MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 8, 1072–1076
https://doi.org/10.1080/23802359.2024.2391089

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

The complete chloroplast genome of *Swertia davidii* (gentianaceae) and its phylogenetic analysis

Zhishi Zhang[a]*, Yongbiao Deng[b]*, Lan Wu[c], Haoren Zhu[d], Yulong Song[a], Aotian He[a], Jian Dai[a], Jinhui Qin[a], Lihuan Luo[a] and Chao Xiong[a]

[a]School of Life Science and Technology, Wuhan Polytechnic University, Wuhan, China; [b]Department of Pharmacognosy, Pharmacy School, Guilin Medical University, Guilin, China; [c]Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China; [d]School of Computer and Cyber Science, Communication University of China, Beijing, China

ABSTRACT

To elucidate the genetic information and evolutionary relationships of *Swertia*, we initiated the sequencing of the complete chloroplast genome of *Swertia davidii* Franch. 1888, complemented by comparative analyses with closely related species. The chloroplast genome of *S. davidii* was 153,516 bp in length and exhibited a typical quadripartite structure. It contained two regions with Inverted Repeat lengths of 25,767 bp, located between one Large Single-Copy region (83,617 bp) and one Short Single-Copy region (18,365 bp). The chloroplast genome of *S. davidii* encoded 132 genes, including 87 protein-coding genes, 37 transfer RNA genes, and 8 ribosomal RNA genes. The overall GC content was 38.15%. Maximum likelihood phylogenetic analysis of *Swertia* based on 26 available plastomes showed a close relationship between *S. davidii* and *S. kouitchensi*. This study will contribute to the genetic preservation of the species and the phylogenetic study of *Swertia*.

Abbreviations: cp: chloroplast; LSC: large single copy; SSC: small single copy; IR: inverted repeat; ML: maximum likelihood; rRNA: ribosomal RNA; tRNA: transfer RNA; SSR: simple sequence repeat; PCG: protein-coding gene.

ARTICLE HISTORY
Received 27 November 2023
Accepted 6 August 2024

KEYWORDS
Swertia; plastome; phylogeny; systematic evolution

Introduction

Swertia davidii Franch. 1888, commonly referred to as "Shuilingzhi or Yudancao", belongs to the Gentianaceae family. This perennial herb is predominantly distributed across the Chinese provinces of Yunnan, eastern Sichuan, western Hubei, and northern Hunan. Since *S. davidii* exhibits a variety of pharmacological effects, such as analgesic, anti-inflammatory, anti-tumor, and bacteriostatic properties (Hunan Academy of Chinese Medicine 1979; Vaidya et al. 2013), it holds significance as an herbal medicine in the central region of China (Flora of China Editorial Committee 1988; Li et al. 2002). In recent years, the pharmacologically active phytochemicals of *S. davidii* have been identified. The whole plant of *S. davidii* yielded structurally diverse compounds, including uvaol, oleanolic acid, isoscoparin, gentiopicroside, swertiamarin, gentiannine, and mangiferin (Zeng et al. 2013; Liang et al. 2014). In addition to ethnomedicinal and phytochemical research, several studies have also described the taxonomy, genetic diversity, germplasm resource conservation, and development. However, few studies have focused on the development of species-specific molecular markers for *S. davidii*, which we believe to be an important tool to assist in the germplasm resource exploitation of this species. An available complete chloroplast (cp) genome of *S. davidii* will contribute to the progress of this work.

With the advent of next-generation sequencing, sequencing the whole cp genome has become more feasible. In the past decade, cp genome has been extensively used in phylogeny, adaptive evolution, genetic diversity, and species identification across various plant families (Wang et al. 2022). In this study, we sequenced, assembled, and analyzed the complete cp genome of *S. davidii*. This will help to elucidate the phylogenetic relationship of this species in *Swertia* and to exploit species-specific molecular markers for germplasm resource conservation and species identification.

