



OPEN ACCESS

EDITED BY

Saraj Bahadur,
Hainan University, China

REVIEWED BY

Huasheng Peng,
China Academy of Chinese Medical Sciences,
China
Zhixi Fu,
Ministry of Education, China
Xiaoyun Wang,
Jiangxi University of Traditional Chinese
Medicine, China

*CORRESPONDENCE

Yi Tong

✉ tongyi@gzucm.edu.cn

RECEIVED 01 February 2025

ACCEPTED 31 March 2025

PUBLISHED 01 May 2025

CITATION

Lin Z, Fan Z, Lan Y, Fu L, Xu F and Tong Y
(2025) Comparative and
phylogenetic analysis of complete
chloroplast genomes of *Phrynium* s. s.
and *Stachyphrynium* (Marantaceae)
in China, including a new species.
Front. Plant Sci. 16:1569683.
doi: 10.3389/fpls.2025.1569683

COPYRIGHT

© 2025 Lin, Fan, Lan, Fu, Xu and Tong. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Comparative and phylogenetic analysis of complete chloroplast genomes of *Phrynium* s. s. and *Stachyphrynium* (Marantaceae) in China, including a new species

Zhiyi Lin¹, Zhichao Fan¹, Yaoqing Lan¹, Lin Fu², Feilong Xu³
and Yi Tong^{1*}

¹School of Pharmaceutical Science, Guangzhou University of Chinese Medicine, Guangzhou, China,

²Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China, ³Information Department, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China

Plants of the genera *Phrynium* and *Stachyphrynium* traditionally used as ethnomedicine or for wrapping sticky rice dumpling in the tropical and south subtropical Asia, have a long history of ethnobotanical use. China represents the northernmost distribution of Marantaceae in Asia. Due to the notably similar leaf morphology between these genera, herbarium specimens are frequently misidentified, especially during the vegetative stages. Their morphological uniformity and unclear interspecific genetic relationships pose significant challenges to taxonomic classification and species identification. To date, systematic taxonomic revisions and phylogeny of their indigenous species remain lacking. In this study, we conducted comparative chloroplast genomes analyses of seven *Phrynium* and two *Stachyphrynium* species in China. The chloroplast genomes exhibited conserved structure, gene content, gene order and codon usage bias, but diverged in genomes size and the SC/IR boundaries. Four variable regions were identified as potential molecular markers for species identification. Phylogenetic analyses using CDS and nrDNA strongly support *Phrynium* and *Stachyphrynium* in China as two distinct monophyletic groups, with *Phrynium* subdivided into two clades. These findings advance our understanding of their molecular relationships and provide critical insights for identification, utilization, and conservation as medicinal plants. Finally, we describe and illustrate the new species *Phrynium pyramidale* Y. Tong & Z. Y. Lin.

KEYWORDS

Phrynium, *Stachyphrynium*, chloroplast genome, nuclear ribosomal DNA, phylogenomic analysis, *Phrynium pyramidale*

1 Introduction

Phrynium Willd. and *Stachyphrynium* K. Schum. are members of the family Marantaceae (order Zingiberales), classified within the Commelinids clade of the monocotyledonous flowering plants according to APG IV system (The Angiosperm Phylogeny Group, 2016; Cole et al., 2019). Currently, *Phrynium* comprises approximately 20 species distributed across the tropical and some subtropical regions of Asia and Africa. In China, six species have been documented (Wu and Helen, 2000; Fang, 2002; Fu et al., 2017), with an additional new species described in this study. Both genera are widely used in traditional Chinese medicine. The herbs and roots are employed for their purported therapeutic effects in clearing heat, detoxifying, cooling the blood, and stopping bleeding, aiding in treating conditions such as hoarseness, sore throat and oral ulcer (Guangdong Institute of Botany, 1977; Wu, 1991; State Administration of Traditional Chinese Medicine, 1999; Guangxi Institute of Botany, 2016). Additionally, these plants are cultivated as ornamental plants in garden ditches and shaded areas (Cen et al., 2017). The leaves of *Phrynium* are traditionally used by Chinese ethnic groups to wrap sticky rice dumplings, owing to their natural preservative properties (Cen et al., 2017).

The genus *Phrynium* is characterized by several distinctive morphological features, including congested and capitate inflorescences, spirally arranged bracts, sepals longer than the corolla tubes, 3-locular ovary and dehiscent fruits (Wu and Helen, 2000; Clausager and Borchsenius, 2003; Suksthan et al., 2010). Seven *Phrynium* species are recorded in China, whereas only two *Stachyphrynium* species occur there. Morphologically, *Stachyphrynium* differs from *Phrynium* by its solitary spikes, two-ranked bracts and corolla tubes longer than the sepals (Wu and Helen, 2000; Clausager and Borchsenius, 2003). Since Willdenow (1797) established *Phrynium*, taxonomic studies have focused on clarifying generic boundaries within Marantaceae and intrageneric classification. Multiple analyses indicate that *Phrynium* is non-monophyletic, forming a clade with *Phacelophrynium*, *Monophrynium*, and *Cominsia* (Andersson, 2001; Prince and Kress, 2006; Suksathan et al., 2009; Ley and Claßen-Bockhoff, 2011). Consequently, Suksathan et al. (2009) merged these genera into an expanded *Phrynium* during their taxonomic revision of Asian Marantaceae. However, Asian Marantaceae remain understudied compared to South American counterparts. Existing works (e.g., Roxburgh, 1820; Ridley, 1899; Schumann, 1902; Gagnepain, 1932; Holttum, 1951; Clausager and Borchsenius, 2003) offer limited insights into species delimitation and distribution. No modern systematic revisions exist for these genera. Given the absence of *Phacelophrynium*, *Monophrynium*, and *Cominsia* in China, this study focuses on *Phrynium* sensu stricto (s.s.). Current classifications rely on floral and bract morphology (Clausager and Borchsenius, 2003; Suksthan et al., 2010; Niissalo et al., 2016). Leaf morphology is highly uniform, complicating pre-flowering identification. *Stachyphrynium placentarium* was historically placed in *Phrynium* due to its densely capitate inflorescence (Merrill, 1919; Clausager and Borchsenius, 2003). Notably, illustrations of *Phrynium rheedei* and *Stachyphrynium placentarium* in Chinese Materia Medica are mislabeled and should be interchanged (State

Administration of Traditional Chinese Medicine, 1999). Herbarium specimens of *Phrynium* are frequently misidentified, underscoring the urgent need for a comprehensive taxonomic revision of *Phrynium* and *Stachyphrynium* in China. We recommend integrating phylogenomic approaches to resolve interspecific relationships and develop reliable molecular barcodes.

Phylogenomics, particularly through complete chloroplast genomes and nuclear genomes analysis, has proven effective in addressing phylogenetic challenges in taxonomically complex plant groups (Jansen et al., 2007; Moore et al., 2007, 2010). The chloroplast, a vital photosynthetic organelle, contains an independent genome typically structured as a double-stranded circular DNA molecule with a highly conserved quadripartite structure: a large single-copy region (LSC), a small single-copy region (SSC), and two inverted repeat regions (IRs). The contraction and expansion of the IR region contribute to the diversity of chloroplast genomes among different species (Aii et al., 1997; Plunkett and Downie, 2000; Wang et al., 2008). Most angiosperm chloroplast genomes range from 120 to 160 kb in length and exhibit low molecular weight, high copy number, structural conservation, maternal inheritance, and low recombination rate (Palmer et al., 1988; Green, 2011; Jansen and Ruhlman, 2012). Consequently, they are widely used in species identification, phylogenetics, and molecular evolution studies (Smith et al., 1993; Andersson, 2001; Prince and Kress, 2006; Suksathan et al., 2009). However, chloroplast genomes have limitations in resolving rapid radiation events and detecting hybridization or polyploidy (McKain et al., 2018). In contrast, nuclear genome possess key features such as biparental inheritance, abundant informative loci, and higher mutation rates (Kellogg and Bennetzen, 2004), enhancing their resolution for lower taxonomic ranks. Due to their large size and structural complexity, nuclear studies often focus on repetitive sequences (e.g., ribosomal DNA [rDNA]) or multigene families (Richard et al., 2008). Currently, rDNA is as pivotal as chloroplast genomes in plant systematics and taxonomy (Kress et al., 2001, 2002; Qiu et al., 2010; Cron et al., 2012).

Chloroplast genomes and rDNA have been widely utilized for constructing phylogenetic relationships and species identification in large, taxonomically challenging plant groups. Fu et al. (2024) demonstrated that both plastomes and nuclear rDNA (nrDNA) could significantly improve species resolution in *Acer*. Similarly, Chao et al. (2023) employed chloroplast genomes and 45S nrDNA to conduct phylogenetic analyses of *Bupleurum* in Sichuan Province, southwest China, providing strong evidence for the monophyly of each species.

In the family Marantaceae, Andersson (2001) inferred the phylogenetic relationships among 22 genera through nucleotide sequence variations of the *rps16* intron and morphological characteristics. Prince and Kress (2006) conducted phylogenetic analyses using *matK*, the 3' intergenic spacer region, and the *trnL-F* intergenic spacer region. These analyses identified five major clades. Suksathan et al. (2009) explored the phylogenetic relationships of Asian Marantaceae using the *rps16* intron, ITS1, and 5S-NTS, and determined that *Phrynium* is a non-monophyletic genus. However, the relationship between *Phrynium* and *Stachyphrynium* remains

unclear, and most *Phrynium* species in China are lacking in these studies. Previous phylogenetic studies primarily focused on chloroplast DNA fragments and ITS regions rather than complete chloroplast genomes to address phylogenetic challenges within Marantaceae. In the context of *Phrynium* and *Stachyphrynium*, only the complete chloroplast genomes sequence of *P. rheedei* (a synonym of *P. pubinerve*) has been deposited in GenBank, without any accompanying publication or analysis. At present, there are several ongoing challenges, including an ambiguous understanding of the chloroplast genomes structure and characteristics of the Asian Marantaceae, a limited comprehension of interspecific relationships within *Phrynium* in China, and deficiency in effective molecular identification methods for both genera.

In this study, we newly sequenced the chloroplast genomes of seven *Phrynium* species and two *Stachyphrynium* species, covering nearly all representatives of these genera in China. Phylogenetic analysis of these taxa were conducted based on complete chloroplast genomes and ribosomal DNA (rDNA). This study aimed to: 1) characterize and compare structural variations in the chloroplast genomes of the newly sequenced species; 2) identify highly variable DNA barcodes for species delimitation within *Phrynium* and *Stachyphrynium*; 3) clarify the phylogenetic positions of *Phrynium* and *Stachyphrynium*, and resolve interspecific relationships within Chinese *Phrynium*.

Additionally, during a field investigation in Menghai County, Yunnan Province, China, a specimen unequivocally belonging to the genus *Phrynium* was collected. Its bracts are arranged in a pagoda-like formation, distinguishing it from other known Asian species. Through comprehensive field observations, literature review, specimen comparison and phylogenetic analysis, we finally identify it as a new species, formally described herein.

2 Materials and methods

2.1 Taxon sampling and morphological inspection

A total of 40 plant samples were collected, representing seven species of *Phrynium* (*P. pyramidale*, *P. pubinerve*, *P. imbricatum*, *P. hainanense*, *P. tonkinense*, *P. pedunculiferum* and *P. yunnanense*), two species of *Stachyphrynium* (*S. spicatum* and *S. placentarium*), and *Marantochloa purpurea*. *Marantochloa purpurea* was selected as an outgroup from a closely related tribe (Prince and Kress, 2006; Suksathan et al., 2009). These samples were collected from diverse regions throughout China, specifically including Yunnan Province, Hainan Province, Guangxi Zhuang Autonomous Region, and Xizang Autonomous Region. The voucher specimens are deposited in the herbarium of Guangzhou University of Chinese Medicine (GUCM), with detailed information provided in Table 1.

Key morphological characteristics of taxonomic significance in *Phrynium* were documented through photography, measurements, and the recording of variation ranges. This analysis focused on sheath, leaf shape and size, bract shape, texture and arrangement, as well as flower size and color, and the shape and size of sepals, with

particular emphasis on comparing *P. pyramidale* to the two related species.

2.2 DNA extraction and genome sequencing

Genomic DNA was extracted from 50 mg of silica gel-dried leaves using a modified CTAB method (Doyle and Doyle, 1987). The purity and integrity of the isolated DNA were evaluated using the B-500 ultra-micro spectrophotometer (Metash, China) and agarose gel electrophoresis. High-quality genomic DNA was randomly fragmented into 300-400 bp segments using ultrasound treatment for library construction. Sequencing was performed on the DNBSEQ platform (150 bp paired-end) at BGI-Wuhan, generating a total of 3.0 Gb of clean data for each sample.

2.3 Assembly and annotation

The chloroplast genomes assembly was conducted using GetOrganelle v1.7.5.0 (Jin et al., 2020), with *Phrynium rheedei* (MZ958828.1) and *Thalia dealbata* (PP059847.1) as reference genomes. The assembly of nrDNA followed the same methodology.

Preliminary annotation was executed using Plastid Genome Annotator (PGA) software (Qu et al., 2019). The results of the annotation were subsequently analyzed and refined through manual corrections in Geneious Prime R9.0.2 (Kearse et al., 2012), which included the identification of start and stop codons for each expressed gene, in comparison to the reference genome. Transfer RNAs (tRNAs) were identified using tRNAscan-SE v2.0.7 (Chan et al., 2021). The comprehensive chloroplast maps were generated with OGDRAW (Lohse et al., 2007), and the complete chloroplast sequences were submitted to GenBank (<http://www.ncbi.nlm.nih.gov/>). The annotation of nrDNA was also conducted using the same software, with comparisons made to the reference genome.

2.4 Genome structure analysis

Geneious R9.0.2 was used to determine the total length of the chloroplast genomes sequences of all species within the genera *Phrynium* and *Stachyphrynium*, as well as to assess the lengths of LSC, IR, SSC regions, along with GC content in each of these regions. The annotated chloroplast genomes sequences were subsequently uploaded to the IRscope online software (Amiryousefi et al., 2018) for a visual analysis of the contraction and expansion of the IR boundaries, resulting in the generation of simplified diagrams.

2.5 Codon usage analysis

PhyloSuite v1.2.2 (Zhang et al., 2020) was employed to extract protein-coding sequences (CDS) from chloroplast genomes, with

TABLE 1 Voucher information and GenBank accession numbers of materials.

