MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 8, 1000–1004
https://doi.org/10.1080/23802359.2024.2387251

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

The complete chloroplast genome of *Flemingia stricta* Roxb. ex Ait. 1812 (Phaseoleae, Fabaceae)

Yisha Cai[a,b], Haimei Chen[b] and Jinghong Zhang[a]

[a]School of Medicine, Huaqiao University, Fujian, China; [b]State Key Laboratory for Quality Ensurance and Sustainable Use of Dao-Di Herbs, Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education of China, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

ABSTRACT

Flemingia stricta Roxb. ex Ait. 1812 belongs to the Phaseoleae tribe within the Fabaceae family and has significant pharmaceutical value. In this study, we reported the complete chloroplast genome of *F. stricta* using the Illumina DNA sequencing data. The chloroplast genome was 152,940 bp and encoded 111 unique genes, including 77 protein-coding genes (PCGs), 30 transfer RNA (tRNA) genes, and 4 ribosomal RNA (rRNA) genes. The phylogenetic analysis confirmed that *F. stricta* was closely related to *Flemingia prostrata* and *Flemingia macrophylla*. The chloroplast genome of *F. stricta* could provide critical information for the molecular breeding of *F. stricta* and be used as a reference genome for other species of Phaseoleae.

ARTICLE HISTORY
Received 18 March 2024
Accepted 26 July 2024

KEYWORDS
chloroplast genome; Phaseoleae; *Flemingia stricta*; phylogenetic analysis

Introduction

Flemingia stricta Roxb. ex Ait. 1812, an erect or spreading subshrub belonging to the Phaseoleae tribe within the Fabaceae family, is widely distributed in tropical areas of Asia, Africa, and Oceania (Li 2007). As a traditional Chinese medicine, it has been used to treat rheumatism, arthropathy, chronic nephritis, and some gynecopathy in folk (Wu 2005). There are various secondary metabolites in *F. stricta*, among which flavonoid and triterpenes compounds have neuroprotection, anti-inflammation, anti-oxidation, cytotoxicity, hormone-like effects, and antimicrobial activities in modern medical fields (Li et al. 2012). High-throughput sequencing technology greatly facilitates the study of plant genome structure, which is helpful for the exploration of plant secondary metabolites and plant genetic variation at the molecular level. Previous studies have yielded multiple research endeavors focused on the secondary metabolites of *F. stricta*. However, the chloroplast genome sequence of *F. stricta* remains unreported, severely restricting follow-up research.

The chloroplast genomes play an essential role in plants. At first, the chloroplast is a vital organelle in plant cells, containing thylakoid structures and serving as the primary site of photosynthesis (Heber and Heldt 1981; Jensen and Leister 2014). Secondly, as a high conservation of plant chloroplast, it has been widely used as a molecular marker in modern molecular biology research (Li et al. 2020; Yang et al. 2021) and in reconstructing phylogenetic relationships of species (Shen et al. 2022; Fan and Ma 2022). Lastly, the plant chloroplast genome is typically a circular double-stranded DNA molecule with a relatively conserved genetic structure, involved in photosynthesis, energy production, maintenance of cellular structure, and the synthesis of biological molecules, which are crucial for the survival, growth, and reproduction of plants (Sabater 2018). Therefore, it is necessary to explore the chloroplast genome of *F. stricta*. In this study, we sequenced and assembled the complete chloroplast genome of *F. stricta,* deciphered the possible structure of *F. stricta* chloroplast genome and performed phylogenetic relationships analysis. The results can help people better explore scientific problems at the molecular level, laying an important theoretical foundation for genetic engineering and genetic breeding in medicinal plants.

Materials and methods

Fresh leaves of *F. stricta* were collected from the Yunnan Branch, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Jinghong, Yunnan Province, China (latitude 220 N, longitude 100.47 E) (Figure 1). The specimen was deposited

CONTACT Haimei Chen hmchen@implad.ac.cn State Key Laboratory for Quality Ensurance and Sustainable Use of Dao-Di Herbs, Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education of China, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China; Jinghong Zhang zjh@hqu.edu.cn School of Medicine, Huaqiao University, Fujian, 362021, China.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2387251.