Materials and methods

One individual plant of *S. davidii* was collected from Zhangjiajie, Hunan Province, China (29°17′08″N, 110°09′30″E) (Figure 1). The species identity has been authenticated by Dr. Lan Wu, an associate researcher affiliated to the Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences. Voucher specimens (HB08CD01) were deposited at

CONTACT Chao Xiong xiongchao080190@126.com School of Life Science and Technology, Wuhan Polytechnic University, Wuhan, China
*Both authors contributed equally to this work.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2391089.

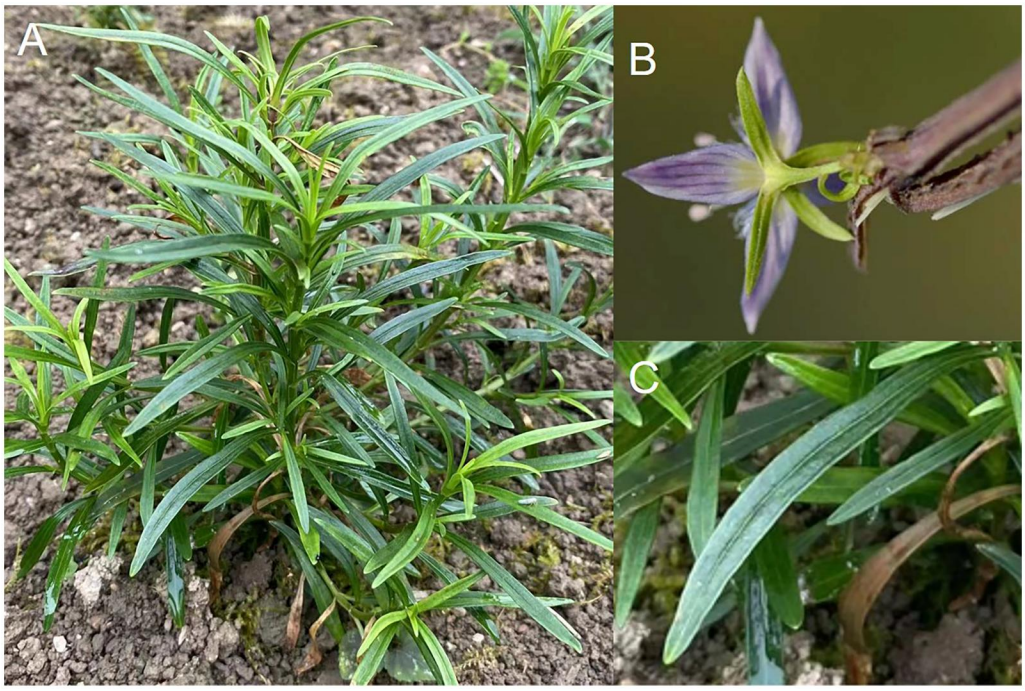

Figure 1. Morphological characteristics of *S. davidii*. (A) *S. davidii* in its natural habitat, (B) flower, (C) leaf. The photographs were taken by corresponding author Chao Xiong in Zhangjiajie, Hunan Province, China (29°17′08″N, 110°09′30″E). *S. davidii* is characterized by perennial herbaceous growth, featuring erect, quadrilateral stems, narrowly winged ribs, and obliquely ascending slanted branches. Its basal and lower stem leaves possess long stalks, while the leaf blade is narrowly elliptical in shape.

Wuhan Polytechnic University (the designated contact person: Chao Xiong, email: xiongchao080190@126.com). Genomic DNA extraction was performed using the Plant Genomic DNA Kit (Tiangen, Co. Ltd., China).