Taxon	Locality	Collection date	Voucher	GenBank accession number	
				cp genomes	nrDNA
<i>Phrynium pyramidale</i> Y. Tong & Z. Y. Lin	Menghai, Yunnan	2023/07/08	Tong Y 230708014	PQ835434	PQ869816
	Menghai, Yunnan	2020/08/21	Tong Y 20082111	PQ835428	PQ869815
	Mengla, Yunnan	2023/07/11	Tong Y 230711003	PQ835426	PQ869814
	Menglun, Yunnan	2023/04/01	Tong Y 23040101	PQ835430	PQ869817
	Malipo, Yunnan	2022/03/19	Guo XB 22031903	PQ835429	PQ869820
	Malipo, Yunnan	2023/07/15	Tong Y 230715021	PQ835433	PQ869818
	Hekou, Yunnan	2023/07/13	Tong Y 230713002	PQ835436	PQ869819
	Pingbian, Yunnan	2023/07/14	Tong Y 230714012	PQ835431	PQ869821
	Pingbian, Yunnan	2023/07/14	Tong Y 230714003	PQ835432	PQ869822
	Motuo, Xizang	2023/12/30	Liu BY 23123001	PQ835427	PQ869823
	Vietnam	2023/12/28	Liao S 20231277	PQ835435	PQ869824
<i>P. pubinerve</i> Blume	Mengla, Yunnan	2023/07/10	Tong Y 230710003	PQ835425	PQ865401
<i>P. tonkinense</i> Gagnep.	Malipo, Yunnan	2020/08/16	Tong Y 20081604	PQ835441	PQ865465
	Malipo, Yunnan	2023/02/24	Tong Y 23022410	PQ835437	PQ865467
	Malipo, Yunnan	2022/03/21	Guo XB 22032103	PQ835438	PQ869826
	Malipo, Yunnan	2022/03/19	Guo XB 22031904	PQ835439	PQ865470
	Honghe, Yunnan	2023/07/14	Tong Y 230714002	PQ835440	PQ865471
<i>P. imbricatum</i> Roxb.	Wanning, Hainan	2023/08/20	Lin ZY 230820003	PQ835419	PQ865473
	Wanning, Hainan	2023/08/20	Lin ZY 230820014	PQ835417	PQ865474
	Baoting, Hainan	2023/08/22	Lin ZY 230822004	PQ835416	PQ869827
	Baoting, Hainan	2023/08/21	Lin ZY 230821004	PQ835418	PQ869825
<i>P. hainanense</i> T. L. Wu & S. J. Chen	Wanning, Hainan	2023/08/20	Lin ZY 230820008	PQ835415	PQ865476
	Baoting, Hainan	2023/08/22	Lin ZY 230822005	PQ835414	PQ865477
<i>P. pedunculiferum</i> D. Fang	Napo, Guangxi	2023/10/29	Tong Y 231029010	PQ835423	PQ865478
	Napo, Guangxi	2020/07/04	Zeng SJ 4736	PQ835420	PQ865481
	Menghai, Yunnan	2020/08/19	Tong Y 20081911	PQ835424	PQ865480
	Menghai, Yunnan	2023/07/08	Tong Y 230708019	PQ835421	PQ865479
	Baoshan, Yunnan	2023/07/13	Tong Y 230713004	PQ835422	PQ865483
<i>P. yunnanense</i> Y. S. Ye & L. Fu	Hekou, Yunnan	2020/08/17	Tong Y 20081708	PQ835442	PQ865482
<i>Stachyphrynium spicatum</i> (Roxb.) K. Schum.	Mengla, Yunnan	2024/09/27	Lin ZY 24092701	PQ835452	PQ865484
<i>S. placentarium</i> (Lour.) Clausager & Borchs.	Baoting, Hainan	2023/08/22	Lin ZY 230822002	PQ835445	PQ865485
	Wanning, Hainan	2023/08/20	Lin ZY 230820011	PQ835450	PQ869828
	Malipo, Yunnan	2020/08/16	Tong Y 20081603	PQ835443	PQ865486
	Hekou, Yunnan	2023/08/18	Tong Y 20081808	PQ835444	PQ865488
	Mengla, Yunnan	2023/07/09	Tong Y 230709016	PQ835449	PQ865489
	Xichou, Yunnan	2023/02/22	Tong Y 23022228	PQ835447	PQ865490
	Longzhou, Guangxi	2023/04/25	Tong Y 23042503	PQ835448	PQ865491

(Continued)

TABLE 1 Continued

Taxon	Locality	Collection date	Voucher	GenBank accession number	
				cp genomes	nrDNA
	Napo, Guangxi	2020/08/14	Tong Y 20081409	PQ835446	PQ865492
	Zhaoqing, Guangdong	2019/06/02	Tong Y 19060201	PQ835451	PQ869829
<i>Marantochloa purpurea</i> (Ridl.) Milne-Redh.	Guangzhou (cultivated), Guangdong	2023/11/12	Tong Y 23111201	PQ835413	PQ865494

sequences less than 300 bp being excluded to enhance the accuracy of the results. The relative synonymous codon usage (RSCU) of these sequences was analyzed using CodonW v.1.4.2 (<http://codonw.sourceforge.net/>), and the values of RSCU were utilized to generate the heat maps using TBtools-II v2.119 (Chen et al., 2023). A codon exhibiting an RSCU value exceeding 1 signifies a higher frequency of usage, while a value below 1 indicates less frequent usage (Sharp and Li, 1987).

2.6 Simple sequence repeat analysis

MISA-web (<http://pgrc.ipk-gatersleben.de/misa>) was employed to identify the simple sequence repeats (SSRs) within chloroplast genomes. The parameters for the minimum repeated size were established for mononucleotides to hexanucleotides at thresholds of 10, 5, 4, 3, 3 and 3, respectively (Beier et al., 2017).

2.7 Comparative analysis

The Mauve v.2.4.0 plugin (Darling et al., 2004) in Genious R9.0.2. was used to conduct multiple sequence alignments for the collinearity analysis of *Phrynium* and *Stachyphrynium*, with the aim of identifying potential rearrangements and inversions. Comparative analysis of the chloroplast genomes was carried out using mVISTA (Frazer et al., 2004) online program (<https://genome.lbl.gov/vista/index.shtml>), utilizing the Shuffle-LAGAN mode to illustrate the variations present among the genomes. Additionally, the sliding window analysis of the whole chloroplast genomes and nrDNA were executed using DnaSP v6.0 (Rozas et al., 2017) to determine the nucleotide diversity (Pi). For chloroplast genomes, 600 bp window length and 200 bp step size were employed, while for nrDNA, 50 bp window length and 20 bp step size was utilized.

2.8 Phylogenetic analysis

Phylogenetic relationships were reconstructed using Bayesian inference (BI) and maximum likelihood (ML) methods based on CDS and nrDNA sequences. In addition to the sequences from *Phrynium* and *Stachyphrynium*, the outgroup *Marantochloa purpurea* of Marantaceae, cultivated in China, was incorporated into the phylogenetic analysis. (Table 1). Sequences were aligned using MAFFT v7 (Kato and Standley, 2013) with default

parameters and then trimmed with trimAL (Capella-Gutierrez et al., 2009) in TBtools, which served to remove poorly aligned regions and enhance the accuracy of the resulting.

For the ML analyses, the optimal nucleotide substitution model was determined using ModelFinder (Kalyanamoothy et al., 2017) based on the Akaike Information Criterion (AIC). ML trees were constructed using IQ-TREE v1.6.1 (Lanfer et al., 2020) with 1,000 bootstrap replicates. For the BI analyses, jModelTest2 (Darriba et al., 2012) was used for the model selection with Bayesian Information Criterion (BIC) as the standard, and MrBayes v3.2.7 (Ronquist et al., 2012) was utilized to construct the BI trees with default settings. The Markov Chain Monte Carlo (MCMC) analysis was run for 2 million generations, with tree sampling occurring every 1,000 generations. Convergence of the MCMC was assumed when the average standard deviation of split frequencies (ASDF) fell below 0.01. The initial 25% of the trees were discarded due to the unreliability of early MCMC sampling data, and the remaining trees were used to generate consensus trees and estimate Bayesian posterior probabilities (PPs).

3 Results

3.1 Features of chloroplast genome and nrDNA

In this study, 40 chloroplast genomes were newly sequenced and assembled, including 30 from *Phrynium* and 10 from *Stachyphrynium*. The genome size of *Phrynium* ranged from 166,401 bp (*P. tonkinense*) to 172,065 bp (*P. pedunculiferum*) (Table 2). The chloroplast genome maps for two species of *Phrynium* and one species of *Stachyphrynium* are shown in Figure 1, with the remaining species are provided in the supplementary materials (Supplementary Figure 1). These genomes exhibit a characteristic quadripartite structure, consisting of a pair of IR regions (26,258-35,228 bp) that are separated by the LSC (89,787-92,551 bp) and SSC (11,071-21,631 bp) regions, with the SSC region displaying the most significant variation in length. The total GC content varied from 35.8 to 36.4%, with a predominant concentration at 36.2%. Generally, the GC content in the IR regions (38.8-42.8%) is higher than that in both the LSC region (33.9-34.3%) and the SSC region (28.6-30.9%). In contrast, the genome lengths of *Stachyphrynium* are notably shorter than those of *Phrynium*, ranging from 162,035 bp (*S. placentalium*) to 167,870 bp (*S. spicatum*). The observed variation in sequence

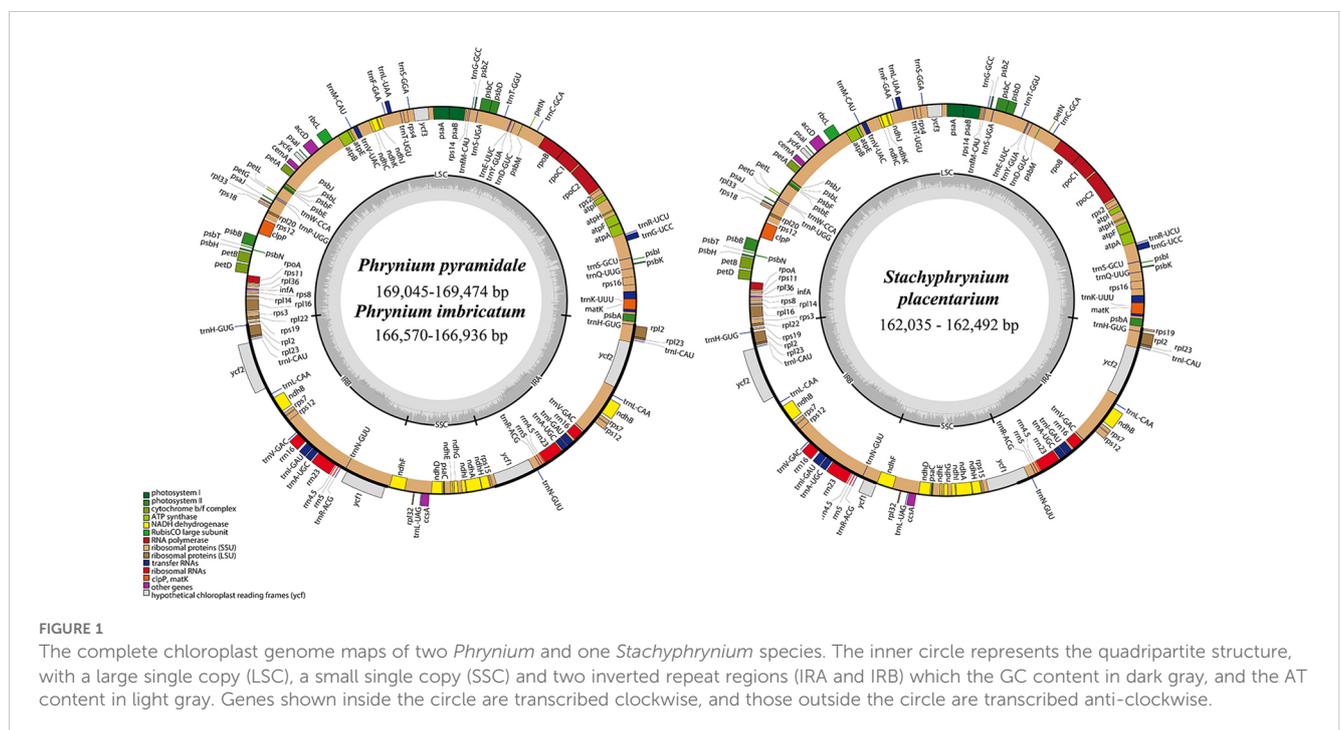
TABLE 2 Summary of the complete chloroplast genomes characteristics of *Phrynium* and *Stachyphrynium* species.

Species	Total length (bp) GC (%)	LSC length (bp) GC (%)	SSC length (bp) GC (%)	IR length (bp) GC (%)	Number of genes		
					Protein	tRNA	rRNA
<i>P. pyramidale</i>	169,045-169,474 36.2	91,223-91,494 34.1	14,868-14,928 30.0-30.1	31,362-31,660 40.5-40.6	86	30	4
<i>P. pubinerve</i>	171,807 35.8	90,281 34.2	11,071 29.6	35,228 38.8	86	30	4
<i>P. tonkinense</i>	166,401-166,815 36.0-36.1	91,878-92,551 33.9	20,656-20,859 28.8-29.0	26,258-27,037 42.4-42.8	87	30	4
<i>P. imbricatum</i>	166,570-166,936 36.4	89,787-90,003 34.3	14,129-14,147 30.9	31,219-31,501 40.5-40.7	86	30	4
<i>P. hainanense</i>	166,773-166,843 36.1	91,838-91,883 34	20,806-20,830 28.8-28.9	27,065 42.4	87	30	4
<i>P. pedunculiferum</i>	171,625-172,065 35.9	91,339-91,864 34.0	15,536-15,614 29.1-29.3	32,304-32,375 40.1	87	30	4
<i>P. yunnanense</i>	166,694 36.0	91,858 33.9	21,631 28.6	26,602 42.5	87	30	4
<i>S. spicatum</i>	167,870 36.2	89,441 34.2	14,291 30.8	32,069 40.1	87	30	4
<i>S. placentarium</i>	162,035-162,492 36.5	88,228-88,566 34.2-34.4	18,725-18,877 29.9	27,471-27,606 42.2	87	30	4

length can be primarily attributed to differences in the lengths of the SSC region (14,291-18,877 bp) and the IR regions (27,471-32,069 bp). The overall GC content exhibited a high degree of similarity (36.2-36.5%), with specific GC content values of 40.1% to 42.2% in the IR regions, 34.2% to 34.4% in the LSC region, and 29.9% to 30.8% in the SSC region (Table 2).

The chloroplast genomes of *Phrynium* and *Stachyphrynium* demonstrate a significant degree of similarity, with gene

arrangements maintained in the same order. These genomes encoded 112-113 unique genes, which include 86-87 protein-coding genes (PCGs), 30 tRNA genes, and 4 ribosomal RNA genes (Table 3). A gene count revealed that *P. imbricatum* and *P. pyramidale* contained a single copy of the *rps19* within the LSC region, whereas other species exhibited two copies in each IR region. Additionally, a pseudogene *ycf1* was identified in *P. hainanense*, *P. tonkinense*, *P. yunnanense* and *S. placentarium*,



while the *ndhA* gene was found to be deleted in *P. pubinerve* (Figure 1; Supplementary Figure 1).

The nrDNA sequence lengths of seven species of *Phrynum* are 5,792-5,821 bp comprising five fragments, i.e., the 18S rRNA gene, ITS1, 5.8S rRNA gene, ITS2 and the 26S rRNA gene. The lengths of these fragments are 1,811 bp, 195-206 bp, 164 bp, 207-229 bp, and 3,411-3,414 bp, respectively. The overall nrDNA GC content is 57.2-57.6%, with the GC content in the ITS regions being the highest among the five fragments (65.1-68.0%). The total length of the nrDNA in *Stachyphrynum*, along with the lengths of its individual components, closely resembles that of the *Phrynum* (Table 4).

3.2 Contraction and expansion of IR regions

To investigate the mechanism driving variation in chloroplast genomes, we conducted a detailed comparative analysis of the four junctions (JLB, JSB, JSA, and JLA) between the LSC, SSC, and two

IR regions across seven species of *Phrynum* and two species of *Stachyphrynum*. This analysis focused on the locations of the IR boundaries and their adjacent genes. The findings indicate that the two junctions between the LSC and IR regions, JLB (LSC/IRb) and JLA (IRa/LSC) exhibit relative stability. In contrast, the junctions JSB (IRb/SSC) and JSA (SSC/IRa) display greater variability among the species examined (Figure 2). With the exception of the JLBs found in *P. pyramidale* and *P. imbricatum*, which were situated between *rps19* and *trnH* genes, all other JLBs in *Phrynum* and *Stachyphrynum* species were located between the *rpl22* and *rps19* genes. The *rpl22* gene was identified in the LSC region, positioned 21-103 bp from the IRb region, while the *rps19* gene was found 18 bp from the JLB instead of *rpl22*. Similarly, aside from the JLAs of *P. pyramidale* and *P. imbricatum*, which were located between the *trnH* and *psbA* genes, all other types were positioned between the *rps19* and *psbA* genes. In *Stachyphrynum*, JLAs were found 74 to 98 bp from *rps19*, whereas in *Phrynum*, they were located 176 to 230 bp from *psbA*. In the chloroplast genomes of *P. hainanense*, *P. tonkinense*, *P. yunnanense* and *S.placentarium*, the *ycf1* gene expanded from the SSC to the IRa region, with a range of 607-

TABLE 3 Gene contents of *Phrynum* and *Stachyphrynum* complete chloroplast genomes.