Figure 1. *F. stricta* was collected from the Yunnan Branch, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Jinghong, Yunnan Province, China. (A) The panorama of *F. stricta*: Shrubs, tall, stems robust. (B) The leaves of *F. stricta*: stipules lanceolate, leaf-lets oblong or lanceolate to obliquely ovate-lanceolate, distinctly veined, puberulent.

at the Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing under the voucher number Implad201911327. The total DNA was extracted using the Plant Genomic DNA Kit (Tiangen Biotech, Beijing, Co., Ltd., China), and then stored in a refrigerator at −80 °C. The total DNA was used for Illumina sequencing. Among Illumina sequencing, the DNA library was constructed from 1 μg genomic DNA by using a NEBNext DNA library prep kit and sequenced with a Hiseq 2500 platform (Illumina, San Diego, CA, USA).

A total of 7.17 G clean data was obtained. Table S1 presents comprehensive details regarding the sequencing data. At first, we utilized the GetOrganelle software to assemble the chloroplast genome with the parameters "-R 15 -k 21,45,65,85,105 -F embplant_pt". Subsequently, we employed the Bandage software to adjust it into a circular chloroplast genome. The chloroplast genome was annotated by CPGAVAS2 (Shi et al. 2019), and the annotation results were manually corrected by using Apollo software (Lewis et al. 2002). Then, the structure of the chloroplast genome was plotted using CPGView webserver (Liu et al. 2023). Finally, the chloroplast genome sequences and annotations were submitted to GenBank with accession numbers PP230165.1. Using genes from related species as queries, we searched the chloroplast genome of each target species by BLASTN to further determine if the gene is truly missing.

We used 15 chloroplast genomes of Phaseoleae species for phylogenetic analysis and selected *Lotus japonicus* (NC_002694.1) (Jiang et al. 2021) as the outgroup. All chloroplast genomes were downloaded from the NCBI GenBank database. For phylogenetic analysis, firstly, the complete chloroplast genomes of 16 species were aligned by MAFFT software (Katoh and Standley 2013) with default parameters. The poorly aligned positions and divergent regions of the alignment were eliminated using Gblocks (Talavera and Castresana 2007). Then, the aligned sequences were used to construct the phylogenetic tree using IQTREE2 (Minh et al. 2020) with the maximum-likelihood method and best-fit model TVM + F+R3. And the bootstrap analysis was assessed using UFBoot with 1,000 replicates (Hoang et al. 2017). Finally, the resulting tree was visualized by iTOL (Letunic and Bork 2021).

Results

The chloroplast genome *F. stricta* consists of 1 circular chromosome which is 152,940 bp in length. And the average depth of coverage is over 500 (Figure S1). The GC content of the chromosome is 35.08% (Figure 2). It is closely related in size to the chloroplast genome of *Flemingia macrophylla* chloroplast genome (Genome size: 152,937 bp, NCBI Accession Number: NC_065865.1). The chloroplast genome of *F. stricta* annotated a total of 129 genes (111 unique genes), including 84 protein-coding genes (77 unique genes), 37 transfer RNA (tRNA) genes (30 unique genes), and 8 ribosomal RNA (rRNA) genes (4 unique genes). These protein-coding genes were categorized into 15 functional groups (Table 1). Among them, a total of 11 unique protein-coding genes (PCGs) (*rps*16, *atp*F, *rpo*C1, *pet*B, *pet*D, *rpl*16, *rpl*2, *ndh*B, *ndh*A, *clp*P) contain one intron and one PCG contains two introns (*ycf*3) (Figure S2). Additionally, *rps*12 is a trans-spliced gene as shown in Figure S3.

