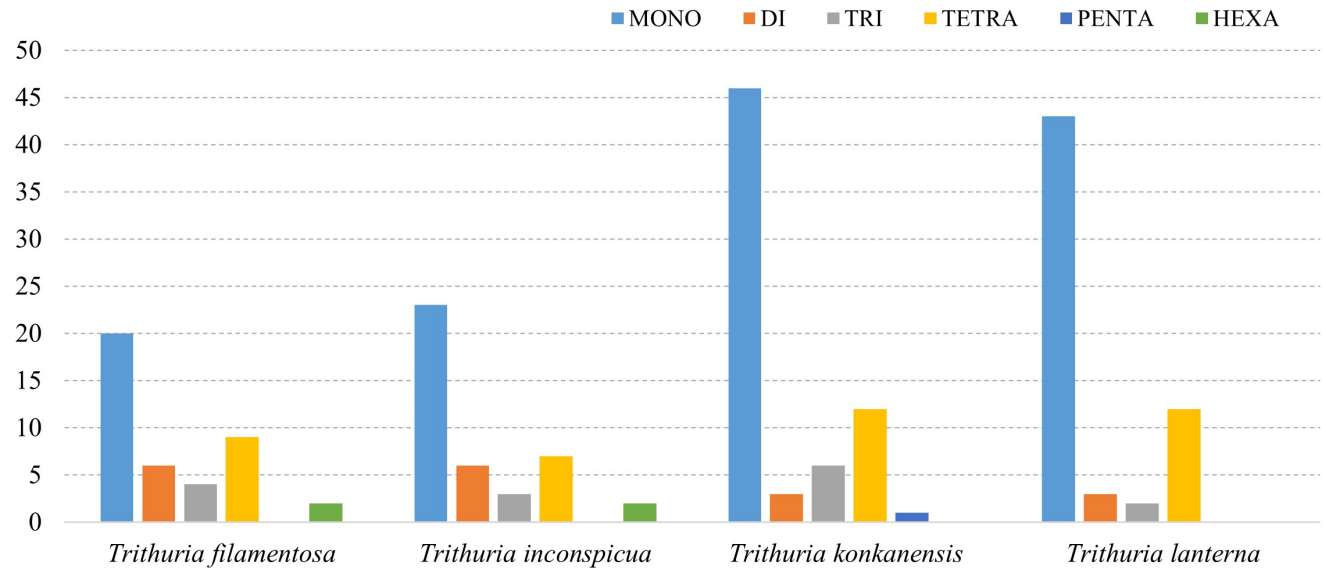

Trithuria inconspicua
Plastid genome
165,389 bp
LSC
IRB
IRA
SSC
photosystem I
photosystem II
cytochrome b/f complex
ATP synthase
NADH dehydrogenase
RubisCO large subunit
RNA polymerase
ribosomal proteins (SSU)
ribosomal proteins (LSU)
transfer RNAs
ribosomal RNAs
clpP, matK
other genes
hypothetical chloroplast reading frames (ycf)
ORFs
origin of replication
polycistronic transcripts
introns