The DNA library with an insert size of 300 bp was constructed for *S. davidii* using the AIR™ Paired-End DNA Sequencing Kit (Bioscientific). Subsequent sequencing was performed utilizing the Illumina NovaSeq6000 platform (Illumina, United States), generating paired-end reads with an average length of 150 bp, totaling 6 GB of raw data. Initially, NOVOPlasty v2.7.2 was employed to assemble the cp genome with a k-mer length of 19 bp (Ma et al. 2020), using the cp genome sequence of *S. leducii* (NC_045301) as reference. Subsequently, the online annotation tool CPGAVAS2 was employed to annotate the *S. davidii* cp genome (Shi et al. 2019; Yang et al. 2022). The start and stop codon positions of the protein-coding genes were manually adjusted based on the annotation of the *S. leducii* cp genome. Thirdly, CPGView was employed to generate a schematic map of the cis-splicing genes and trans-splicing genes, along with the circular genomic map for the *S. davidii* cp genome (Liu et al. 2023). The annotated sequence of the *S. davidii* cp genome has been submitted to the GenBank database under the accession number OR142207.

The complete cp genomes of 26 *Swertia* species were aligned using MAFFT v7.222 to investigate the phylogenetic relationship of *S. davidii* within the *Swertia* genus (Kazutaka and Standley 2013). To construct a comprehensive phylogenetic tree, *Centaurium erythraea* and *Gentianopsis grandis* were chosen as outgroups. The maximum-likelihood (ML) tree was constructed using RAxML v7.0.4 with the GTR + CAT model and 1000 bootstrap replicates (Stamatakis 2006).

Results

For the assembled genomes, the minimum and average read mapping depths were 584×, and 4704×, respectively (Supplementary Figure 1). The chloroplast genome length of *S. davidii* is 153,516 bp and has 132 genes in total, comprising 37 tRNA (transfer RNA) genes, 8 rRNA (ribosomal RNA) genes, and 87 protein-coding genes (PCGs). 13 of the 132 genes (*ndh*A, *ndh*B, *pet*B, *pet*D, *atp*F, *rpl*14, *rpl*16, *rpl*2, *trn*A-UGC, *trn*G-UCC, *trn*I-GAU, *trn*K-UUU, and *trn*L-UAA) have one intron, while 3 (*rps*12, *clp*P, and *ycf*3) have two introns. The complete chloroplast of *S. davidii* also contains two Inverted Repeats (IRs) of 25,767 bp, a Large Single-Copy (LSC) region of 83,617 bp and a Small Single-Copy (SSC) region of 18,365 bp (Figure 2). Additionally, five protein-coding genes (*ndh*B, *rps*7, *rps*12, *ycf*2, and *ycf*15) and five tRNAs (*trn*A-UGC, *trn*I-CAU, *trn*N-GUU, *trn*R-ACG, and *trn*V-GAC) are duplicated in the IR regions. Among these, twelve genes (*atp*F, *rpo*C1, *ycf*3, *clp*P, *pet*B, *pet*D, *rpl*16, *rpl*2, *ndh*B, and *ndh*A) are cis-splicing genes, *rpl*2 and *ndh*B were duplicates. Ten of these genes have one intron and two exons, while two have two introns and three exons. The *rps*12 gene is also a trans-splicing gene (Supplementary Figure 2). The overall GC content in the chloroplast is 38.15%, with the GC contents of the LSC, SSC, and IR regions being 36.24%, 31.9%, and 43.49%, respectively.

In the phylogenetic analysis, 84 protein-coding sequences from the cp genomes of 26 *Swertia* species were utilized to construct a phylogenetic tree through ML analysis. (Figure 3). The analysis revealed that *S. davidii* belongs to the Gentianaceae family and is closely related to *S. kouitchensis*.