Gene function	Gene category	Name of gene
Self-replication	Large subunit of ribosome	<i>rpl2^a×2</i> , <i>rpl14</i> , <i>rpl16[*]</i> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23^a×2</i> , <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
	Small subunit of ribosome	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7^a×2</i> , <i>rps8</i> , <i>rps11</i> , <i>rps12^{a,b,**}×2</i> , <i>rps14</i> , <i>rps15</i> , <i>rps16[*]</i> , <i>rps18</i> , <i>rps19^a×2</i>
	DNA-dependent RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1[*]</i> , <i>rpoC2</i>
	Ribosomal RNAs	<i>rrn4.5^a×2</i> , <i>rrn5^a×2</i> , <i>rrn16^a×2</i> , <i>rrn23^a×2</i>
	Transfer RNAs	<i>trnA</i> -UGC ^{a,*} ×2, <i>trnC</i> -GCA, <i>trnD</i> -GUC, <i>trnE</i> -UUC, <i>trnF</i> -GAA, <i>trnM</i> -CAU, <i>trnG</i> -GCC, <i>trnG</i> -UCC [*] , <i>trnH</i> -GUG ^a ×2, <i>trnI</i> -CAU ^a ×2, <i>trnI</i> -GAU ^{a,*} ×2, <i>trnK</i> -UUU [*] , <i>trnL</i> -CAA ^a ×2, <i>trnL</i> -UAA [*] , <i>trnL</i> -UAG, <i>trnM</i> -CAU, <i>trnN</i> -GUU ^a ×2, <i>trnP</i> -UGG, <i>trnQ</i> -UUG, <i>trnR</i> -ACG ^a ×2, <i>trnR</i> -UCU, <i>trnS</i> -GCU, <i>trnS</i> -GGA, <i>trnS</i> -UGA, <i>trnT</i> -GGU, <i>trnT</i> -UGU, <i>trnV</i> -GAC ^a ×2, <i>trnV</i> -UAC [*] , <i>trnW</i> -CCA, <i>trnY</i> -GUA
Photosynthesis	Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	Photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i> , <i>ycf3^{**}</i>
	Subunits of NADH-dehydrogenase	<i>ndhA[*]</i> , <i>ndhB^{**}×2</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>
	Subunits of cytochrome b/f complex	<i>petA</i> , <i>petB[*]</i> , <i>petD[*]</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF[*]</i> , <i>atpH</i> , <i>atpI</i>
	Large subunit of rubisco	<i>rbcL</i>
Other genes	Translation initiation factor	<i>infA</i>
	Maturase	<i>matK</i>
	Protease	<i>clpP^{**}</i>
	Envelope membrane protein	<i>cemA</i>
	Subunit of accetyl-CoA-carboxylase	<i>accD</i>
	Cytochromesynthesis	<i>ccsA</i>
Genes of unknown function		<i>ycf1^a×2</i> , <i>ycf2^a×2</i> , <i>ycf4</i>

^aGenes in inverted repeat regions; ^btrans-splinting gene; ^{*}genes containing one intron; ^{**}genes containing two introns; ×2 shows genes with duplicates.

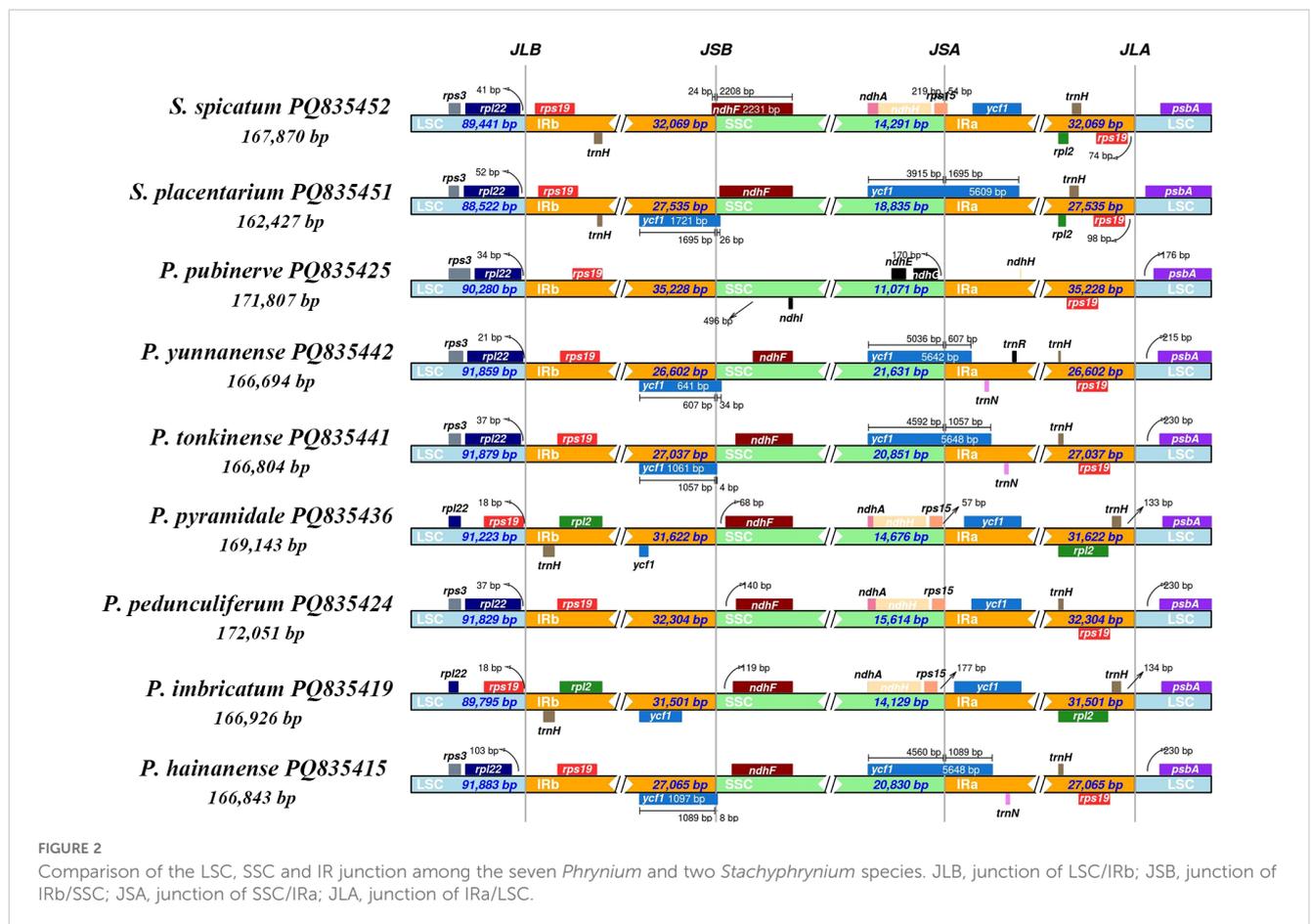
TABLE 4 Summary of nrDNA (18S-ITS1-5.8S-ITS2-26S) characteristics of *Phrynium* and *Stachyphrynium* species.

Species	Total length (bp)	18S rRNA length (bp)	ITS1 length (bp)	5.8S rRNA length (bp)	ITS2 length (bp)	26S rRNA length (bp)	Total GC content (%)	ITS GC content (%)
<i>P. pyramidale</i>	5,812-5,813	1,811	202	164	222-223	3,413	57.4-57.5	66.9-67.1
<i>P. pubinerve</i>	5,821	1,811	206	164	229	3,411	57.6	67.4
<i>P. tonkinense</i>	5,792-5,797	1,811	195-198	164	207-211	3,413	57.4	65.7-66.2
<i>P. imbricatum</i>	5,800	1,811	202	164	210	3,413	57.2-57.3	65.1
<i>P. hainanense</i>	5,795-5,798	1,811	195	164	212-215	3,413	57.5	66.9-67.2
<i>P. pedunculiferum</i>	5,805-5,808	1,811	195-198	164	220-221	3,414	57.4-57.6	66.0-67.6
<i>P. yunnanense</i>	5,799	1,811	195	164	216	3,413	57.6	68.0
<i>S. spicatum</i>	5,809	1,810	195	164	231	3,409	57.6	68.6
<i>S. placentarium</i>	5,806	1,810	195	164	230	3,407	57.6	66.5-67.2

1,695 bp. Additionally, the *ycf1* gene was also present in the IRb region in these species due to partial duplication. This truncation likely results in a nonfunctional protein, leading to the classification of *ycf1* as a pseudogene. In the remaining species, *ycf1* was entirely located within the IRa region and being duplicated. Regarding the *rps19* gene, it was entirely situated within the IR regions in all species, with the exception of *P. pyramidale* and *P. imbricatum*, where it was located in the LSC region.

3.3 Relative synonymous codons usage

The analysis of codon usage frequency and relative synonymous codon usage (RSCU) was conducted using protein-coding genes. The protein-coding genes in the chloroplast genomes of the seven *Phrynium* species comprised between 18,329-20,711 codons, encoding a total of 20 amino acids, while the two *Stachyphrynium* species contained between 20,332-20,432 codons. The results



indicate that RSCU values were comparable across different genera (Figure 3). The codons UAA, UGA and UAG were identified as stop codons. The RSCU analysis revealed that nearly all amino acids, with the exception of methionine (Met) and tryptophan (Trp), were represented by 2 to 6 synonymous codons. Among the 64 codons analyzed for the *Phrynium* and *Stachyphrynium* species, 30 exhibited RSCU values exceeding 1, indicating a preference for certain codons, with the majority (29/30, 96.67%) end with the base A or U. Similarly, 31 codons had RSCU values below 1, with most (28/31, 90.32%) end with the base C or G. The codons AUG and UGG, which correspond to methionine and tryptophan respectively, displayed no significant bias (RSCU = 1.00). Notably, the UUA codon, which encodes leucine, exhibited the highest RSCU value of approximately 1.90, while the AGC codon, encoding serine, demonstrated the lowest RSCU value at approximately 0.29.

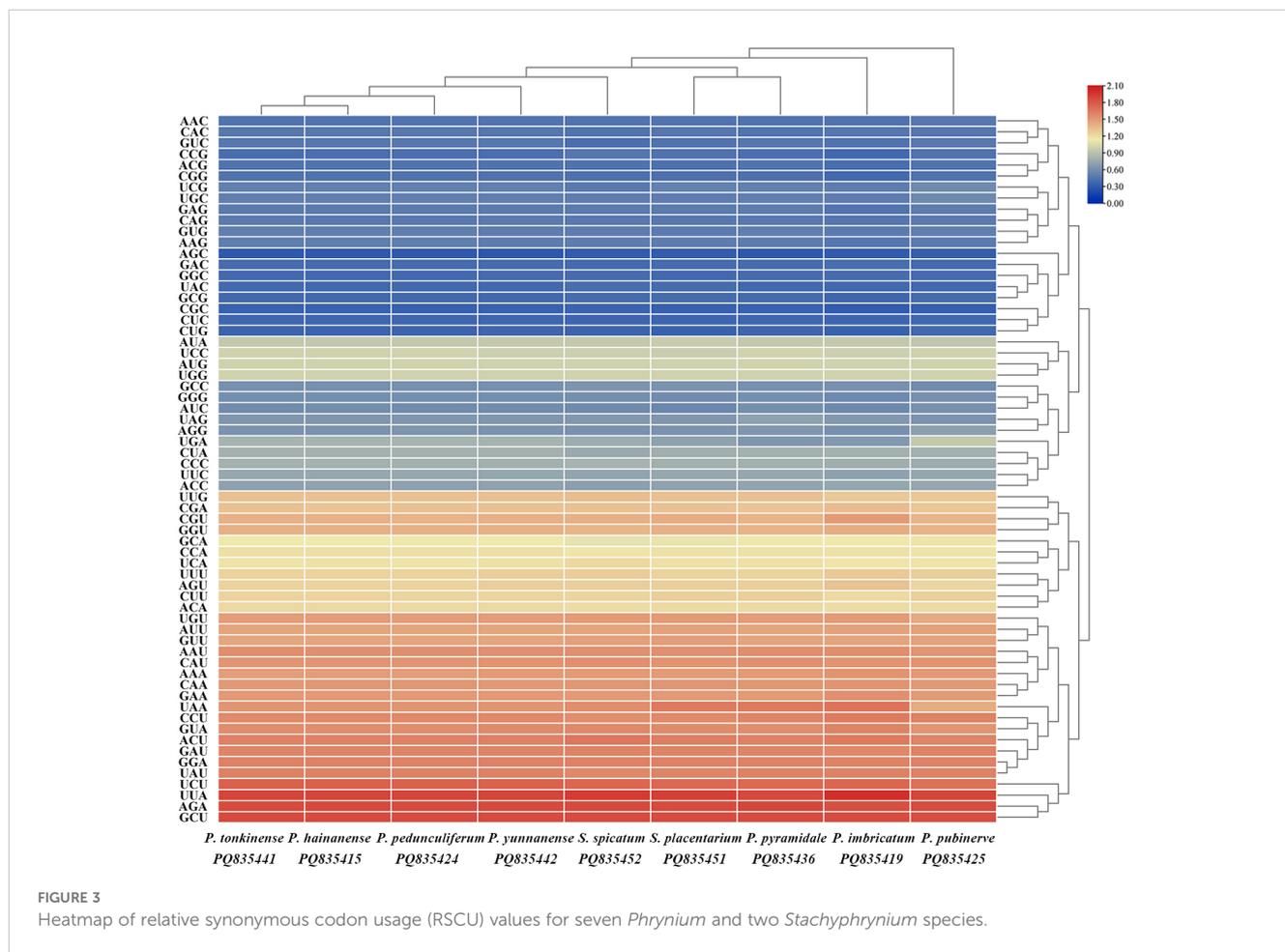
12.95%, 16.07%, 6.12%, and 6.54%, respectively. Notably, SSRs composed of A/T bases represented a significant proportion (27.73-48.0%). The distribution of SSRs in chloroplast was observed to be uneven, with 65.3% of SSRs located in the LSC region, 15.93% in the SSC region, and 18.78% in the IR regions. This distribution indicates a higher level of polymorphic variation within the LSC region (Figure 4). However, the proportion of SSRs in the IR region increased (45.38%) in *P. pubinerve*. In the case of *Stachyphrynium*, a total of 73-101 SSRs were identified, with no hexanucleotide repeats detected in *S. placentalium*. The predominant type of repeat observed was mononucleotides, which constituted approximately 53.42-57.43% of the total SSR loci. Consistent with the findings in *Phrynium*, the majority of SSRs in *Stachyphrynium* were also distributed in the LSC region (74.14%).