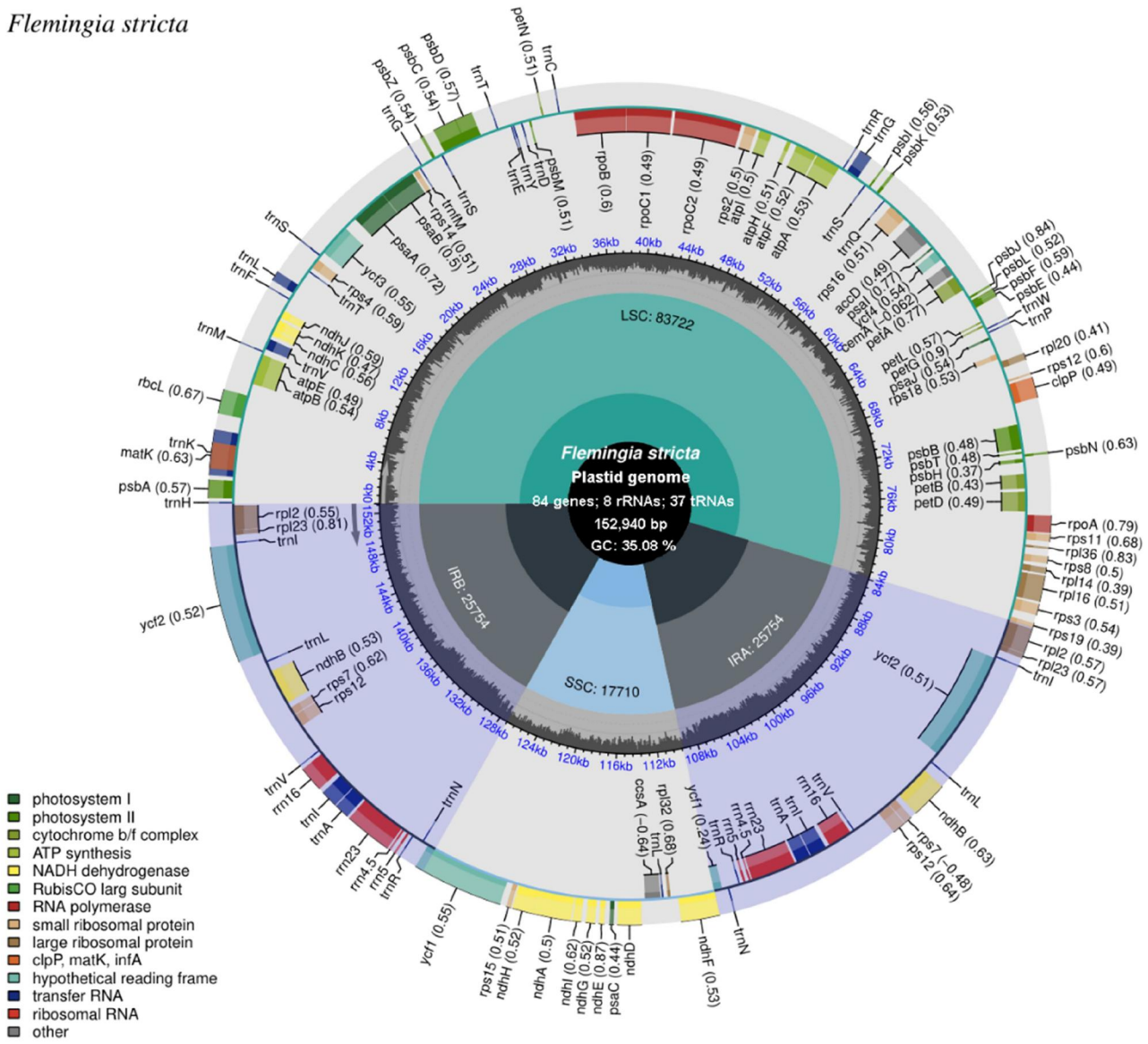

Figure 2. A schematic representation of the *F. stricta* chloroplast genome. The graph was drawn using CPGview (http://www.1kmpg.cn/cpgview). The graph is represented from inside out: (1) the green circle: LSC (large single-copy), the blue circle SSC (small single-copy), the light grey circle: IRA and IRB (inverted repeat); (2) the distribution of GC content on the chromosome; (3) the scale coordinate axis; (4) genes located on the negative strand and positive strands. And the numbers after the gene names indicate GC content of each gene.

Table 1. Genes predicted in the chloroplast genome of *F. stricta*.

Group of genes	Name of genes
Subunits of ATP synthase	*atpA, atpB, atpE, atpF, atpH, atpI*
Subunits of photosystem II	*psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ, ycf3*
Subunits of NADH-dehydrogenase	*ndhA, ndhB*(×2), *ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK*
Subunits of cytochrome b/f complex	*petA, petB, petD, petG, petL, petN*
Subunits of photosystem I	*psaA, psaB, psaC, psaI, psaJ*
Subunit of rubisco	*rbcL*
Large subunit of ribosome	*rpl14, rpl16, rpl2*(×2), *rpl20, rpl23*(×2), *rpl32, rpl36*
DNA-dependent RNA polymerase	*rpoA, rpoB, rpoC1, rpoC2*
Small subunit of ribosome	*rps11, rps12*(×2), *rps14, rps15, rps16, rps18, rps19, rps2, rps3, rps4, rps7*(×2), *rps8*
Subunit of Acetyl-CoA-carboxylase	*accD*
c-type cytochrome synthesis gene	*ccsA*
Envelop membrane protein	*cemA*
Protease	*clpP*
Maturase	*matK*
Conserved open reading frames	*ycf1*(×2), *ycf2*(×2), *ycf4*
tRNA genes	37 tRNA genes (8 contains 1 intron)
rRNA genes	*rrn16S*(×2), *rrn23S*(×2), *rrn5S*(×2), *rrn4.5S*(×2)

Note: (×2) indicates two copies of the gene.