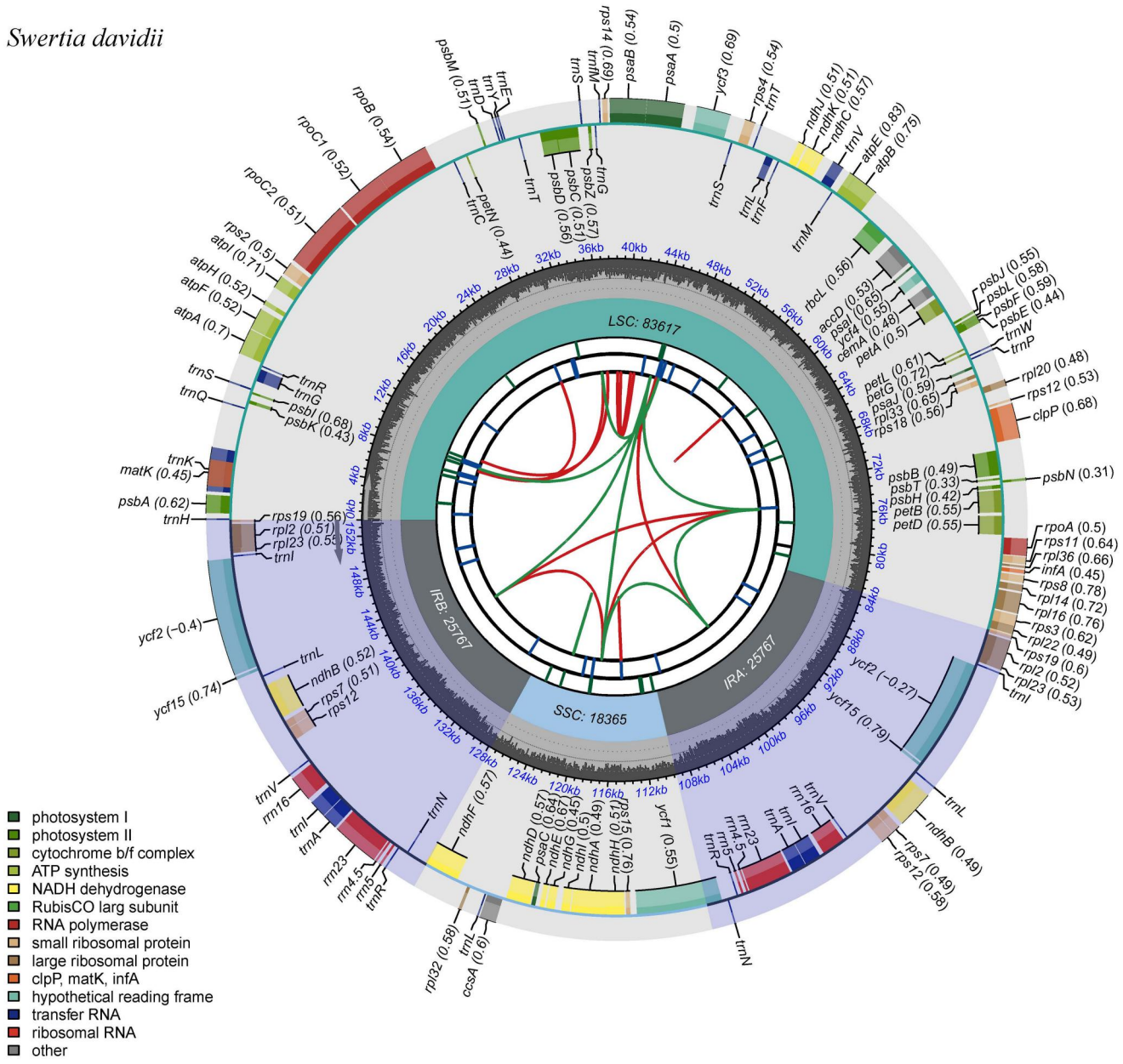

Figure 2. A circular map of the whole chloroplast genome of *S. davidii*. Six tracks total on the map illustrate different aspects of the genome. Starting from the center of the map, the first track shows scattered, which consists of both straight and palindromic motifs, shown by red and green arcs, in that order. Short blue bars on the following track indicate long tandem repeats. Short tandem repeats or microsatellite sequences are represented as colored bars on the third track; each color represents a different type of repeat. Every color has a corresponding type of repeat, and the descriptions that go with each color give important details about the traits of each type of repeat. The following are the colors and the repetition types that correspond to them: black: c (complex repeat); green: p1 (repeat unit size = 1); yellow: p2 (repeat unit size = 2); purple: p3 (repeat unit size = 3); blue: p4 (repeat unit size = 4); orange: p5 (repeat unit size = 5); red: p6 (repeat unit size = 6). The fourth track presents the LSC, SSC, and two IR sections that make up the architecture of the chloroplast genome. The GC content across the genome is depicted in the fifth track. The genome carefully color-codes genes according to their functional classification. Inner genes are shown as having a transcription orientation that is clockwise, whereas outside genes are portrayed as having an anticlockwise orientation. A key that clarifies gene functional classification is provided in the bottom left corner of the visualization to aid in interpretation.

Discussions and conclusions

This study presents the first comprehensive report on the sequence and detailed features of the whole plastome of *S. davidii*. The cp genome of *S. davidii* spans 153,516 bp and contains 87 PCGs, 37 tRNA genes, and 8 rRNA genes. The size, organization, and gene content of this cp genome show similarities to those of other *Swertia* species. However, minor but notable differences were observed in the genome length and intergenic region sequences. Phylogenetic analyses revealed a close relationship between *S. davidii* and *S. kouitchensis*, clustering them into a relatively newly evolved clade. Additionally, all *Swertia* species collectively form a larger clade within the phylogenetic tree. Notably, all nodes received 100% bootstrap support in the ML analyses. This finding clearly emphasizes the significant role of cp genome in resolving the taxonomic uncertainties among *Swertia* species. Additionally, the newly sequenced cp genome will facilitate subsequent research efforts in areas such as species identification and germplasm conservation of the *Swertia* species.