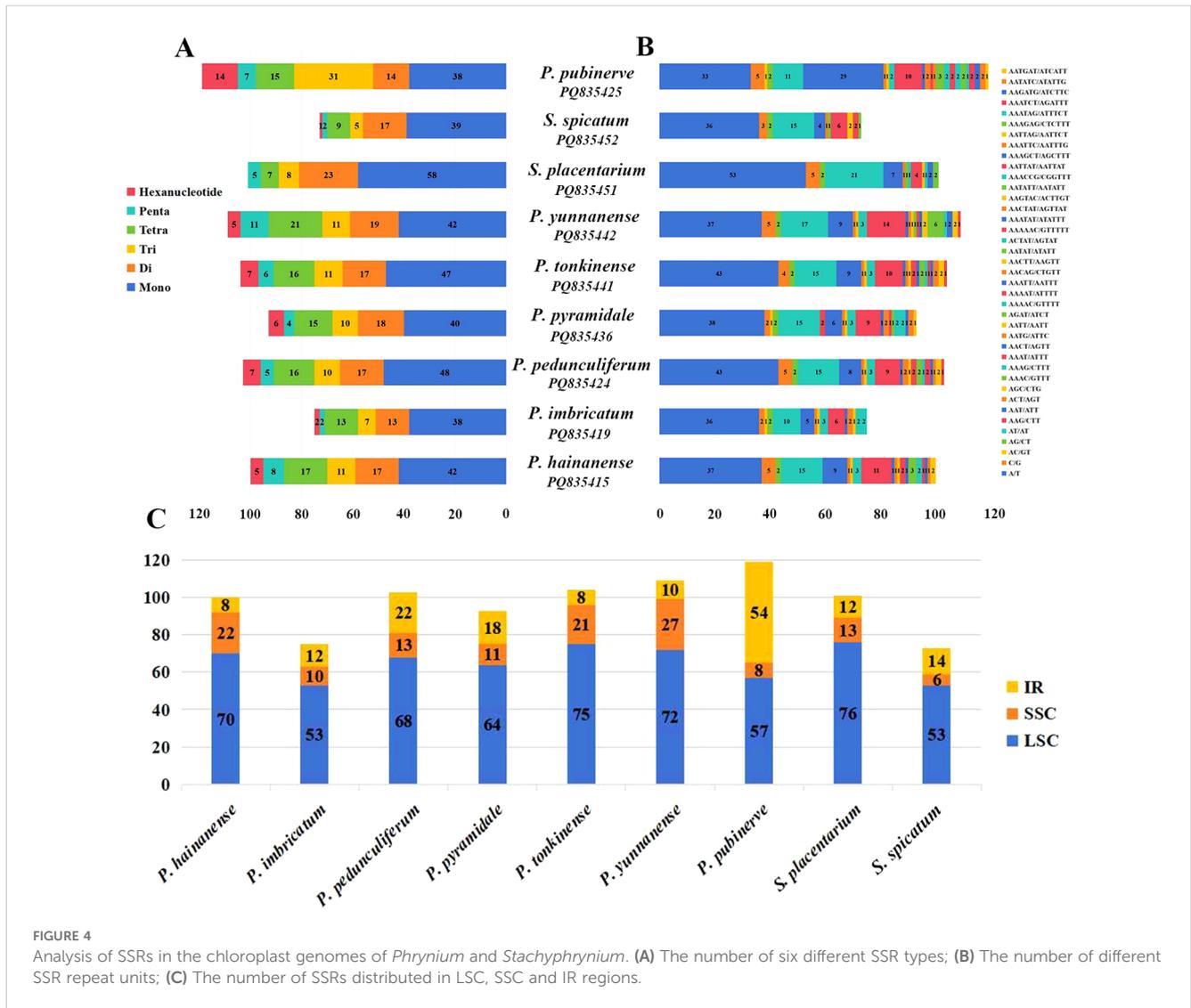
3.4 Simple sequence repeat analysis

A total of 75 to 119 SSRs were identified in chloroplast genomes of *Phrynium* (Figure 4). The average percentages of SSRs, ranging from mononucleotides to hexanucleotides, were 41.96%, 16.36%,

3.5 Comparative analysis

The multiple sequence alignments performed with mVISTA software revealed a notable similarity in the complete chloroplast genomes of the examined species of *Phrynium* and *Stachyphrynium*. The results (Figure 5) indicated that the non-coding regions of the





chloroplast genomes exhibited a higher frequency of variation compared to the coding regions, suggesting that sequence diversity is more pronounced in the non-coding regions. The SSC and LSC regions exhibited significantly greater variability compared to the IR regions, while the rRNA genes remained highly conserved, showing minimal variation. Among the coding regions, the *ycf1* and *ycf2* gene were identified as highly variable, whereas the non-coding regions, including *trnS-trnG*, *atpH-atpI*, *petN-psbM*, *trnE-trnT*, *rps4-trnL*, *psbE-petL*, *accD-psaI*, *ndhF-rpl32*, *psaC-ndhI* and *ndhH-ycf1*, exhibited a higher degree of differentiation.

The collinearity analysis of *Phrynum* and *Stachyphrynium*, conducted using Mauve, demonstrated the presence of locally collinear blocks (LCBs) within these chloroplast genomes (Figure 6). The LSC region and its adjacent regions exhibited a greater consistency in gene order compared to the SSC and IR regions. A significant distinction was observed in *P. pubinerve*, characterized by an inversion of an approximately 8 kb segment that spans from *ndhF* to *ndhI*, transitioning from the SSC to the IR region, with the inversion occurring at the junction of IRb/SSC.

3.6 Sequence divergence and mutational hotspots

Nucleotide diversity (P_i) values of the complete chloroplast genomes and nrDNA were assessed to identify hot spots of sequence divergence in *Phrynum* and *Stachyphrynium*. The sliding window analysis revealed that the P_i values of the chloroplast genomes of seven *Phrynum* species varied from 0-0.02669 over a 600 bp region, with greater divergence observed in LSC and SSC regions compared to the IR regions. A total of ten highly variable loci were identified in the chloroplast genomes of *Phrynum*, including *trnS-GCU-trnG-UCC* (0.02669), *rps12-psbB* (0.01761), *ycf1* (0.01592), *clpP* (0.01592), *atpH-atpI* (0.01444), *ycf2* (0.01287), *rps4-trnL-UAA* (0.01232), *ndhH-trnN-GUU* (0.01216), *rpl16* intron (0.01147), and *trnE-UUC-trnT-GGU* (0.01135). The intergenic spacer regions exhibited significantly higher P_i values compared to the coding regions, with only two coding regions (*ycf1* and *clpP*) displaying elevated values (Figure 7A). The nrDNA sequences from the seven *Phrynum* species contained 170 variable sites and 146

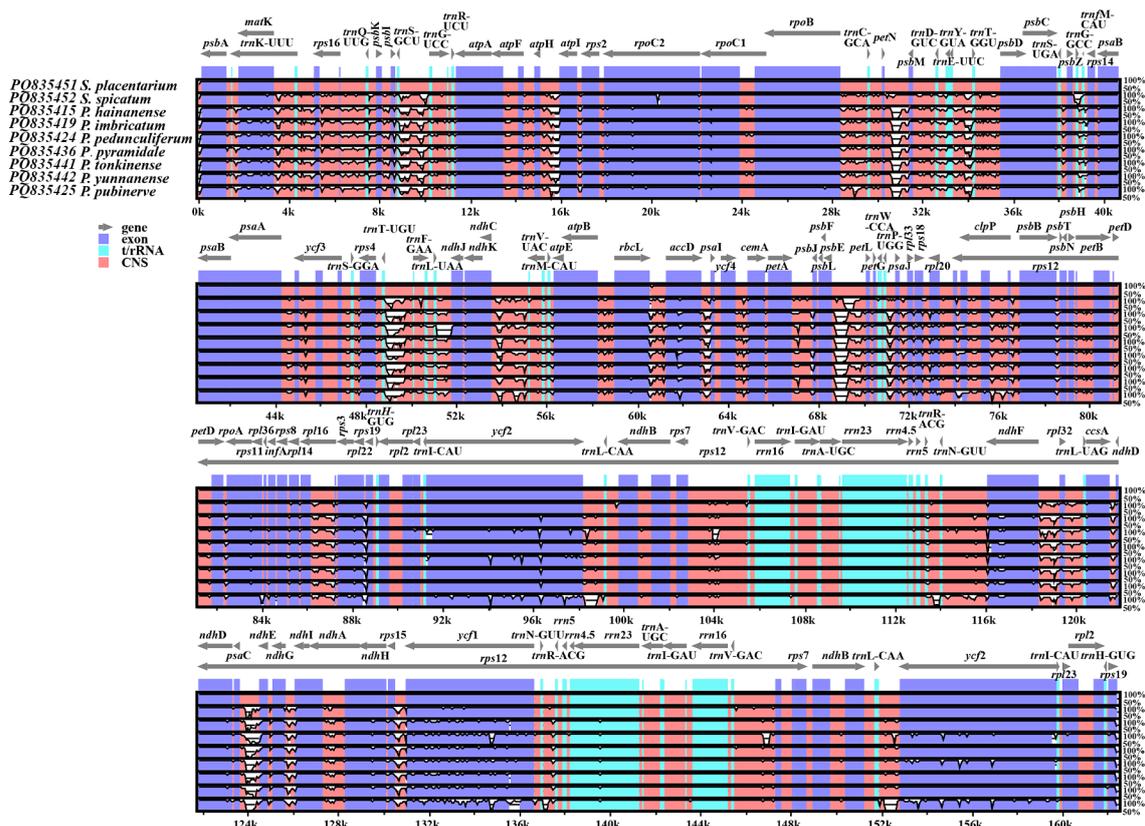


FIGURE 5

Comparison of seven *Phrynium* and two *Stachyphrynium* species of chloroplast genomes with *S. placentarium* as a reference sequence. Gray arrows above the alignment indicate the orientation of each gene. The colored areas indicate the exon, t/rRNA, and conserved non-coding sequences (CNS). The vertical axis indicates the percentage of identity, ranging from 50 to 100%.

parsimony-informative sites, accounting for 2.91% and 2.50% of the total sequences, respectively. The figure (Figure 7B) shows that the majority of the variation is concentrated in the ITS1 region. The Pi values of nrDNA ranged from 0-0.11429, with average values of 0.05982 for ITS1, 0.05321 for ITS2, 0.0104 for 5.8S rRNA, 0.00201 for 18S rRNA and 0.00779 for 26S rRNA.

A similar pattern was observed in *Stachyphrynium*, where regions of the chloroplast genomes with elevated Pi values included *psaC-ndhE* (0.03678), *trnS-GCU-trnG-UCC* (0.03167), *ndhH-ycf1* (0.01993), *ndhE-ndhI* (0.01856), *atpH-atpI* (0.01622), *trnK-UUU* (0.01563), *trnT-UGU-trnL-UAA* (0.01563), *trnE-UUC-trnI-TGGU* (0.014), and *ndhF-rpl32* (0.01363) (Figure 7C). The nrDNA sequences from the two *Stachyphrynium* species contained 69 variable sites, representing 1.18% of the total sequence. The Pi values of nrDNA ranged from 0-0.03678, with the average values of 0.02502 for ITS1, 0.01381 for ITS2, 0.0025 for 5.8S rRNA, 0.00067 for 18S rRNA and 0.00145 for 26S rRNA (Figure 7D).

3.7 Phylogenetic analysis

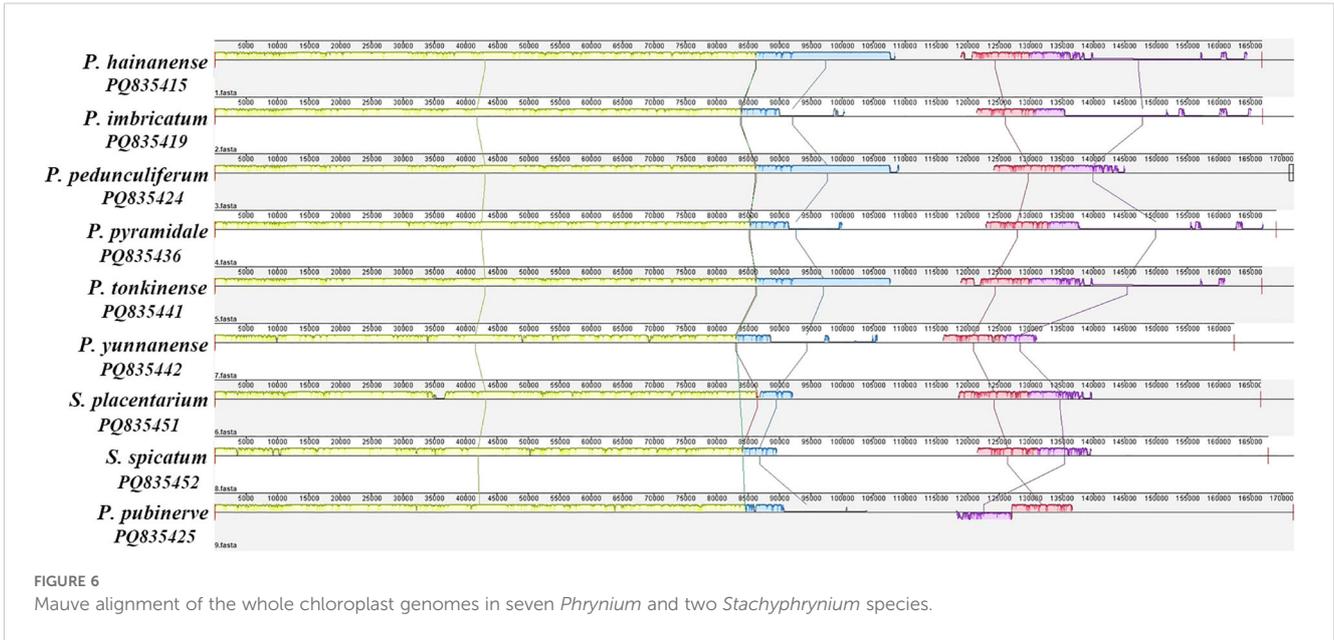
Phylogenetic trees were constructed using the Maximum Likelihood (ML) and Bayesian Inference (BI) methodologies, based on the analysis of CDS and nrDNA sequences, with *Marantochloa purpurea*

(Marantaceae) serving as the outgroup (Figure 8). The topology of the ML and BI trees demonstrated a high degree of similarity, and several nodes on these trees received strong support. The Chinese *Phrynium* species were grouped into a monophyletic clade, which was sister to that of *Stachyphrynium*. This clade could be further subdivided into two distinct clades: one consisting of *P. pyramidale*, *P. pubinerve* and *P. imbricatum* (BS=100%, PP=1.00), and another encompassing all remaining Chinese species. The position of *P. yunnanense* varied across different trees. In the CDS tree, it formed an independent branch that was sister to other species, situated adjacent to *P. pedunculiferum* and *P. tonkinense*. Conversely, in the nrDNA tree, *P. yunnanense* exhibited a closer relationship with *P. pedunculiferum*, subsequently merging with *P. hainanense* and *P. tonkinense*. Additionally, the *Stachyphrynium* species in China also constituted a monophyletic clade.