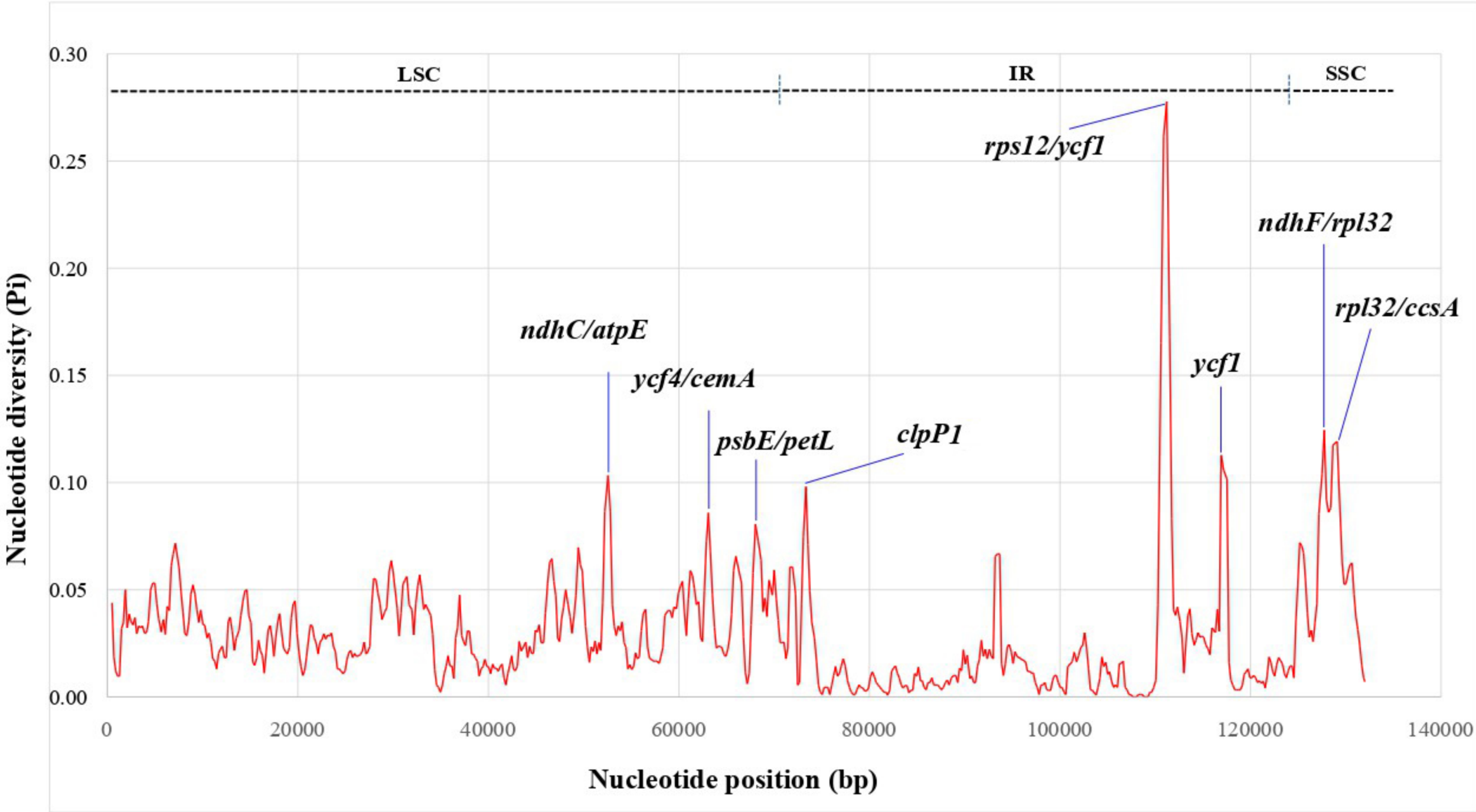

Trithuria konkanensis
Plastid genome
179,088 bp
LSC
IRB
IRA
SSC
photosystem I
photosystem II
cytochrome b/f complex
ATP synthase
NADH dehydrogenase
RubisCO large subunit
RNA polymerase
ribosomal proteins (SSU)
ribosomal proteins (LSU)
transfer RNAs
ribosomal RNAs
clpP, matK
other genes
hypothetical chloroplast reading frames (ycf)
ORFs
origin of replication
polycistronic transcripts
introns