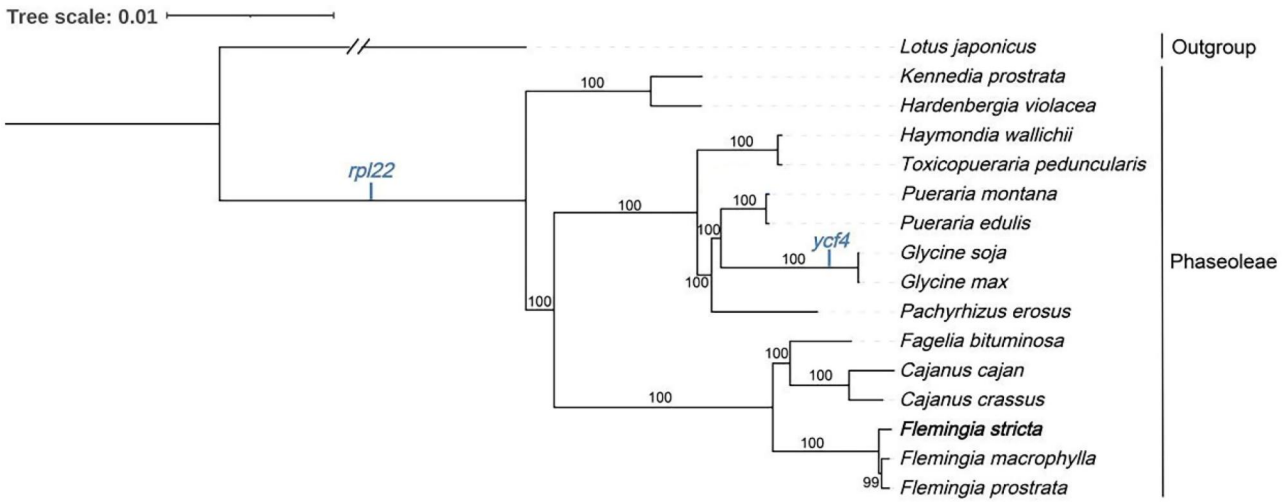

Figure 3. Phylogenetic relationships of *F. stricta* and other 14 species. The numbers indicate the bootstrap values for the maximum likelihood (ML) tree. The number on each branch node represents the bootstrap number. The tribe is labeled next to the species' name. The bold font indicates that all 15 species belong to Phaseoleae. The outgroup belongs to the Cladrastis clade. The following sequences were used: *Lotus japonicus* (NC_002694.1) (Jiang et al. 2021), *Glycine max* (NC_007942.1) (Jin et al. 2019), *Pueraria Montana* (OP963928.1), *Pueraria edulis* (NC_065692.1), *Flemingia stricta* (PP230165.1), *Flemingia prostrata* (NC_065863.1), *Flemingia macrophylla* (NC_065865.1) (Qin et al. 2021), *Glycine soja* (NC_022868.1) (Gao and Gao 2017), *Pachyrhizus erosus* (NC_026682.1), *Toxicopueraria peduncularis* (OR491709.1), *Haymondia wallichii* (NC_052910.1) (Huang et al. 2021), *Cajanus cajan* (KU729879.1) (Kaila et al. 2016), *Cajanus crassus* (NC_057277.1) (Oyebanji et al. 2020), *Fagelia bituminosa* (NC_057451.1) (Oyebanji et al. 2020), *Hardenbergia violacea* (NC_057453.1) (Oyebanji et al. 2020), *Kennedia prostrata* (NC_057454.1) (Oyebanji et al. 2020). For the gene loss in the chloroplast genomes, *rpl22* was not found in all 16 Phaseoleae. The *ycf4* gene was absent in the chloroplast genome of two Glycine species.