4 Discussion

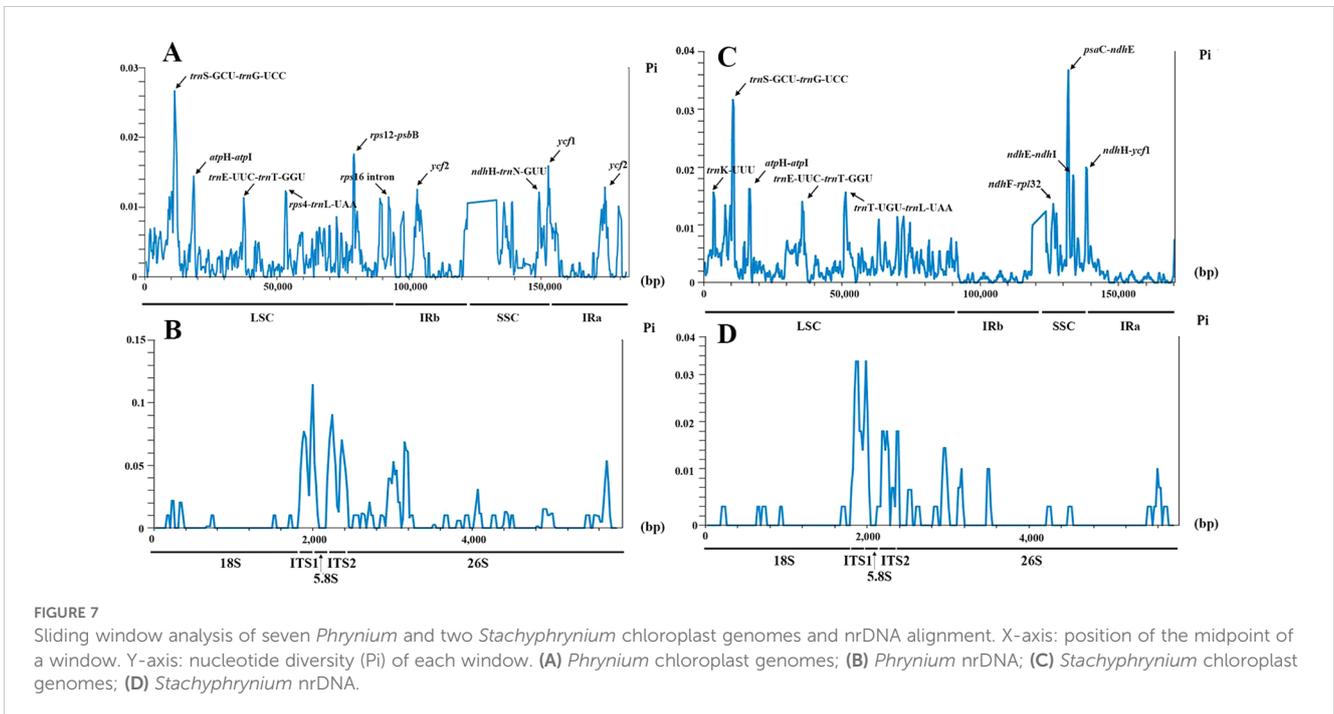
4.1 General characteristics of the chloroplast of *Phrynium* and *Stachyphrynium*

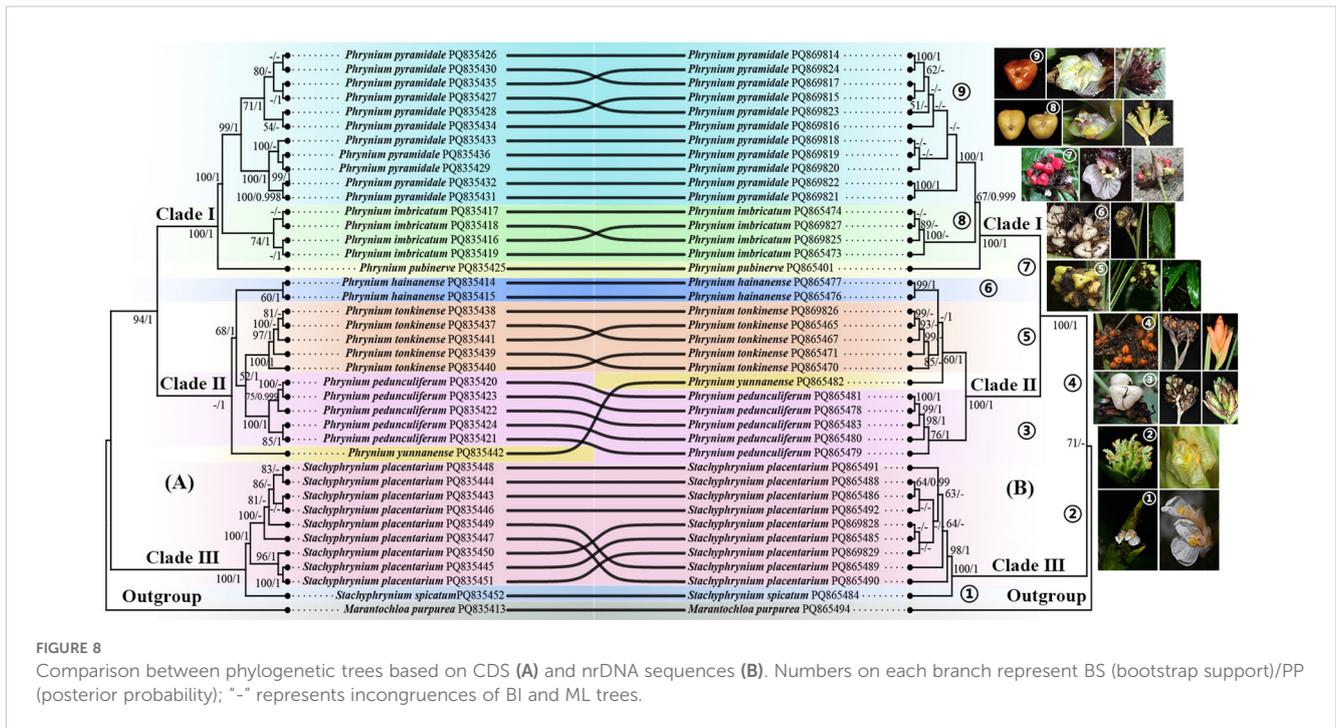
The chloroplast genomes of seven *Phrynium* and two *Stachyphrynium* species, newly reported in this study, exhibits a



significant degree of similarity regarding genome structure, gene order, and gene composition. The distribution of GC content in the chloroplast genomes was uneven, with elevated GC content observed in the IR regions compared to the LSC and SSC regions. This difference may be explained by the presence of four types of rRNAs (*rrn4.5*, *rrn5*, *rrn16*, and *rrn23*) in the IR region, while the reduced GC content in the SSC region is likely due to the presence of NADH dehydrogenase genes, which generally exhibit a lower GC content relative to genes associated with photosynthesis and the genetic system (Jansen and Ruhlman, 2012). Notably, these chloroplast genomes are considerably larger than the average

chloroplast genome size of 150 kb typically observed in most flowering plants (Ruhlman and Jansen, 2014). Although the IR region is often regarded as the most conserved segment of the chloroplast genomes, variations in its boundaries, resulting from contraction and expansion, are commonly observed throughout chloroplast genome evolution. These variations are associated with changes in the length of the chloroplast genomes and may lead to phenomena such as gene loss, gene duplication, gene rearrangement, and pseudogenization in certain angiosperms (Zhu et al., 2016). A notable structural change observed in *P. pubinerve* includes an inversion of the *ndh1* gene, alongside the





absence of the *ndhA* gene. In this species, the *ndhA* gene was transferred from the SSC region to the IR region due to the expansion of the IR region, and the instability at the IR/SSC junction may be associated with the deletion of the *ndhA* gene. The *ndhA* gene is a member of the NADH dehydrogenase gene family. In the light reaction phase of photosynthesis, the protein encoded by this gene actively engages in the photosynthetic electron transport chain, promoting the progress of the light reaction and ensuring the highly efficient functioning of photosynthesis. Some studies have revealed that the hemiparasitic trait in plants can hasten the pseudogenization and loss of plastid NDH genes (Li et al., 2021). The exact reason for the absence of the *ndhA* gene in the chloroplast genomes of *P. pubinerve* remains challenging to explain. The species-specific loss of certain NDH genes across different taxa suggests that not all NDH genes are universally involved in or required for photosynthetic electron transport. The functional deficiency caused by the loss of a single NDH gene (e.g., *ndhA*) may be compensated by other NDH genes or nuclear-encoded genes (Mohanta et al., 2020). Therefore, whether the absence of *ndhA* in *P. pubinerve* is linked to its ecological environment in the shaded and wet forest deserves further study. Inversions within plastomes are recognized as a powerful phylogenetic feature due to their rarity and minimal homoplasy (Schwarz et al., 2015). However, underlying mechanisms driving inversions in plant genomes remain poorly understood. In the chloroplast genomes of *P. hainanense*, *P. yunnanense*, *P. tonkinense* and *S. placentarium*, as the IR regions expands towards the SSC regions, the *ycf1* gene, which spans the SSC/IRa boundary, exhibits partial duplication at the IRb/SSC boundary. This truncation likely results in a nonfunctional protein, and thus *ycf1* is considered as a pseudogene. At the transcriptional level,

duplicated parts may alter the chromatin structure around the *ycf1* gene, interfering with the transcription factors' binding to the promoter. This reduces *ycf1* transcription and mRNA abundance, ultimately impacting protein expression levels. At the protein level, when duplicated parts are located near the crucial domains or functional sites of the encoded protein, they disrupt its folding, resulting in an abnormal structure that impedes related function. The *ycf1* pseudogene could serve as a novel molecular marker, providing additional phylogenetic insights. By comparing the presence/absence of the *ycf1* pseudogene, sequence characteristics, and IR region variations across species, it is possible to reconstruct plant evolutionary relationships and reveal their history (Dong et al., 2015). While pseudogenes typically evolve under neutral selection, some research indicates that they can be repurposed as functional genes with distinct roles following intense purifying selection pressure (Mahmudi et al., 2015). It is noteworthy that the expansion of IR regions appears to occur more progressively in monocots compared to non-monocot angiosperms, with a greater number of LSC genes being converted into IRs (Wang et al., 2008).

4.2 Simple sequence repeats and codon usage analysis

SSRs were extensively distributed in the chloroplast genomes of eukaryotic organisms. Due to their simple structure, relative conservation, and polymorphism, SSRs are recognized as effective molecular markers and are widely used in species identification, genetic diversity, and population genetics research (Powell et al., 1995; Provan et al., 2001; He et al., 2012). In the chloroplast genomes of *Phrynium* and *Stachyphrynium*, a total of 73 to 119

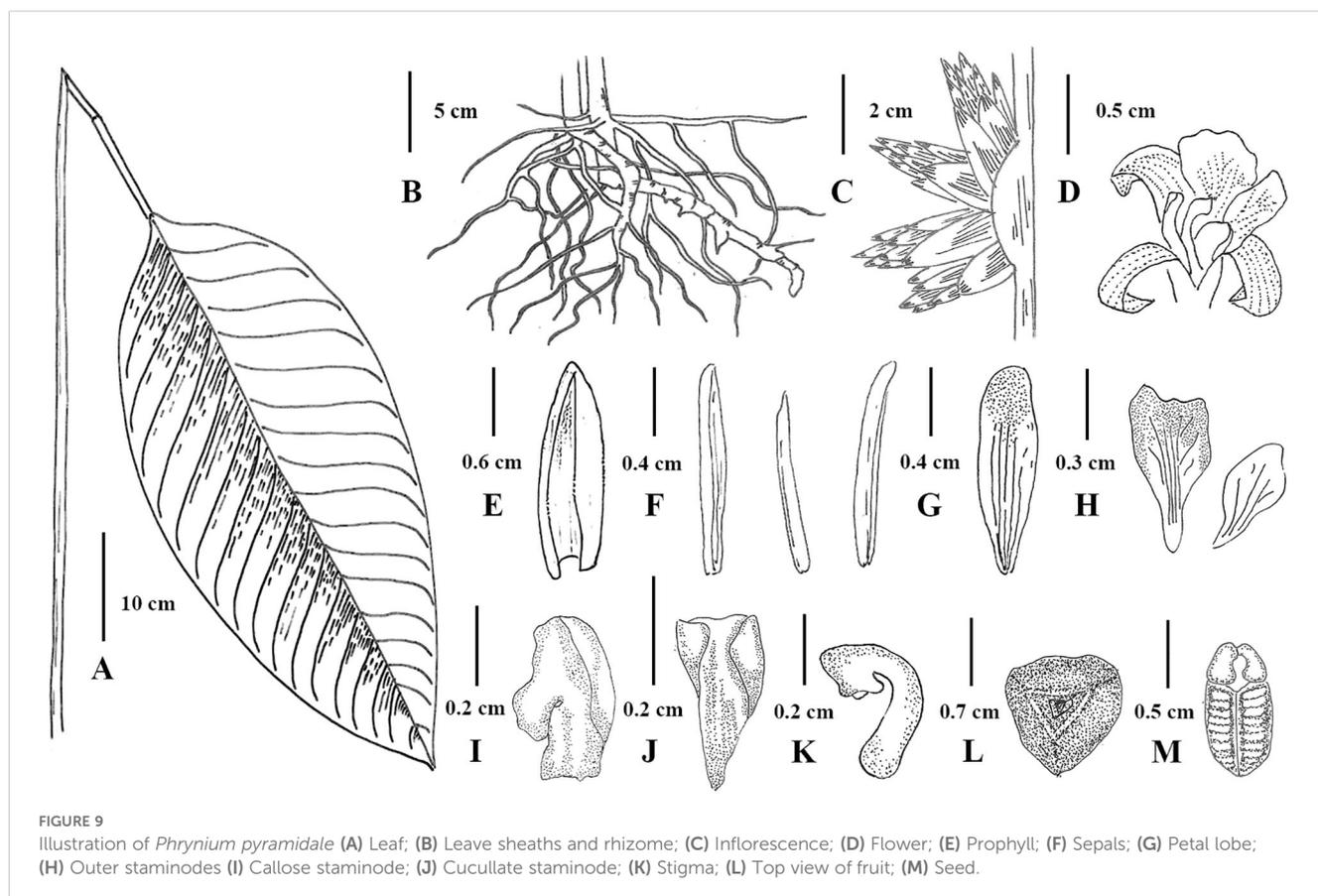
SSRs were identified, encompassing repeat units ranging from mononucleotides to hexanucleotides. The findings for both genera align with the typical characteristics of SSRs observed in the chloroplast genomes of angiosperms, particularly the predominance of mononucleotide A/T repeats, which are generally composed of short polyadenine (poly A) or polythymine (poly T) repeats, while G or C repeats are infrequently encountered (Powell et al., 1996). Mononucleotide repeats may play a more significant role in genetic variation compared to other types of SSRs. Furthermore, SSRs are primarily located in the LSC region, indicating a high level of nucleotide variability in this area, which may provide valuable insights for detecting both intraspecific and interspecific polymorphisms at the population level.

Throughout the course of long-term evolution, species have gradually evolved specific codon usage patterns during the process of genes encoding proteins in response to their physiological requirements. This phenomenon is referred to as codon usage bias (Zhou et al., 2008; Sheng et al., 2021). RSCU is a critical metric for examining codon usage bias. In *Phrynium* and *Stachyphrynium*, the number of high-frequency codons with an RSCU greater than 1 is 29 and 28, respectively, with the majority end with an A/U base. This observation indicates a strong preference for A/U-ending codons in these genomes, a trend that is consistent with findings in other angiosperms (Nanjala et al., 2022; Jiang et al., 2023; Xu et al., 2023). Codons are essential for the

accurate expression of genetic information, and codon usage bias is closely linked to the long evolutionary history of organisms. Consequently, investigating codon utilization in these genera is of considerable significance for the exploration of the evolutionary models of species in *Phrynium* and *Stachyphrynium*.