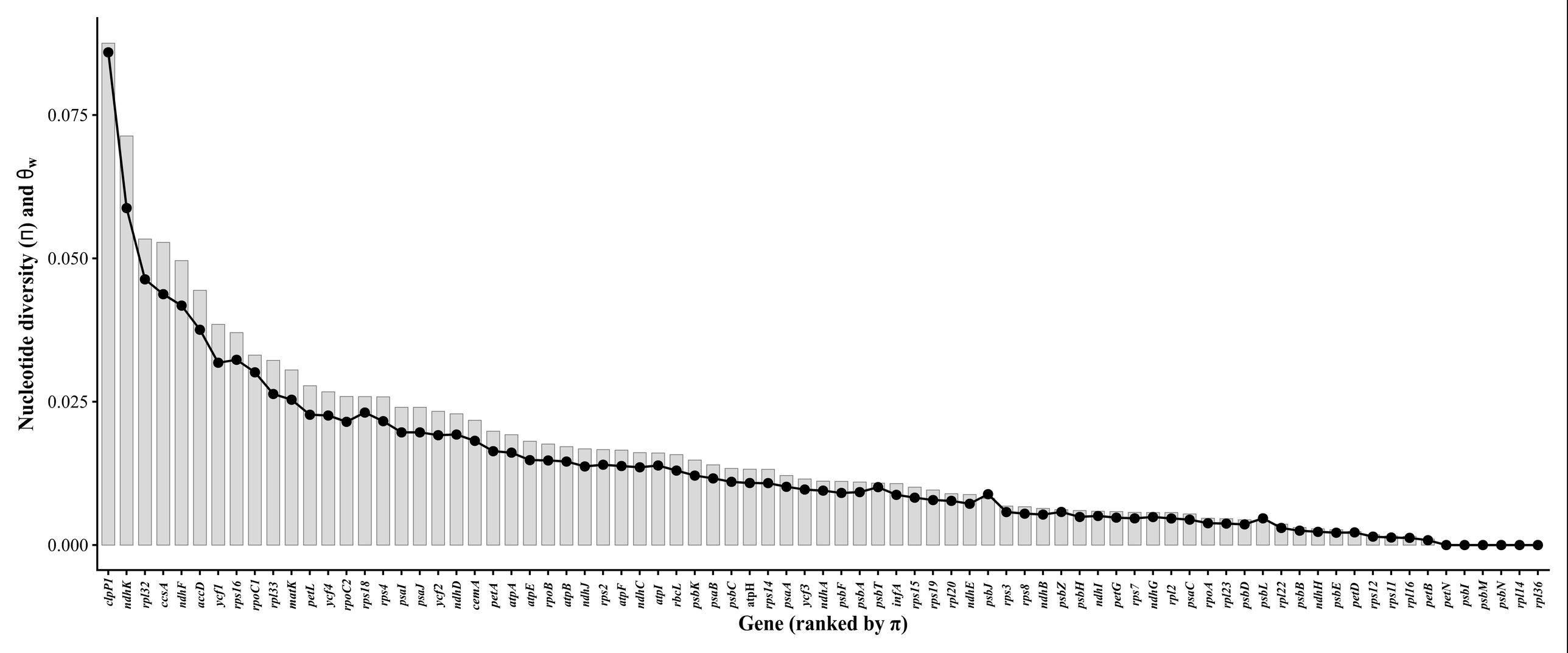
Total
LSC
IR
SSC
p = 0.00082
p = 0.000959
p = 0.000818
p = 0.00082
Plastome length (bp)
200,000
190,000
180,000
170,000
160,000
150,000
140,000
120,000
110,000
100,000
90,000
80,000
70,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
40,000
35,000
30,000
25,000
20,000
15,000
10,000
5,000
ANA Grade
Hydatellaceae
Pinophytina
ANA Grade
Hydatellaceae
Pinophytina
ANA Grade
Hydatellaceae
Pinophytina
ANA Grade
Hydatellaceae
Pinophytina
ANA Grade
Hydatellaceae
Pinophytina


JLB
JSB
JSA
JLA
Brasenia schreberi
158,957 bp
Cabomba aquatica
159,487 bp
Barclaya kunstleri
160,051 bp
Euryale ferox
159,930 bp
Nuphar advena
160,866 bp
Nymphaea alba
159,925 bp
Victoria cruziana
158,994 bp
Trithuria inconspicua
165,389 bp
Trithuria konkanensis
179,088 bp
Trithuria lanterna
179,339 bp
LSC
IRb
SSC
IRa
trnN
rpl2
rps19
rpl22
ndhF
ycf1
ndhD
trnH
psbA
rpl33
rps18
rpl20