To determine the phylogenetic relationship of 15 Phaseoleae species, we constructed the phylogenetic trees using the complete chloroplast genomes of 16 species. The phylogenetic tree indicated that all chloroplast genomes of Phaseoleae species have lost the *rpl22* gene. The chloroplast genomes of *Glycine max* and *Glycine soja* have experienced a loss of the *ycf4* gene. As the *rpl33* gene in the chloroplast genomes of *Flemingia* is a pseudogene, its absence is regarded as a typical phenomenon. The phylogenetic tree based on the complete chloroplast genome sequence indicates that *F. stricta*, along with other *Flemingia* species and the closely related *Cajanus* and *Fagelia* species, forms a distinct branch. Furthermore, the results show that *Flemingia prostrata* is more closely related to *Flemingia macrophylla* than to *F. stricta* (Figure 3). Thus, *F. stricta* can serve as a reference for exploring the chloroplast genome of other species in the *Flemingia* genus. Moreover, all nodes had bootstrap support values $>$ 95, indicating the strong reliability of the phylogenetic relationship of the 15 chloroplast genomes of Phaseoleae species.

Discussion and conclusions

Plant chloroplast genomes are usually assembled and displayed as circular maps based on the widely-held view across the broad community of life scientists that circular genome-sized molecules are the primary form of plant chloroplast DNA (Kozik et al. 2019). In this study, we assembled and annotated the chloroplast genome of *F. stricta* for the first time. The chloroplast genome of *F. stricta* also contains one circular molecular. And we revealed the general feature of *F. stricta* chloroplast genome and conducted phylogenetic analysis. Providing the genetic reference for *F. stricta*is crucial as the species is one of the genera *Flemingia*. *F. stricta*, as one of the traditional Chinese medicines, its roots, leaves or whole plants play an important role in the treatment of rheumatism, arthropathy, chronic nephritis, and so on (Wu 2005). Due to the low natural production of flavonoids and triterpenoids, which are secondary metabolites of plants, the development of new drugs is severely limited (Jiang et al. 2021). Thus, the research on genomes of *F. stricta* lays a very important theoretical foundation for improving its secondary metabolite content. This study reported a complete chloroplast genome and found that 15 species of Phaseoleae have lost the *rpl22* gene, which may provide a reference for the evolutionary relationship of chloroplast in Phaseoleae species, molecular breeding, genetic diversity and molecular-based species identification.

Ethical approval

The sample of *Flemingia stricta* was collected with permission from Fresh leaves of F. stricta were collected from the Yunnan Branch, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Jinghong, Yunnan Province, China, and strictly complied with local and Chinese regulations. No ethical approval is required in this study.

Author contributions

CHM and CYS made substantial contributions to the design of this work and final approval of the version to be published; CYS analyzed the data and drafted this work; CHM and ZJH reviewed this work critically. All authors agreed to be accountable for all aspects of this work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [81872966], CAMS Innovation Fund for Medical Sciences (CIFMS) (2021-I2M-1-071, 2021-I2M-1-022), the National Science & Technology Fundamental Resources Investigation Program of China [2018FY100705], and Provincial Special Project for Construction of Innovation Demonstration Area at Chenzhou City under National Sustainable Development Plan (2023sfq04).

ORCID

Jinghong Zhang http://orcid.org/0009-0006-5900-4286

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at (https://www.ncbi.nlm.nih.gov/) under the accession number PP230165.1. The associated BioProject, BioSample, and SRA numbers are PRJNA1069630, SAMN39617629 and SRR27766012.

References

Fan ZF, Ma CL. 2022. Comparative chloroplast genome and phylogenetic analyses of Chinese Polyspora. Sci Rep. 12(1):15984. doi:10.1038/s41598-022-16290-4.

Gao C-W, Gao L-Z. 2017. The complete chloroplast genome sequence of wild soybean, *Glycine soja*. Conservation Genet Resour. 9(2):329–331. doi:10.1007/s12686-016-0659-z.

Heber U, Heldt H. 1981. The chloroplast envelope: structure, function, and role in leaf metabolism. Annu Rev Plant Physiol. 32(1):139–168. doi:10.1146/annurev.pp.32.060181.001035.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2017. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 35(2):518–522. doi:10.1093/molbev/msx281.

Huang X, Xia C, Zhao Y, Xu L. 2021. The first complete chloroplast genome of *Haymondia wallichii* (Fabaceae) and its phylogenetic analysis. Mitochondrial DNA B Resour. 6(7):1865–1866. doi:10.1080/23802359.2021.1923415.

Jensen PE, Leister D. 2014. Chloroplast evolution, structure and functions. F1000Prime Rep. 6:40. doi:10.12703/P6-40.