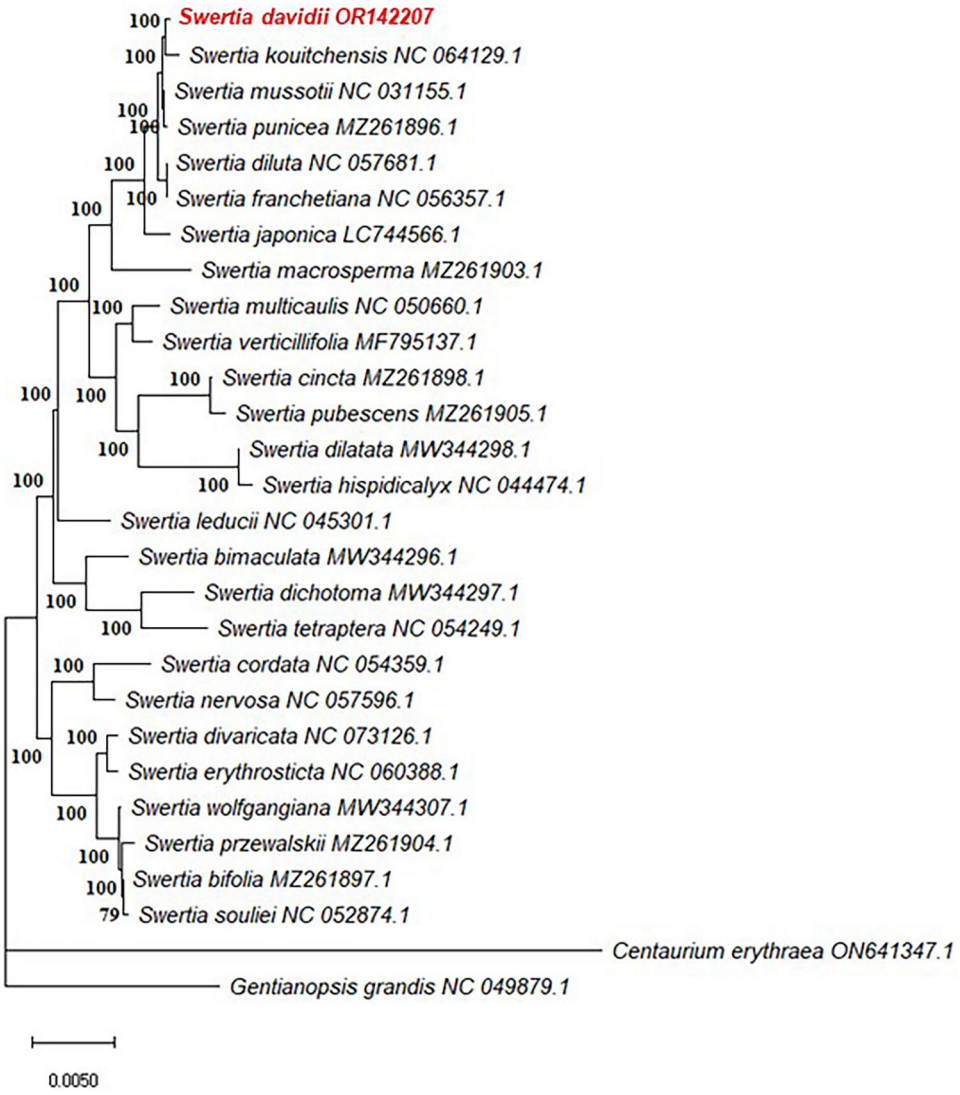

Figure 3. Maximum likelihood phylogenetic tree for *S. davidii* based on 27 complete chloroplast genomes. The number adjacent to each node indicates the bootstrap support value. The following sequences were used: *S. bifolia* (MZ261897.1) (Cao et al. 2022), *S. bimaculate* (MW344295.1) (Xu et al. 2021), *S. cincta* (MZ261898.1) (Cao et al. 2022), *S. cordata* (NC_054359.1) (Yang et al. 2022), *S. davidii* (OR142207, this study), *S. dichotoma* (MW344297.1) (Cao et al. 2022), *S. dilatata* (MW344298.1) (Xu et al. 2021), *S. diluta* (NC_057681.1) (Cao et al. 2022), *S. divaricate* (OQ446461.1) (Hou et al. 2023), *S. erythrosticta* (NC_060388.1) (Xu et al. 2021), *S. franchetiana* (NC_056357.1) (Cao et al. 2022), *S. hispidicalyx* (NC_044474.1) (Zhang et al. 2021), *S. japonica* (LC744566.1) (Yoichi 2023), *S. kouitchensis* (NC_064129.1) (Cao et al. 2022), *S. leducii* (NC_045301.1) (Yang et al. 2022), *S. macrosperma* (MZ261903.1) (Cao et al. 2022), *S. multicaulis* (NC_050660.1) (Yang et al. 2022), *S. mussotii* (NC_031155.1) (Xu et al. 2021), *S. nervosa* (NC_057596.1) (Yang et al. 2022), *S. przewalskii* (MZ261904.1) (Cao et al. 2022), *S. pubescens* (MZ261905.1) (Cao et al. 2022), *S. punicea* (MZ261896.1) (Cao et al. 2022), *S. souliei* (NC_052874.1) (Bi et al. 2020), *S. tetraptera* (NC_054249.1) (Xu et al. 2021), *S. verticillifolia* (MF795137.1) (Yang et al. 2022), *S. wolfgangiana* (MW344307.1) (Cao et al. 2022), *centaurium erythraea* (ON641347.1) (Carvalho Leonardo et al. 2023), and *gentianopsis grandis* (NC_049879.1) (Li et al. 2021).