4.3 Molecular markers

Highly variable genomic regions are widely used as DNA barcoding markers for species identification and phylogenetic studies. The identification of species in *Phrynium* and *Stachyphrynium* presents challenges due to the pronounced morphological similarities among them. Given that certain traditional DNA barcode fragments are inadequate for accurate species identification and phylogenetic analysis within these genera, it is essential to explore more highly variable regions at the generic level that may serve as effective markers for future studies on species differentiation. Through mVISTA and nucleotide diversity analyses, four highly variable intergenic regions (*trnS-trnG*, *atpH-atpI*, *trnE-trnT*, *trnT-trnL*) were identified in seven *Phrynium* species and two *Stachyphrynium* species. Furthermore, the coding regions *ycf1* and *ycf2* exhibited significantly higher variability in *Phrynium*. The *ycf1* gene has been recognized as the most variable region in the chloroplast genomes, exhibiting a mutation rate that exceeds that



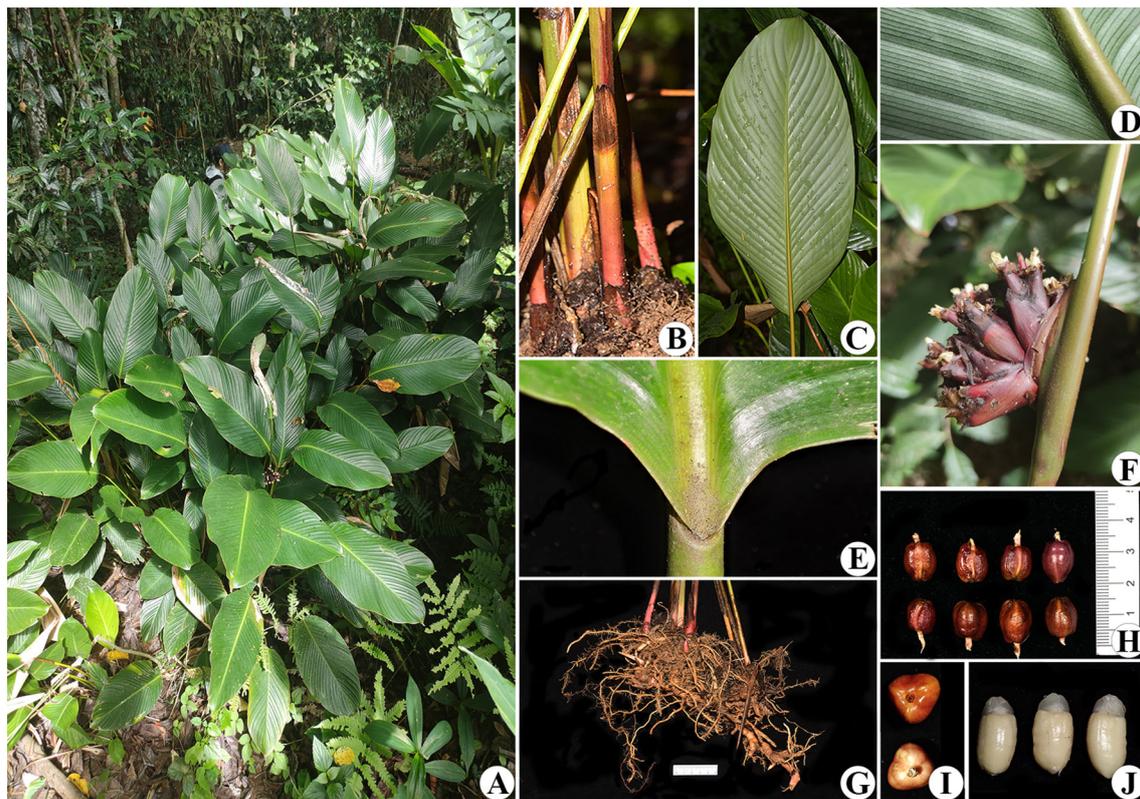


FIGURE 10

Phrynium pyramidale (A) Habit; (B) Leaf sheaths; (C) Abaxial leaf surface; (D) Details of the abaxial leaf surface; (E) Lamina base and pulvinus; (F) Inflorescence with blooming flowers; (G) Rhizome; (H, I) Young fruits; (J) Seeds.

of established chloroplast barcode candidates such as *matK* and *rbcl*. This has led to its increasing application as a potentially valuable molecular marker, particularly in recent studies of orchids (Neubig et al., 2008; Whitten et al., 2014; Dong et al., 2015). Additionally, *ycf2* alone can provide a well-supported phylogeny that generally corresponds with phylogenetic trees derived from more comprehensive multigene or plastid genomes datasets (Huang et al., 2010). The intergenic regions identified in the current study have also been documented in other plant species at the species level. For instance, the *atpH-atpI* and *trnS-trnG* regions exhibited significant variability, facilitating effective species differentiation in Aroideae subfamily (Araceae) (Dong et al., 2012; Li et al., 2022). The *trnT-trnL* region has been employed to analyze the phylogenetic relationships of *Kengyilia* (Poaceae) species (Gao et al., 2014).

Phylogenetic trees of seven species of *Phrynium* and two species of *Stachyphrynium* utilizing four molecular markers that were independently and combinatorially screened were also constructed. The resulting trees exhibited a topology consistent with that of the CDS and the nrDNA tree, revealing that *Phrynium* is further divided into two distinct clades, while *Stachyphrynium* constituted an independent clade. However, certain groups of species (such as *P. tonkinense*, *P. pedunculiferum*) within each tree were not clearly distinguished (Supplementary Figures 2A–F). These results suggest that the identified regions may function as

potential molecular markers and offer significant resources for phylogenetic investigations at both interspecific and intraspecific levels. Although several candidate barcode regions have been recognized, further research is necessary to assess the effectiveness of these highly differentiated markers.

4.4 Phylogenetic relationships

Prince and Kress (2006) conducted phylogenetic analyses of Marantaceae using *matK*, the 3' intergenic spacer region, and the *trnL-F* region. Their findings indicated that *Phrynium* is non-monophyletic, with *Phacelophrynium* nested within it. The *Stachyphrynium* clade, which includes *Stachyphrynium*, was grouped with the *Maranta* clade and formed a sister clade to the *Donax* clade (which includes *Phrynium*). Suksathan et al. (2009) further analyzed phylogenetic relationships within Asian Marantaceae using the *rps16* intron, ITS1 and NTS. The results revealed that *Stachyphrynium* constitutes a strongly supported monophyletic group, while the genera, *Phrynium*, *Phacelophrynium*, *Monophrynium* and *Cominsia* form a well-supported monophyletic clade. Based on these findings, they proposed consolidating these four genera within the *Phrynium* complex into a single genus, as phylogenetic evidence does not

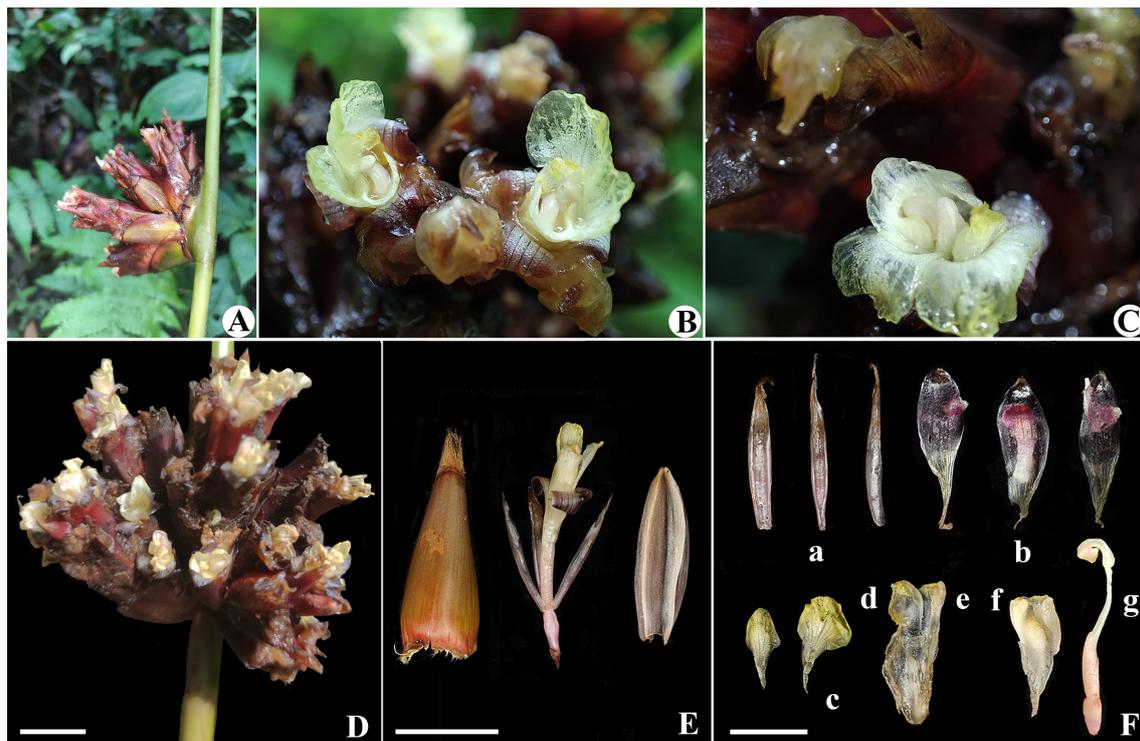


FIGURE 11

Phrynium pyramidale (A) Lateral view of inflorescence; (B, C) Close up of flower; (D) Top view of inflorescence; (E) Fertile bract, flower and prophyll; (F) Flower dissection: (a) Sepals; (b) Petal lobes; (c) Outer staminodes; (d) Callose staminode; (e) Fertile stamen; (f) Cucullate staminode; (g) Pistil. Scalebars=1cm.

support their distinctiveness. Furthermore, *Stachyphrynium* exhibits closer phylogenetic relationship with African genera. Clearly, comprehensive taxon sampling of *Phrynium* is essential to evaluate its monophyly and clarify relationships with other genera.

In this analysis, *Phrynium* and *Stachyphrynium* are positioned on separate branches. *Phrynium* is further subdivided into two clades supported by morphological characteristics, while *Stachyphrynium* forms an independent clade. Within clade I, *P. pyramidale*, *P. imbricatum* and *P. pubinerve* clustered closely. Their synapomorphies include a shorter peduncle (<3 cm). Four species—*P. hainanense*, *P. tonkinense*, *P. pedunculiferum*, and *P. yunnanense*—are grouped into Clade II. Clade II is characterized by having subtriangular fruits and most of them have a long peduncle, except *P. tonkinense*. The shortened peduncle might be a consequence of the continuous adaptive evolution of plants in response to environmental pressures. *P. yunnanense* is distinguished from the other species by its longer peduncle (20–45 cm), bright orange bracts and fruits, along with green flowers. The remaining species exhibit clear distinctions: while both *P. pedunculiferum* and *P. hainanense* produce white fruits, *P. pedunculiferum* is characterized by green, leathery bracts that remain intact at maturity. In contrast, *P. hainanense* has bracts that dissolve into a blackish fibrous mass. Additionally, *P. tonkinense* is notable for its

distinctly longer and narrower lanceolate leaves compared to other species in the genus.

It is a commonly observed that phylogenetic trees constructed from chloroplast and nuclear genomes frequently exhibit topological inconsistencies. These discrepancies may result from incomplete lineage sorting, horizontal gene transfer, recombination or convergent molecular evolution (Steenwyk et al., 2023). The phylogenetic position of *P. yunnanense* exemplifies this issue. However, it receives weak support in both genome-based trees, and as only one sample is included, its phylogenetic placement remains uncertain, warranting further investigation.

4.5 The new species

The new species, *Phrynium pyramidale*, exhibits a suite of unique morphological characteristics, including a pagoda-like inflorescence, reddish-brown fertile bracts that fragment into numerous fibers apically, yellow flowers, and subtriangular brown fruits. Its distinct species status is further supported by molecular evidence. Phylogenetic analyses of 40 samples of *Phrynium* and *Stachyphrynium*, based on chloroplast CDS and nrDNA datasets (Figure 8), reveal that eleven populations of *P. pyramidale* form a strongly supported monophyletic clade, which is sister to *P.*

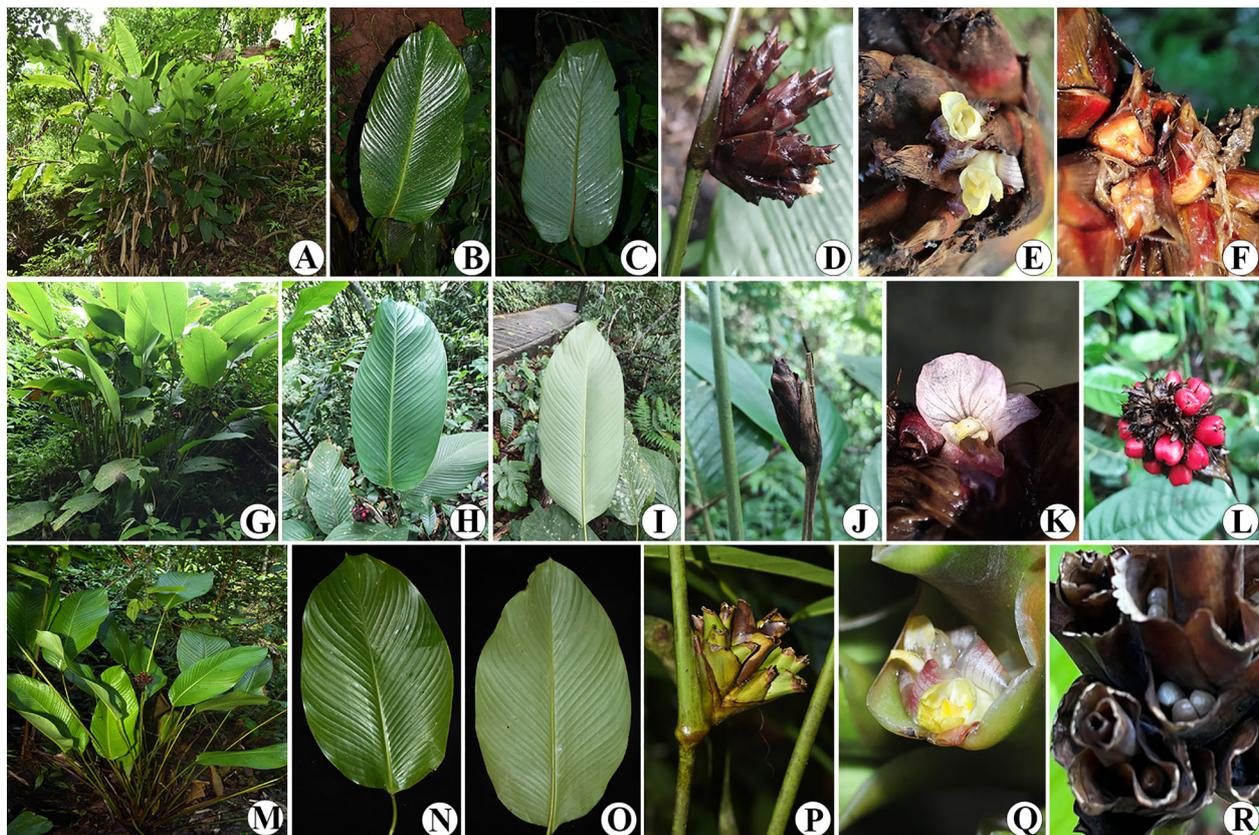


FIGURE 12

Phrynium pyramidale (A) Habit; (B) Adaxial leaf surface; (C) Abaxial leaf surface; (D) Inflorescence; (E) Close up of flower; (F) Young fruits. *Phrynium pubinerve* (G) Habit; (H) Adaxial leaf surface; (I) Abaxial leaf surface; (J) Inflorescence; (K) Close up of flower; (L) Young fruits. *Phrynium imbricatum* (M) Habit; (N) Adaxial leaf surface; (O) Abaxial leaf surface; (P) Inflorescence; (Q) Close up of flower; (R) Young fruits.

pubinerve and *P. imbricatum*. *P. pyramidale* is characterized by bracts arranged in a pagoda-like structure with the apex decaying rapidly but never exceeding one-third of the total length of bracts. Molecular data indicate that *P. imbricatum* and *P. pubinerve* are the closest relatives of *P. pyramidale*. However, *P. imbricatum* can be distinguished by its green, leathery larger bracts (2.8–3.7 × 1.3–2.4 cm), with entire apices and seeds that bearing a large appendages. In contrast, *P. pubinerve* is characterized by purplish-red flowers, bright red fruits and bracts that rapidly wither and dissolve into a blackish fibrous mass.