Jiang F, Gong T, Chen J, Chen T, Yang J, Zhu P. 2021. Synthetic biology of plants-derived medicinal natural products. Sheng Wu Gong Cheng Xue Bao. 37(6):1931–1951. doi:10.13345/j.cjb.210138.

Jiang W, Ma L, Yi D, Pang Y. 2021. The complete chloroplast genome sequence of *Lotus corniculatus* L. Mitochondrial DNA B Resour. 6(1):189–190. doi:10.1080/23802359.2020.1860708.

Jin L, Wan Z, Wei D, Yin D. 2019. Characterization of the complete chloroplast genome of black soybeans, *Glycine max* (L.) Merr. (legume). Mitochondrial DNA B Resour. 4(2):3309–3311. doi:10.1080/23802359.2019.1673227.

Kaila T, Chaduvla PK, Saxena S, Bahadur K, Gahukar SJ, Chaudhury A, Sharma TR, Singh NK, Gaikwad K. 2016. Chloroplast genome sequence of Pigeonpea (*Cajanus cajan* (L.) Millspaugh) and *Cajanus scarabaeoides* (L.) Thouars: genome organization and comparison with other legumes. Front Plant Sci. 7:1847. doi:10.3389/fpls.2016.01847.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780. doi:10.1093/molbev/mst010.

Kozik A, Rowan BA, Lavelle D, Berke L, Schranz ME, Michelmore RW, Christensen AC. 2019. The alternative reality of plant mitochondrial DNA: one ring does not rule them all. PLoS Genet. 15(8):e1008373. doi:10.1371/journal.pgen.1008373.

Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49(W1):W293–W296. doi:10.1093/nar/gkab301.

Lewis SE, Searle SMJ, Harris N, Gibson M, Lyer V, Richter J, Wiel C, Bayraktaroglu L, Birney E, Crosby MA, et al. 2002. Apollo: a sequence annotation editor. Genome Biol. 3(12):Research0082. doi:10.1186/gb-2002-3-12-research0082.

Li C, Zheng Y, Huang P. 2020. Molecular markers from the chloroplast genome of rose provide a complementary tool for variety discrimination and profiling. Sci Rep. 10(1):12188. doi:10.1038/s41598-020-68092-1.

Li H, Zhai F, Liu Z. 2012. Chemical constituents and bioactivities of the plants of genus Flemingia Roxb. et Ait. (Leguminosae). Comb Chem High Throughput Screen. 15(8):611–622. doi:10.2174/138620712802650504.

Li J. 2007. Flora of China. Harvard Papers Bot. 13(2):301–302. doi:10.3100/1043-4534-13.2.301.

Liu S, Ni Y, Li J, Zhang X, Yang H, Chen H, Liu C. 2023. CPGView: a package for visualizing detailed chloroplast genome structures. Mol Ecol Resour. 23(3):694–704. doi:10.1111/1755-0998.13729.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 37(5):1530–1534. doi:10.1093/molbev/msaa015.

Oyebanji O, Zhang R, Chen S-Y, Yi T-S. 2020. New insights into the plastome evolution of the Millettioid/Phaseoloid clade (Papilionoideae, Leguminosae). Front Plant Sci. 11:151. doi:10.3389/fpls.2020.00151.

Qin X-M, Li H, Cui N, Jiang S-Y, Liang Y-N, Wang M-L, Huang X-Y. 2021. The complete chloroplast genome of *Flemingia macrophylla* (Willd.) Prain (Fabaceae) from Guangxi, China. Mitochondrial DNA B Resour. 6(12):3378–3380. doi:10.1080/23802359.2021.1997124.

Sabater B. 2018. Evolution and function of the chloroplast. Current investigations and perspectives. Int J Mol Sci. 19(10):3095. doi:10.3390/ijms19103095.

Shen J, Li X, Chen X, Huang X, Jin S. 2022. The complete chloroplast genome of *Carya cathayensis* and phylogenetic Analysis. Genes (Basel). 13(2):369. doi:10.3390/genes13020369.

Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 47(W1):W65–w73. doi:10.1093/nar/gkz345.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 56(4):564–577. doi:10.1080/10635150701472164.

Wu J-N. 2005. An illustrated Chinese materia medica. New York, NY: Oxford University Press.

Yang H, Wang L, Chen H, Jiang M, Wu W, Liu S, Wang J, Liu C. 2021. Phylogenetic analysis and development of molecular markers for five medicinal Alpinia species based on complete plastome sequences. BMC Plant Biol. 21(1):431. doi:10.1186/s12870-021-03204-1.