The phylogenetic patterns of *Swertia* have been elucidated using variable cpDNA loci for species identification. These loci encompass gene fragments such as *mat*K (Xi et al. 2014), *mat*K and *rbc*L (Cao et al. 2021), *mat*K and *trn*L(UAA) (von Hagen and Kadereit 2002), *trn*L(UAA), *trn*L(UAA)-*trn*F(GAA), and *trn*S(UGA)-*ycf*9 (Chassot et al. 2001). However, these results contained several flawed solutions. Moreover, several clades in previously reported phylogenetic trees remained unclear. In this study, detailed genomic data will enable scientists to accurately distinguish closely related *Swertia* species.

The comprehensive genome database provides a solid foundation for future studies on the evolutionary history of *Swertia* species. Using the complete cp genomes, we can reconstruct the phylogenetic tree with high taxonomic resolution, aiding in understanding the evolutionary history and the factors driving evolutionary changes in *Swertia*.

Ethical approval

This article does not contain any studies with human participants or animals performed by any authors. The species described in this paper is not endangered, protected, or personally owned. The plant material was collected in accordance with guidelines provided by the authors' institution (School of Life Science and Technology, Wuhan Polytechnic University) and national regulations.

Author's contributions

Chao Xiong conceived and designed this study. Zhishi Zhang and Yongbiao Deng contributed to the analytical methods, performed

experiments, and wrote the first draft of the manuscript. Lan Wu conducted the identification of *Swertia davidii* and revised the manuscript. Haoren Zhu contributed to sample preparation, figure production, and acquisition data. Yulong Song, Aotian He, Jian Dai, Jinhui Qin, and Lihuan Luo conducted data analysis. All authors have read and agreed to the published version of the manuscript.

Disclosure statement

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

Funding

This work was supported by the Research Funding of Wuhan Polytechnic University under Grant [no. 2021RZ024], Research and Innovation Initiatives of WHPU under Grant [no. 2021Y14].

Data availability statement

The genome data that support the findings of this study are openly available in GenBank (https://www.ncbi.nlm.nih.gov) under accession no. OR142207. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA983562, SRR24921675, and SAMN35731316, respectively.

References

Bi H, Yang Y, Hu Q, Ma Y. 2020. Characterization of the complete chloroplast genome of *Swertia souliei* (Gentianaceae). Mitochondrial DNA Part B Resour. 5(2):1901–1902. doi:10.1080/23802359.2020.1754944.

Carvalho Leonardo I, Alberti A, Denoeud F, Barreto Crespo MT, Capelo J, Bustos Gaspar F. 2023. The complete plastome of *Centaurium erythraea* subsp. *majus* (Hoffmanns. & Link) M. Laínz (Gentianaceae), the first chloroplast genome belonging to the *Centaurium* genus. Mitochondrial DNA Part B Resour. 8(1):86–90. doi:10.1080/23802359.2022.2160670.

Cao Q, Gao Q, Ma X, Zhang F, Xing R, Chi X, Chen S. 2022. Plastome structure, phylogenomics and evolution of plastid genes in *Swertia* (Gentianaceae) in the Qing-Tibetan Plateau. BMC Plant Biol. 22(1):195. doi:10.1186/s12870-022-03577-x.

Cao Q, Xu LH, Wang JL, et al. 2021. Molecular phylogeny of subtribe Swertiinae. Bull Botanical Res. 41(03):408–418.

Chassot P, Nemomissa S, Yuan Y-M, Küpfer P. 2001. High paraphyly of *Swertia* L. (Gentianaceae) in the *Gentianella*-lineage as revealed by nuclear and chloroplast DNA sequence variation. Plant Syst Evol. 229(1–2):1–21. doi:10.1007/s006060170015.

Flora of China Editorial Committee. 1988. Flora of China. Beijing: Science and Technology Press, 62: p. 380–381.

Hou Y, Zhang Y, Yu J, Wang X. 2023. Characteristics of the complete chloroplast genome of *Swertia divaricata* Harry Sm. (Gentianaceae) and its phylogenetic inference. Mitochondrial DNA Part B Resour. 8(10):1132–1136. doi:10.1080/23802359.2023.2270211.