5 Conclusion

In this study, we conducted a comprehensive analysis of chloroplast genomes across *Phrynium* and *Stachyphrynium*, including genome assembly, comparative genomics, and phylogenetics. Results reveal strong structural conservation regarding structure, gene content, gene arrangement, and codon usage bias. However, divergences were detected in genome size and the SC/IR boundaries. We further identified genomic hotspot regions contributing to intergeneric differentiation, which could

serve as potential DNA barcodes for species identification. Phylogenetic analysis based on CDS (coding sequences) and nrDNA (nuclear ribosomal DNA) strongly support the monophyly of *Phrynium* and *Stachyphrynium* with *Phrynium* further subdivided into two distinct clades. These findings deepen our understanding of their molecular evolution and provide critical insights for identification, utilization, and conservation of these medicinal plants. Additionally, we describe and illustrate a new species, *P. pyramidale*.

6 Taxonomy

Phrynium pyramidale Y. Tong & Z. Y. Lin, sp. nov. (Figure 9–12)

Diagnosis:—*Phrynium pyramidale* is similar to *Phrynium pubinerve* and *Phrynium imbricatum*, but it is distinguished by having pagoda-like inflorescence, the reddish-brown fertile bracts which only splitting up into many fibers at the tip, the yellow flowers and the subtriangular brown fruits. (Table 5).

Type:—CHINA. Yunnan Province: Xishuangbanna Dai Autonomous Prefecture, Xiding Country, Zhanglang Village,

TABLE 5 Morphological comparison of *Phrynium pyramidale*, *P. pubinerve* and *P. imbricatum*.

Characters	<i>P. pyramidale</i>	<i>P. pubinerve</i>	<i>P. imbricatum</i>
Sheath	glabrous	densely pubescent	glabrous
Leaves	2-6 per shoot, abaxially fresh green, adaxially light green	2-4 per shoot, abaxially light green, adaxially yellowish green	2-3 per shoot, abaxially dark green, adaxially yellowish green
Pulvinus	6-12 cm long	2-12.5 cm long	3-12 cm long
Bracts	broadly elliptical, 2-2.9 × 1.6-2.2 cm, yellowish green	broadly ovate to narrowly elliptic, 1.9-6.7 × 1.1-3.7 cm, yellowish green or brown	ovate-oblong, 2.8-3.7 × 1.3-2.4 cm, reddish-brown or light green
Fertile bracts	reddish-brown, broadly lanceolate or broadly elliptical, glabrous	broadly elliptic, reddish-brown or yellowish green	ovate to ovate-lanceolate, reddish-brown or light green
Flowers	yellow, 2.2-2.5 cm long	purplish-red, 1.8-2.0 cm long	yellow, 2.5-2.8 cm long
Sepals	subulate, 11-13 × 2.5-3 mm	linear, 8.7-11 × 2.0-2.5 mm	linear, 10.2 × 1.1 mm
Petal lobes	elliptic to oblong, 11-13 × 3.5-4 mm, pale purple	oblanceolate, 6.2-8.8 × 2.5-3.1 mm, purplish-red	elliptic-oblong, 6-7.3 × 3 mm, purplish-red
Ovary	glabrous	densely hairy	glabrous
Fruits	reddish brown, elongate to triangular, 12-14 × 10-12 mm	bright red, elongate to triangular, 11-15 × 10-12 mm	light brown, elongate to triangular, 11-13 × 7-10 mm
Seeds	grey, trigonous, 9-11 × 5-7 mm, shallowly grooved	grey, subquadrangular, 9-11 × 6-7 mm, slightly rough, pubescent, with translucent, bifid appendages	white, slightly trigonous ellipsoid, 10-11 × 5-6 mm, shallowly grooved, with very large appendages, 3 mm in diam

growing on a gentle slope situated beneath the forest adjacent to the temple, 21°54'58.78"N, 100°07'08.12"E, 1,542 m a.s.l., 8 July 2023, *Tong Yi TY230708014* (holotype: GUCM!, isotype: IBSC!).

Description:—Rhizomatous ground herb 0.6-2 m tall. Basal leaves 2-6 per shoot; sheath 20-55 cm long, red to brown, glabrous; petiole 40-135 cm long, light green, glabrous; pulvinus 6-12 cm long, yellowish green, glabrous; lamina oblong to elliptic, 40-55 × 13-27 cm, thinly leathery, adaxially fresh green, glabrous, abaxially light green, glabrous to sparsely hairy, base rounded, apex acuminate, acumen 0.5-1.5 cm. Inflorescence interfoliar, erect, peduncle largely obscured by an enveloping leaf sheath, 0.8-1.5 cm long, light green, glabrous; synflorescence capitate, densely congested, 5-8 cm in diameter; bracts subtending the proximal branches broadly elliptical, 2-2.9 × 1.6-2.2 cm, yellowish green, soon becoming dry and turning brown, broadly elliptical; fertile bracts 1.5-3.4 × 1.2-2.3 cm, reddish-brown, glabrous, broadly lanceolate or broadly elliptical, spirally arranged, apex initially acute, soon splitting up into many fibres at the tip; flower pairs 3 per special paraclade, associated prophylls 17-20 × 8.5-10 mm, ovate to lanceolate, with two distinct keels on the abaxial surface and incurved at margins, interphylls 14-18 × 4 mm. Flowers 2.2-2.5 cm long; sepals 3, subulate, 11-13 × 2.5-3 mm, yellowish-white, semi-translucent; corolla tube 8 mm, cylindrical; staminodial tube as long as the corolla tube; petal lobes elliptic to oblong, 11-13 × 3.5-4 mm, semi-translucent pale purple, dotted with small purple spots, deflexed and curled; outer staminodes 2, slightly unequal, petaloid lobes broadly obovate, 5-7 × 2.5-3.5 mm, light yellow; cucullate staminode 3-3.5 × 1.5-2 mm, cream-white, yellow at the tip; callose staminode 3.5-4 × 2.5-3 mm; fertile stamen small, 3 mm long; style with free part 6 mm long, curved inwards, cream-white; stigma light yellow, ca. 2 mm in diameter; ovary oblong, 3-4

mm long, 3-loculed. Fruits reddish brown, shiny, elongate to triangular, 12-14 × 10-12 mm; seeds 3, grey, trigonous, shallowly grooved, 9-11 × 5-7 mm, arillate, aril cream-white, extending into 2 curled subulate, translucent appendages.

Etymology:—The specific epithet refers to the pagoda-like inflorescence.

Habitat, distribution and phenology:—*Phrynium pyramidale* is currently known from Yunnan and Xizang. And it also found in Laos, India and Thailand (Figure 13). It inhabits in moist valleys and dense forests, at an altitude of about 1,000-1,500 m. Flowering from June to August; fruiting from July to September.

Conservation status:—*Phrynium pyramidale* is widespread in China, south to Laos and Vietnam, and southwest to India. It exhibits well-developed vegetative reproduction, and is also highly adaptable to the local natural environment and capable of surviving in human-disturbed habitats. It is commonly cultivated by local communities around homes for use in wrapping traditional rice dumplings during festivals. Its populations are often large and frequently encountered within its distribution range. Therefore, according to the IUCN Red List Categories and Criteria (IUCN, 2024), it should be assessed as Least Concern (LC).

Specimens Examined:—CHINA: Yunnan: Menghai, Mengsong Township, Sanmai Village, 22°03'00"N, 100°39'34"E, 1,541 m a.s.l., 24 Apr. 2012, *Menghai Census Team 5328220074* (IMDY); Yuanjiang, Donge Town, Namujing, 23°42'N, 101°49'E, 1,000 m a.s.l., 28 May. 1983 (fl.), *Tao G.D. 38659* (HITBC, KUN); Puer, Jinggu, Zhengxing town, 23°22'29.17"N, 100°55'58.71"E, 981 m a.s.l., 17 Dec. 2019, *ZXX191871* (KUN); Wenshan, Maguan, Gulijing County, Macaoping Village, 22°41'35.83"N, 103°54'22.66"E, 950 m a.s.l., 12 May. 2019 (fl.), *Z. L. Dao, L. Cai, P.*

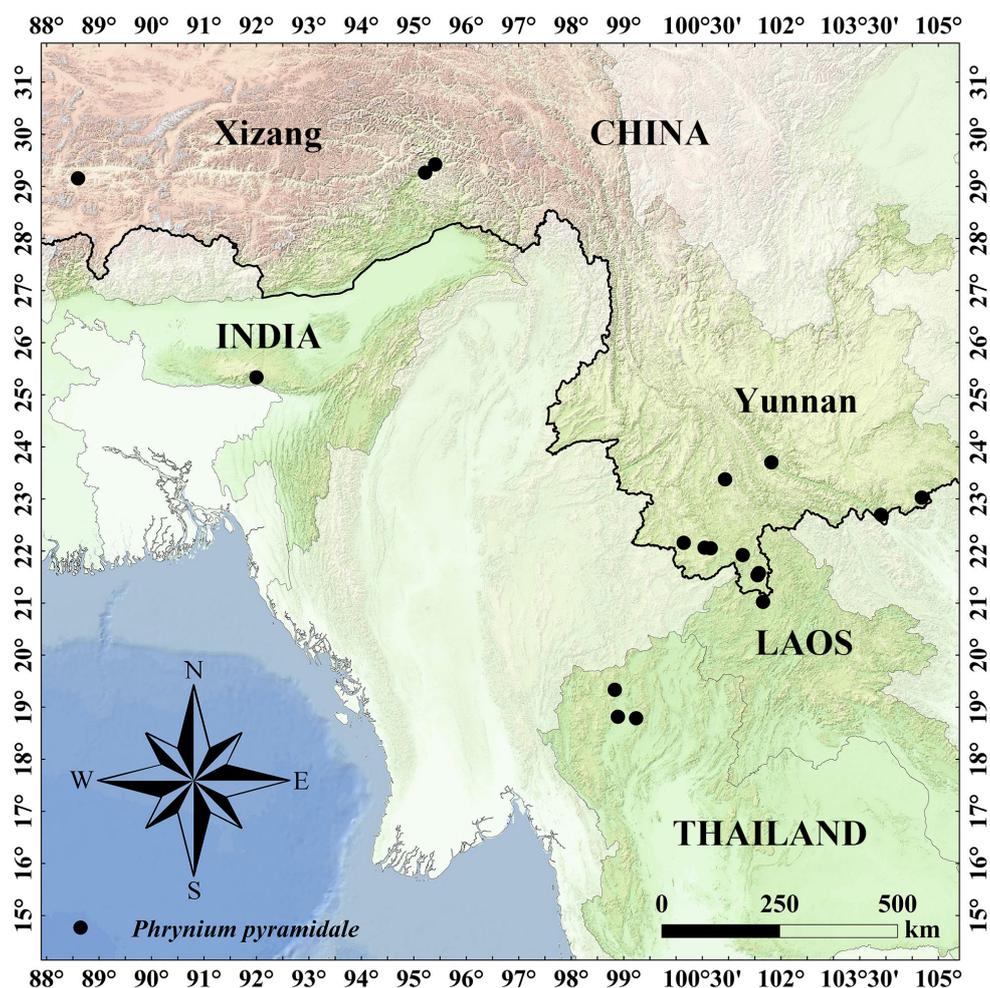


FIGURE 13
Distribution map of *Phrynium pyramidale* (black circle).

Zhang KIBDZL168B09 (KUN); 1,450 m a.s.l., 29 May. 2013 (fl.), L. Wu 3935 (BNU); Lincang, Mayidui Village, Near Nanbo River, 1,350 m a.s.l., 24 Aug. 1957 (fr.), J. S. Xin 333 (IBSC, KUN, LBG, PE); Xishuangbanna, Jinghong, Mengyang Town, Kunlun Highway, 1,400 m a.s.l., 7 Jul. 1957 (fr.), Sino-Soviet Joint Expedition 5956 (IBSC, IBK, KUN). Xizang: Linzhi, Motuo, Miri Village, 29° 25'18.04"N, 95°24'26.30"E, 833 m a.s.l., 27 May. 2013 (fl.), C. Liu, J. Cai, T. Zhang 13CS6428 (KUN); From Tea garden to Buqunhu Lake, 29°15'37.75"N, 95°12'57.20"E, 1,101 m a.s.l., 14 Nov. 2016, C. Liu, J. D. Ya, H. J. He 16CS11896 (KUN). LAOS: Oudomxai, Ban Nam Pheng, 21°01'12"N, 101°39'25"E, 800 m a.s.l., 23 Jun. 1999, M. F. Newman 875 (E). INDIA: Assam, Nongpoh, Khasi Hills, 2,000 ft., 13 May. 1949, Koelz WN 22692 (L); Meghalaya, Cherrapunjee, Khasi Hills, 4,000 ft., 15 May. 1952. Thakur Rup Chand 5647 (MICH). THAILAND: Chiangmai: Doi Sutep, 18°48'58.5"N, 98° 53'30.52"E, 1,120 m a.s.l., 17 Apr. 1958, Sørensen, T; Larsen, K; Hansen, B 2838 (L); Ru-See Cave area, 18°47'14.68"N, 99° 14'17.23"E, 1,075 m a.s.l., 10 Oct. 1987, Maxwell, JF 87-1148 (L); Chang Kian Valley, 18°48'58.5"N, 98°53'30.52"E, 1,000 m a.s.l., Maxwell, JF 88-619 (L); Chiang Dao base hill, 19°20'N, 98°50'E, 600-700 m a.s.l., 6 May. 1973, R. Geesink 5723 (AAU, L, P).

Data availability statement

The original contributions presented in this research are accessible to the public. The data that support the findings of this study can be found in the Genbank database at <https://www.ncbi.nlm.nih.gov/under> the accession number PQ835413-PQ835452, PQ865401, PQ865465, PQ865467, PQ865470-PQ865471, PQ865473-PQ865474, PQ865476-PQ865486, PQ865488-PQ865492, PQ865494, PQ869814-PQ869829.

Author contributions

ZL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. ZF: Data curation, Formal analysis, Resources, Writing – original draft. YL: Formal analysis, Resources, Visualization, Writing – original draft. LF: Formal Analysis, Resources, Visualization, Writing – review & editing. FX: Resources, Visualization, Writing – review & editing. YT: Funding acquisition, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by Science and Technology Planning Project of Guangdong Province, China (Grant No.2021B1212100007), the Fund for Improving Medical Service and Security Capacity of Department of Social Security, Ministry of Finance (Grant No. (2019)39), the National Natural Science Foundation of China (No.32270408), and the Special Subsidies for Public Health Services of TCM “the National Survey of TCM Resources” [DSS, MOF. 66/2017] from National Administration of Traditional Chinese Medicine of People’s Republic of China.