Count of unique SSRs
40
30
20
10
Species
SSR type
Amborella trichopoda
Barclaya kunstleri
Barclaya longifolia
Barclaya rotundifolia
Brasenia schreberi
Cabomba aquatica
Cabomba caroliniana
Cabomba furcata
Euryale ferox
Illicium anisatum
Illicium brevistylum
Illicium burmanicum
Illicium difengpi
Illicium dunnianum
Illicium floridanum
Illicium henryi
Illicium jiadifengpi
Illicium lanceolatum
Illicium majus
Illicium micranthum
Illicium oligandrum
Illicium simonsii
Illicium ternstroemioides
Illicium tsaii
Illicium tsangii
Illicium verum
Kadsura ananosma
Kadsura coccinea
Kadsura heteroclita
Kadsura interior
Kadsura longipedunculata
Kadsura oblongifolia
Nuphar advena
Nuphar japonica
Nuphar longifolia
Nuphar lutea
Nuphar pumila
Nuphar shimadae
Nymphaea alba subsp. alba
Nymphaea amazonum
Nymphaea ampla
Nymphaea atrans
Nymphaea capensis
Nymphaea colorata
Nymphaea conardii
Nymphaea dimorpha
Nymphaea gigantea
Nymphaea glandulifera
Nymphaea gracilis
Nymphaea heudelotii
Nymphaea hybrid cv.
Nymphaea immatabilis
Nymphaea jamesoniana
Nymphaea lasiophylla
Nymphaea lotus
Nymphaea mexicana
Nymphaea micrantha
Nymphaea nouchali
Nymphaea odorata
Nymphaea potamophila
Nymphaea pubescens
Nymphaea rubra
Nymphaea rudgeana
Nymphaea sp.
Nymphaea tenerinervia
Nymphaea tetragona
Nymphaea thermarum
Nymphaea × daubenyana
Schisandra chinensis
Schisandra henryi
Schisandra macrocarpa
Schisandra plena
Schisandra propinqua
Schisandra propinqua subsp. sinensis
Schisandra repanda
Schisandra sphenanthera
Trithuria filamentosa
Trithuria inconspicua
Trithuria konkanensis
Trithuria lanterna
Victoria cruziana
Cycas revoluta
Ginkgo biloba


MONO
DI
TRI
TETRA
PENTA
HEXA
50
45
40
35
30
25
20
15
10
5
0
Trithuria filamentosa
Trithuria inconspicua
Trithuria konkanensis
Trithuria lanterna


LSC
IR
SSC
rps12/ycf1
ndhF/rpl32
rpl32/ccsA
ycf1
ndhC/atpE
ycf4/cemA
psbE/petL
clpP1
Nucleotide diversity (Pi)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
20000
40000
60000
80000
100000
120000
140000
Nucleotide position (bp)


Nucleotide diversity (Π) and θw
0.075
0.050
0.025
0.000
clpP1
ndhK
rpl32
ccsA
ndhF
accD
ycf1
rps16
rpoC1
rpl33
matK
petL
ycf4
rpoC2
rps18
rps4
psaI
psaJ
ycf2
ndhD
cemA
petA
atpA
atpE
rpoB
atpB
ndhJ
rps2
atpF
ndhC
atpI
rbcL
psbK
psaB
psbC
atpH
rps14
psaA
ycf3
ndhA
psbF
psbA
psbT
infA
rps15
rps19
rpl20
ndhE
psbJ
rps3
rps8
ndhB
psbZ
psbH
ndhI
petG
rps7
ndhG
rpl2
psaC
rpoA
rpl23
psbD
psbL
rpl22
psbB
ndhH
psbE
petD
rps12
rps11
rpl16
petB
petN
psbI
psbM
psbN
rpl14
rpl36
Gene (ranked by π)

Nucleotide diversity (π)

0.075

0.050

0.025

0.000

clpP1

ndhK

rpl32

ccsA

ndhF

accD

ycf1

rps16

0.00

0.25

0.50

0.75

1.00

Haplotype diversity (Hd)