Hunan Academy of Chinese Medicine. 1979. Hunan Pharmacology. Changsha: Hunan People Press,

Kazutaka 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.

Li H, Zhao J, Zhou N. 2021. The complete chloroplast genome sequence of *Gentianopsis grandis* (Harry Sm.) Ma (Gentianaceae) and phylogenetic analysis. Mitochondrial DNA Part B Resour. 6(8):2297–2298. doi:10.1080/23802359.2021.1947910.

Li LI, Huang H, Chen Y. 2002. Preliminary studies on natural resource of *Swertia davidii*. Chin Tradit Herb Drugs. 33(5):466–468.

Liang J, Li S, Chen C. 2014. Chemical constituents in *Swertia davidii*. Chin Tradit Herb Drugs. 45(7):919–923.

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.

Ma QG, Zhang JG, Zhang JP. 2020. The complete chloroplast genome of *Catalpa ovata* G. Don. (Bignoniaceae). Mitochondrial DNA Part B Resour. 5(2):1800–1801. doi:10.1080/23802359.2020.1750979.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22(21):2688–2690. doi:10.1093/bioinformatics/btl446.

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.

Vaidya H, Goyal RK, Cheema SK. 2013. Anti-diabetic activity of swertiamarin is due to an active metabolite, gentianine, that upregulates PPAR-g gene expression in 3T3-L1 cells. Phytother Res. 27(4):624–627. doi:10.1002/ptr.4763.

von Hagen KB, Kadereit JW. 2002. Phylogeny and fower evolution of the *Swertiinae* (Gentianaceae-Gentianeae): homoplasy and the principle of variable proportions. Syst Botany. 27(3):548–572.

Wang Y, Wen F, Hong X, Li Z, Mi Y, Zhao B. 2022. Comparative chloroplast genome analyses of *Paraboea* (Gesneriaceae): insights into adaptive evolution and phylogenetic analysis. Front Plant Sci. 13:1019831. doi:10.3389/fpls.2022.1019831.

Xi HC, Sun Y, Xue CY. 2014. Molecular phylogeny of Swertiinae (Gentianaceae-Gentianeae) based on sequence data of ITS and *mat*K. Plant Divers. 36(002):145–156.

Xu X, Li J, Chu R, Luan M, Wang H, Song K, Wei S, Shi Y, Zhu S, Wei Z, et al. 2021. Comparative and phylogenetic analyses of *Swertia* L. (Gentianaceae) medicinal plants (from Qinghai, China) based on complete chloroplast genomes. Genet Mol Biol. 45(1):e20210092. doi:10.1590/1678-4685-GMB-2021-0092.

Yang L, Li J, Zhou G. 2022. Comparative chloroplast genome analyses of 23 species in *Swertia* L. (Gentianaceae) with implications for its phylogeny. Front Genet. 13:895146. doi:10.3389/fgene.2022.895146.

Yoichi W. 2023. The complete chloroplast genome sequence of *Swertia japonica* (Schult.) Makino (Gentianaceae). Mitochondrial DNA Part B Resour. 8(11):1179–1182. doi:10.1080/23802359.2023.2275335.

Zeng W, Han H, Tao Y, Yang L, Wang Z, Chen K. 2013. Identification of bio-active metabolites of gentiopicroside by UPLC/Q-TOF MS and NMR. Biomed Chromatogr. 27(9):1129–1136. doi:10.1002/bmc.2917.

Zhang Y, Yu J, Xia M, Chi X, Khan G, Chen S, Zhang F. 2021. Plastome sequencing reveals phylogenetic relationships among *Comastoma* and related taxa (Gentianaceae) from the Qinghai-Tibetan Plateau. Ecol Evol. 11(22):16034–16046. doi:10.1002/ece3.8274.