Conflict of interest

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

References

- Aii, J., Kishima, Y., Mikami, T., and Adachi, T. (1997). Expansion of the IR in the chloroplast genomes of buckwheat species is due to incorporation of an SSC sequence that could be mediated by an inversion. *Curr. Genet.* 31, 276–279. doi: 10.1007/s002940050206
- Amiryousefi, A., Hyvonen, J., and Pocza, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* 34, 3030–3031. doi: 10.1093/bioinformatics/bty220
- Andersson, L. (2001). Phylogeny and classification of marantaceae. *Bot. J. Linn Soc.* 135, 275–287. doi: 10.1006/bojl.2000.0418
- Beier, S., Thiel, T., Munch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198
- Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Cen, A., Deng, W., Mo, X., Xiao, P., Zou, F., Xiao, C., et al. (2017). The investigation and study on wild resource of *Phrynium* willd in qianxinan prefecture. *Seed* 01, 60–63. doi: 10.16590/j.cnki.1001-4705.2017.01.060
- Chan, P. P., Lin, B. Y., Mak, A. J., and Lowe, T. M. (2021). tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* 49, 9077–9096. doi: 10.1093/nar/gkab688
- Chao, Z., Hong, X., Xie, X., Huang, R., and Tian, E. (2023). Phylogenomic analysis of *Bupleurum* in Western Sichuan, China, including an overlooked new species. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1294670
- Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., et al. (2023). TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Mol. Plant* 16, 1733–1742. doi: 10.1016/j.molp.2023.09.010
- Clausager, K., and Borchsenius, F. (2003). The marantaceae of sabah, northern borneo. *Kew Bull.* 58, 647–678. doi: 10.2307/4111147
- Cole, T., Hilger, H., and Stevens, P. (2019). Angiosperm phylogeny poster (APP) – Flowering plant systematics. *PeerJ* 7, e2320v6. doi: 10.7287/peerj.preprints.2320v6
- Cron, G. V., Pirone, C., Bartlett, M., Kress, W. J., and Specht, C. (2012). Phylogenetic relationships and evolution in the strelitziaceae (Zingiberales). *Syst. Bot.* 37, 606–619. doi: 10.1600/036364412x648562
- Darling, A. C., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704
- Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772. doi: 10.1038/nmeth.2109
- Dong, W., Liu, J., Yu, J., Wang, L., and Zhou, S. (2012). Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One* 7, e35071. doi: 10.1371/journal.pone.0035071
- Dong, W., Xu, C., Li, C., Sun, J., Zuo, Y., Shi, S., et al. (2015). *ycf1*, the most promising plastid DNA barcode of land plants. *Sci. Rep.* 5, 8348. doi: 10.1038/srep08348
- Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.
- Fang, D. (2002). A new species of *phrynium* willd. (Marantaceae) from guangxi. *J. Trop. Subtrop Bot.* 10, 250–252. doi: 10.3969/j.issn.1005-3395.2002.03.010
- Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32, W273–W279. doi: 10.1093/nar/gkh458
- Fu, L., Ye, Y. S., and Liao, J. P. (2017). *Phrynium yunnanense* (Marantaceae), a new species from Yunnan, China. *Phytotaxa* 307, 89–94. doi: 10.11646/phytotaxa.307.1.9
- Fu, N., Xu, Y., Jin, L., Xiao, T. W., Song, F., Yan, H. F., et al. (2024). Testing plastomes and nuclear ribosomal DNA sequences as the next-generation DNA barcodes for species identification and phylogenetic analysis in *Acer*. *BMC Plant Biol.* 24, 445. doi: 10.1186/s12870-024-05073-w
- Gagnepain, F. (1932). *Flore générale de l’Indo-Chine* Vol. 6 (Paris: Masson & Cie), 121–137.
- Gao, G., Tang, Z., Wang, Q., Gou, X., Ding, C., Zhang, L., et al. (2014). Phylogeny and maternal donor of *Kengyilia* (Triticeae: Poaceae) based on chloroplast *trnT-trnL* sequences. *Biochem. Syst. Ecol.* 57, 102–107. doi: 10.1016/j.bse.2014.07.004
- Green, B. R. (2011). Chloroplast genomes of photosynthetic eukaryotes. *Plant J.* 66, 34–44. doi: 10.1111/j.1365-3113X.2011.04541.x
- Guangdong Institute of Botany (1977). “*Marantaceae*,” in *Flora hainanica*, vol. 4. (Science Press, Beijing), 104–106.
- Guangxi Institute of Botany (2016). “*Marantaceae*,” in *Flora of guangxi*, vol. 5. (Guangxi: Guangxi Science and Technology Publishing House), 114–117.
- He, S., Wang, Y., Volis, S., Li, D., and Yi, T. (2012). Genetic diversity and population structure: implications for conservation of wild soybean (*Glycine soja* Sieb. et Zucc) based on nuclear and chloroplast microsatellite variation. *Int. J. Mol. Sci.* 13, 12608–12628. doi: 10.3390/ijms131012608
- Holttum, R. E. (1951). The marantaceae of malaya. *Gardens’ Bull.* 13, 254–266.
- Huang, J. L., Sun, G. L., and Zhang, D. M. (2010). Molecular evolution and phylogeny of the angiosperm *ycf2* gene. *J. Syst. Evol.* 48, 240–248. doi: 10.1111/j.1759-6831.2010.00080.x
- IUCN (2024). *The IUCN red list of threatened species. Version 2024-2*. Available online at: <https://www.iucnredlist.org> (Accessed 28 Dec 2024).

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher’s note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2025.1569683/full#supplementary-material>

- Jansen, R. K., Cai, Z., Raubeson, L. A., Daniell, H., Depamphilis, C. W., Leebens-Mack, J., et al. (2007). Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19369–19374. doi: 10.1073/pnas.0709121104
- Jansen, R. K., and Ruhlman, T. A. (2012). “Plastid genomes of seed Plants,” in *Genomics of Chloroplasts and Mitochondria* (Springer, Dordrecht Advances), 103–126. doi: 10.1007/978-94-007-2920-9_5
- Jiang, D., Cai, X., Gong, M., Xia, M., Xing, H., Dong, S., et al. (2023). Complete chloroplast genomes provide insights into evolution and phylogeny of *Zingiber* (Zingiberaceae). *BMC Genomics* 24, 30. doi: 10.1186/s12864-023-09115-9
- Jin, J. J., Yu, W.-B., Yang, J. B., Song, Y., dePamphilis, C. W., Yi, T.-S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biol.* 21, 241. doi: 10.1186/s13059-020-02154-5
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Kellogg, E. A., and Bennetzen, J. L. (2004). The evolution of nuclear genome structure in seed plants. *Am. J. Bot.* 91, 1709–1725. doi: 10.3732/ajb.91.10.1709
- Kress, W. J., Prince, L. M., Hahn, W. J., and Zimmer, E. A. (2001). Unraveling the evolutionary radiation of the families of the zingiberales using morphological and molecular evidence. *Syst. Biol.* 50, 926–944. doi: 10.1080/106351501753462885
- Kress, W. J., Prince, L. M., and Williams, K. J. (2002). The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data. *Am. J. Bot.* 89, 1682–1696. doi: 10.3732/ajb.89.10.1682
- Lanfear, R., von Haeseler, A., Woodhams, M. D., Schrempf, D., Chernomor, O., Schmidt, H. A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015
- Ley, A. C., and Claßen-Bockhoff, R. (2011). “Ontogenetic and phylogenetic diversification in Marantaceae,” in *Flowers on the Tree of Life* (Cambridge: Cambridge University Press), 236–255.
- Li, B., Liu, T., Ali, A., Xiao, Y., Shan, N., Sun, J., et al. (2022). Complete chloroplast genome sequences of three aroidae species (Araceae): lights into selective pressure, marker development and phylogenetic relationships. *BMC Genomics* 23, 218. doi: 10.1186/s12864-022-08400-3
- Li, X., Yang, J.-b., Wang, H., Song, Y., Corlett, R. T., Yao, X., et al. (2021). Plastid NDH pseudogenization and gene loss in a recently derived lineage from the largest hemiparasitic plant genus *pedicularis* (Orobanchaceae). *Plant Cell Physiol.* 62, 971–984. doi: 10.1093/pcp/pcab074
- Lohse, M., Dreichsel, O., and Bock, R. (2007). OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* 52, 267–274. doi: 10.1007/s00294-007-0161-y
- Mahmudi, O., Sennblad, B., Arvestad, L., Nowick, K., and Lagergren, J. (2015). Gene-pseudogene evolution: a probabilistic approach. *BMC Genomics* 16 Suppl 10, S12. doi: 10.1186/1471-2164-16-S10-S12
- McKain, M. R., Johnson, M. G., Uribe-Convers, S., Eaton, D., and Yang, Y. (2018). Practical considerations for plant phylogenomics. *Appl. Plant Sci.* 6, e1038. doi: 10.1002/aps3.1038
- Merrill, E. D. (1919). *Philippine journal of science* Vol. 15 (Manila: The Bureau of Science of the Government of the Philippine Islands), 230.
- Mohanta, T. K., Mishra, A. K., Khan, A., Hashem, A., Abd Allah, E. F., and Al-Harrasi, A. (2020). Gene loss and evolution of the plastome. *Genes (Basel)* 11, 1133. doi: 10.3390/genes11101133
- Moore, M. J., Bell, C. D., Soltis, P. S., and Soltis, D. E. (2007). Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19363–19368. doi: 10.1073/pnas.0708072104
- Moore, M. J., Soltis, P. S., Bell, C. D., Burleigh, J. G., and Soltis, D. E. (2010). Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4623–4628. doi: 10.1073/pnas.0907801107
- Nanjala, C., Wang, O., Odago, W., Mutinda, E. S., Waswa, E. N., Oulo, M. A., et al. (2022). Plastome structure of 8 *Calanthe* s.l. species (Orchidaceae): comparative genomics, phylogenetic analysis. *BMC Plant Biol.* 22, 387. doi: 10.1186/s12870-022-03736-0
- Neubig, K. M., Whitten, W. M., Carlswald, B. S., Blanco, M. A., Endara, L., Williams, N. H., et al. (2008). Phylogenetic utility of *ycf1* in orchids: a plastid gene more variable than *matK*. *Plant Syst. Evol.* 277, 75–84. doi: 10.1007/s00606-008-0105-0
- Niissalo, M. A., Khew, G. S., Webb, E. L., and Leong-ŠKorničKová, J. (2016). Notes on Singaporean native Zingiberales II: revision of Marantaceae, with a new generic record and notes on naturalised and commonly cultivated exotic species. *Phytotaxa* 289, 201–224. doi: 10.11646/phytotaxa.289.3.1
- Palmer, J. D., Jansen, R. K., Michaels, H. J., Chase, M. W., and Manhart, J. R. (1988). Chloroplast DNA variation and plant phylogeny. *Ann. Mo Bot. Gard* 75, 1180–1206. doi: 10.2307/2399279
- Plunkett, G. M., and Downie, S. R. (2000). Expansion and contraction of the chloroplast inverted repeat in apiaceae subfamily apioideae. *Syst. Bot.* 25, 648–667. doi: 10.2307/2666726
- Powell, W., Machray, G. C., and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* 1, 215–222. doi: 10.1016/1360-1385(96)86898-1
- Powell, W., Morgante, M., Andre, C., McNicol, J., Machray, G., Doyle, J., et al. (1995). Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Curr. Biol.* 5, 1023–1029. doi: 10.1016/S0960-9822(95)00206-5
- Prince, L. M., and Kress, W. J. (2006). Phylogenetic relationships and classification in Marantaceae: insights from plastid DNA sequence data. *Taxon* 55, 281–296. doi: 10.2307/25065578
- Provan, J., Powell, W., and Hollingsworth, P. M. (2001). Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends Ecol. Evol.* 16, 142–147. doi: 10.1016/s0169-5347(00)02097-8
- Qiu, Y. L., Li, L., Wang, B., Xue, J. Y., Hendry, T. A., Li, R. Q., et al. (2010). Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *J. Syst. Evol.* 48, 391–425. doi: 10.1111/j.1759-6831.2010.00097.x
- Qu, X. J., Moore, M. J., Li, D. Z., and Yi, T. S. (2019). PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods* 15, 50. doi: 10.1186/s13007-019-0435-7
- Richard, G. F., Kerrest, A., and Dujon, B. (2008). Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol. Mol. Biol. R.* 72, 686–727. doi: 10.1128/MMBR.00011-08
- Ridley, H. N. (1899). The scitamineae of the malay peninsula. *J. Straits Branch R. Asiatic Soc.* 32, 175–182.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029
- Roxburgh, W. (1820). *Flora indica, or, Descriptions of Indian plants* Vol. 1 (Serampore: Printed at the Mission Press), 1–9.
- Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248
- Ruhlman, T. A., and Jansen, R. K. (2014). The plastid genomes of flowering plants. *Methods Mol. Biol.* 1132, 3–38. doi: 10.1007/978-1-62703-995-6_1
- Schumann, K. (1902). “Marantaceae” in *Das Pflanzenreich*, vol. 4. (Leipzig: Verlag von Wilhelm Engelmann), 1–184.
- Schwarz, E. N., Ruhlman, T. A., Sabir, J. S. M., Hajrah, N. H., Alharbi, N. S., Al-Malki, A. L., et al. (2015). Plastid genome sequences of legumes reveal parallel inversions and multiple losses of *rps16* in papilionoids. *J. Syst. Evol.* 53, 458–468. doi: 10.1111/jse.12179
- Sharp, P. M., and Li, W. H. (1987). The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15, 1281–1295. doi: 10.1093/nar/15.3.1281
- Sheng, J., She, X., Liu, X., Wang, J., and Hu, Z. (2021). Comparative analysis of codon usage patterns in chloroplast genomes of five *Miscanthus* species and related species. *PeerJ* 9, e12173. doi: 10.7717/peerj.12173
- Smith, J. F., Kress, W. J., and Zimmer, E. A. (1993). Phylogenetic analysis of the zingiberales based on *rbcL* sequences. *Ann. Mo Bot. Gard* 80, 620–630. doi: 10.2307/2399850
- State Administration of Traditional Chinese Medicine and Chinese Materia Medica Editorial Committee (1999). *Zhonghua Bencao (The Chinese Herbal)* Vol. 8 (Shanghai: Shanghai Scientific and Technical Publishers), 670–671.
- Steenwyk, J. L., Li, Y., Zhou, X., Shen, X. X., and Rokas, A. (2023). Incongruence in the phylogenomics era. *Nat. Rev. Genet.* 24, 834–850. doi: 10.1038/s41576-023-00620-x
- Suksathan, P., Gustafsson, M. H., and Borchsenius, F. (2009). Phylogeny and generic delimitation of Asian Marantaceae. *Bot. J. Linn Soc.* 159, 381–395. doi: 10.1111/j.1095-8339.2009.00949.x
- Suksthan, P., Madulid, D., and Borchsenius, F. (2010). Marantaceae in the Philippines. *Taiwania* 55, 28–36. doi: 10.6165/ta.2010.55(1).28
- The Angiosperm Phylogeny Group (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn Soc.* 181, 1–20. doi: 10.1111/boj.12385
- Wang, R. J., Cheng, C. L., Chang, C. C., Wu, C. L., Su, T. M., and Chaw, S. M. (2008). Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol. Biol.* 8, 36. doi: 10.1186/1471-2148-8-36
- Whitten, W. M., Neubig, K. M., and Williams, N. H. (2014). Generic and Subtribal relationships in neotropical cymbidieae (orchidaceae) based on *matK/ycf1* plastid data. *Lankester* 13, 375–392. doi: 10.15517/lanlv3i13.14425

- Willdenow, C. L. V. (1797). "Species Plantarum," in *Editio quarta*, vol. 1. Ed. G. C. Impensis. (Berlin: impensis Nauk), 495.
- Wu, D. (1991). "Marantaceae," in *Flora of Guangdong*. Eds. F. H. Chen and D. L. Wu (Guangdong: Guangdong Science and Technology Press), 437–441.
- Wu, D., and Helen, K. (2000). "Marantaceae," in *Flora of China*, vol. 24. Eds. Z. Y. Wu, P. H. Raven and D. Y. Hong (Science Press; St. Louis: Missouri Botanical Garden, Beijing), 379–382.
- Xu, S., Teng, K., Zhang, H., Gao, K., Wu, J., Duan, L., et al. (2023). Chloroplast genomes of four *Carex* species: Long repetitive sequences trigger dramatic changes in chloroplast genome structure. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1100876
- Zhang, D., Gao, F., Jakovlic, I., Zou, H., Zhang, J., Li, W. X., et al. (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour* 20, 348–355. doi: 10.1111/1755-0998.13096
- Zhou, M., Long, W., and Li, X. (2008). Patterns of synonymous codon usage bias in chloroplast genomes of seed plants. *For Stud. China* 10, 235–242. doi: 10.1007/s11632-008-0047-1
- Zhu, A., Guo, W., Gupta, S., Fan, W., and Mower, J. P. (2016). Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytol.* 209, 1747–1756. doi: 10.1111/nph.13743