Nymphaea micrantha (MT107635)
Nymphaea sp. (MW057724)
Nymphaea dimorpha (MW057738)
Nymphaea nouchali (MT991061)
Nymphaea ampla (KU189255)
Nymphaea gracilis (MW057734)
Nymphaea heudelotii (MW057733)
Nymphaea thermarum (MW057722)
Nymphaea capensis (MK111409)
Nymphaea daubenyana (MW057739)
Nymphaea pubescens (OQ161187)
Nymphaea colorata (MT107631)
Nymphaea atrans (MZ325841)
Nymphaea immutabilis (MW057732)
Nymphaea gigantea (MT107637)
Nymphaea conardii (MW057737)
Nymphaea glandulifera (MW057736)
Nymphaea jamesoniana (KT749898)
Nymphaea potamophila (MT107633)
Nymphaea amazonum (MW057741)
Nymphaea tenerinervia (MW057723)
Nymphaea lasiophylla (MW057731)
Nymphaea rudgeana (MW057725)
Nymphaea lotus (MW057729)
Nymphaea rubra (MT107632)
Euryale ferox (KY392765)
Victoria cruziana (KY001813)
Nymphaea hybrid cv. (MW644616)
Nymphaea mexicana (MW057728)
Nymphaea odorata (MW057726)
Nymphaea alba subsp. alba (MK040444)
Nymphaea tetragona (MT107634)
Barclaya kunstleri (KY392762)
Barclaya rotundifolia (MW057721)
Barclaya longifolia (KY284156)
Nuphar japonica (MH161174)
Nuphar shimdae (MH050797)
Nuphar pumila (MH161176)
Nuphar lutea (MH161175)
Nuphar advena (DQ354691)
Nuphar longifolia (MH050795)
Nymphaeaceae
Cabomba caroliniana (KT705317)
Cabomba furcata (MG967470)
Cabomba aquatica (MG720559)
Brasenia schreberi (NC_031343)
Cabombaceae
Trithuria filamentosa (KF696682)
Trithuria inconspicua (HE963749)
Trithuria konkanensis
Trithuria lanterna (MN365661)
Hydatellaceae
Illicium jiadifengpi (OL802930)
Illicium tsaii (MZ315036)
Illicium simonsii (OL802933)
Illicium tsangii (OL802934)
Illicium anisatum (OP032107)
Illicium ternstroemioides (MZ315039)
Illicium majus (OL802932)
Illicium burmanicum (MZ315037)
Illicium verum (KY085896)
Illicium brevistylum (OL802927)
Illicium lanceolatum (MZ315040)
Illicium dunnianum (OL802929)
Illicium henryi (OL879874)
Illicium difengpi (MW039601)
Illicium micranthum (MZ315035)
Illicium floridanum (KY085892)
Illicium oligandrum (NC_009600)
Kadsura heteroclita (MN823698)
Kadsura interior (MN698966)
Kadsura longipedunculata (MH535482)
Kadsura oblongifolia (PX548869)
Schisandra propinqua (PQ037862)
Schisandra propinqua subsp. sinensis (MZ846205)
Schisandra plena (PQ282136)
Schisandra macrocarpa (PP706075)
Kadsura ananosma (MN823697)
Kadsura coccinea (MN480469)
Schinsandra sphenanthera (MK193856)
Schisandra henryi (PV747847)
Schisandra chinensis (KU362793)
Schisandra repanda (OK377339)
Schisandraceae
Amborella trichopoda (AJ506156)
Amborellaceae
Cycas revoluta (NC_020319)
Cycadaceae
Ginkgo biloba (MN443423)
Ginkgoaceae
0.04


Nymphaea potamophila (MT107633)
Nymphaea jamesoniana (KT749898)
Nymphaea glandulifera (MW057736)
Nymphaea conardii (MW057737)
Nymphaea tenerinervia (MW057723)
Nymphaea amazonum (MW057741)
Nymphaea rudgeana (MW057725)
Nymphaea lasiophylla (MW057731)
Nymphaea rubra (MT107632)
Nymphaea lotus (MW057729)
Victoria cruziana (KY001813)
Euryale ferox (KY392765)
Nymphaea mexicana (MW057728)
Nymphaea hybrid cv. (MW644616)
Nymphaea odorata (MW057726)
Nymphaea tetragona (MT107634)
Nymphaea alba subsp. alba (MK040444)
Nymphaea thermarum (MW057722)
Nymphaea heudelotii (MW057733)
Nymphaea pubescens (OQ161187)
Nymphaea daubenyana (MW057739)
Nymphaea capensis (MK111409)
Nymphaea colorata (MT107631)
Nymphaea sp. (MW057724)
Nymphaea micrantha (MT107635)
Nymphaea dimorpha (MW057738)
Nymphaea nouchali (MT991061)
Nymphaea gracilis (MW057734)
Nymphaea ampla (KU189255)
Nymphaea immutabilis (MW057732)
Nymphaea atrans (MZ325841)
Nymphaea gigantea (MT107637)
Barclaya rotundifolia (MW057721)
Barclaya kunstleri (KY392762)
Barclaya longifolia (KY284156)
Nuphar pumila (MH161176)
Nuphar japonica (MH161174)
Nuphar shimdae (MH050797)
Nuphar lutea (MH161175)
Nuphar longifolia (MH050795)
Nuphar advena (DQ354691)
Cabomba furcata (MG967470)
Cabomba caroliniana (KT705317)
Cabomba aquatica (MG720559)
Brasenia schreberi (NC_031343)
Trithuria lanterna (MN365661)
Trithuria konkanensis
Trithuria inconspicua (HE963749)
Trithuria filamentosa (KF696682)
Illicium tsangii (OL802934)
Illicium simonsii (OL802933)
Illicium tsaii (MZ315036)
Illicium jiadifengpi (OL802930)
Illicium ternstroemioides (MZ315039)
Illicium anisatum (OP032107)
Illicium majus (OL802932)
Illicium verum (KY085896)
Illicium burmanicum (MZ315037)
Illicium lanceolatum (MZ315040)
Illicium brevistylum (OL802927)
Illicium henryi (OL879874)
Illicium dunnianu (OL802929)
Illicium difengpi (MW039601)
Illicium micranthum (MZ315035)
Illicium floridanum (KY085892)
Illicium oligandrum (NC_009600)
Schisandra propinqua subsp. sinensis (MZ846205)
Schisandra propinqua (PQ037862)
Schisandra plena (PQ282136)
Schisandra macrocarpa (PP706075)
Kadsura oblongifolia (PX548869)
Kadsura longipedunculata (MH535482)
Kadsura interior (MN698966)
Kadsura heteroclita (MN823698)
Kadsura coccinea (MN480469)
kadsura ananosma (MN823697)
Schisandra henryi (PV747847)
Schinsandra sphenanthera (MK193856)
Schisandra chinensis (KU362793)
Schisandra repanda (OK377339)
Amborella trichopoda (AJ506156)
Ginkgo biloba (MN443423)
Cycas revoluta (NC_020319)
Nymphaeaceae
Cabombaceae
Hydatellaceae
Schisandraceae
Amborellaceae
Ginkgoaceae
Cycadaceae
0.05

Nymphaea ampla (AY771812)
Nymphaea ampla (FJ026604)
Nymphaea gracilis (FM242151)
Nymphaea elegans (AY771811)
Nymphaea elegans (FJ026601)
Nymphaea heudelotii (FJ026603)
Nymphaea micrantha (EU428065)
Nymphaea hybrid cv. (PV423496)
Nymphaea capensis (AY620421)
Nymphaea nouchali (FJ198404)
Nymphaea atrans (FJ026554)
Nymphaea georginae (FJ026563)
Nymphaea macrosperma (FJ026577)
Nymphaea carpentariae (FJ026556)
Nymphaea elleniae (FJ026560)
Nymphaea hastifolia (FJ026568)
Nymphaea violacea (FJ026582)
Nymphaea lotus (OM677814)
Nymphaea petersiana (MW798799)
Nymphaea zenkeri (MK452743)
Nymphaea lotus f. *thermalis* (FM242153)
Nymphaea petersiana (FM242156)
Nymphaea rubra (GU199454)
Nymphaea pubescens cv. *alba* (AY707894)
Nymphaea nouchali (KF170027)
Nymphaea rubra (AY707902)
Nymphaea rudgeana (EU428068)
Nymphaea oxypetala (FM242150)
Nymphaea amazonum (FM242149)
Nymphaea jamesoniana (FM242152)
Nymphaea alba (EU428030)
Nymphaea alba (FM242148)
Nymphaea mexicana (AY707893)
Nymphaea pygmaea (LC775191)
Nymphaea tetragona (LC775225)
Nymphaea odorata (FM242155)
Nymphaea candida (AY707900)
Nymphaea odorata (AY707901)
Euryale ferox (AY620422)
Victoria amazonica (AY858642)
Nuphar sinensis (AF136288)
Nuphar pumila (AY620428)
Nuphar lutea (EU428043)
Nuphar spenneriana (KY972342)
Nuphar pumila subsp. *sinensis* (KY972345)
Nuphar rubrodisca (AH006157)
Nuphar ulvacea (AH006160)
Nuphar orbiculata (AH006163)
Nuphar microphylla (AH006167)
Nuphar ozarkana (AH006159)
Nuphar polysepala (MF964142)
Nuphar variegata (KY972350)
Barclaya wellyi (OQ171038)
Barclaya kunstleri (MH535476)
Barclaya rotundifolia (MH535481)
Barclaya hirta (OQ171039)
Barclaya motleyi (MH535477)
Barclaya panchorensis (MH535479)
Barclaya rugosa (MH535480)
Barclaya longifolia (MH535471)
Cabomba caroliniana (AY620424)
Cabomba furcata (AY620425)
Brasenia schreberi (AY620426)
Trithuria austinensis (JQ284161)
Trithuria inconspicua (JQ284177)
Trithuria filamentosa (JQ284173)
Trithuria australis (JQ284171)
Trithuria cowieana (JQ284154)
Trithuria lanterna (JQ284158)
Trithuria polybracteata (JQ284155)
Trithuria konkanensis (JQ284156)
Trithuria bibracteata (JQ284178)
Trithuria submersa (JQ284184)
Schisandra viridis (AF263438)
Schisandra arisanensis (KP689642)
Schisandra sphenanthera (AF263437)
Schisandra sphaerandra (KY523607)
Schisandra elongata (KP689650)
Schisandra glaucescens (KP689649)
Schisandra glaucescens (AF263439)
Schisandra henryi (AF263435)
Schisandra pubescens (AF263436)
Schisandra rubriflora (AF263440)
Schisandra grandiflora (AF163707)
Schisandra chinensis (AF263441)
Schisandra glabra (AF163720)
Schisandra bicolor (DQ342255)
Schisandra repanda (KX815928)
Schisandra bicolor var. *tuberculata* (AF263442)
Kadsura ananosma (AF263447)
Kadsura coccinea (AF263445)
Schisandra propinqua (AF163717)
Schisandra propinqua var. *sinensis* (AF263444)
Schisandra plena (AF263443)
Schisandra macrocarpa (PV082525)
Kadsura japonica (AF163712)
Kadsura heteroclita (AF263446)
Kadsura longipedunculata (AF263448)
Kadsura angustifolia (AF163715)
Kadsura oblongifolia (AF163714)
Kadsura scandens (AF163719)
Illicium verum (AF163724)
Illicium arborescens (AF163723)
Illicium dunnianum (AF163731)
Illicium fargesii (AF263449)
Illicium anisatum (AF163729)
Illicium angustisepalum (AF163721)
Illicium difengpi (AF163728)
Illicium micranthum (AF163727)
Illicium majus (AF163733)
Illicium lanceolatum (AF163726)
Illicium henryi (AF163725)
Illicium floridanum (AF163734)
Illicium parviflorum (AF163735)
Amborella trichopoda (LC669891)
Gnetum ula (AY449609)

Nymphaeaceae
Cabombaceae
Hydatellaceae
Schisandraceae
Amborellaceae
Gnetaceae

0.2

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Fig.S1.pptx
- Table1.docx
- Fig.S3.jpg
- Fig.S2.png
- Table2.docx
- TableS1.docx
- TableS2.xlsx
- TableS4.xlsx
- TableS3.xlsx