Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

🔓 OPEN ACCESS   Check for updates

Characterization and phylogenetic analysis of the complete chloroplast genome of *Carpesium lipskyi* (Asteraceae, Inuleae)

Tingyu Li[a,b], Hui Chen[a], Xinyu Chen[a], Tianmeng Qu[a], Xinyi Zheng[a], Liang Pang[c,d], Xianhua Gu[e*] and Zhixi Fu[a,b,f*]

[a]Ministry of Education, Key Laboratory of Land Resources Evaluation and Monitoring in Southwest (Sichuan Normal University), Chengdu, China; [b]College of Life Sciences, Sichuan Normal University, Chengdu, China; [c]Sichuan Tianshengyuan Environmental Services Co., Ltd, Chengdu, China; [d]Sichuan Tianshengyuan Information Technology Co., Ltd, Chengdu, China; [e]Sichuan Leshan Ecological Environment Monitoring Center Station, Leshan, China; [f]Sustainable Development Research Center of Resources and Environment of Western Sichuan, Sichuan Normal University, Chengdu, China

ABSTRACT

The species of *Carpesium lipskyi* C.Winkl. 1998 is an important traditional Chinese medicine in China. In this study, the complete chloroplast (cp) genome of *C. lipskyi* was determined and analyzed. The result showed that the complete cp genome of *C. lipskyi* was 151,244 bp in length, consisting of a large single-copy (LSC) region of 82,908 bp, a small single-copy (SSC) region of 18,430 bp, and a pair of inverted repeats (IRs) of 24,953 bp. The overall GC content of the *C. lipskyi* is 37.68%. The species of *C. lipskyi* possessed 127 genes, including 83 protein-coding genes, 36 transfer RNA genes, and eight ribosomal RNA genes. The present study found that *Inula* is sister groups with the closest genetic relationship. The obtained knowledge could provide useful information for future phylogenetic, taxonomic, and evolutionary studies on Inuleae.

Introduction

The genus *Carpesium* (Asteraceae) includes approximately 21 species worldwide. These plants are mostly perennial herbs, such as *Carpesium lipskyi* (Chen and Anderberg 2011). Most of them are distributed across Asia and Europe, particularly in the mountainous areas of Southwest China, and 17 species are in China (Chen and Anderberg 2011; Shi, Hou, et al. 2022; Shi, Xu, et al. 2022; Zhong et al. 2022). It was mainly distributed in Gansu, Qinghai, Sichuan, Yunnan province and growed in forest margins and hillside thickets at altitudes of 2000–3500 m (Chen and Anderberg 2011). It has served as a folk medicine for the treatment of sore throat, stomachache and insect and snake bites (Chen and Anderberg 2011).

The complete chloroplast (cp) genomes present an effective means of improving the rate of species identification and has been developed as a tool for plant phylogenetic studies at different taxonomic levels (Jansen et al. 2007; Parks et al. 2009; Gitzendanner et al. 2018). Several studies have reported the complete cp genomes of Asteraceae (Vargas et al. 2017; Zhang et al. 2021; Liu et al. 2023). Previously, some sequences of *C. lipskyi* such as *ndhF*, *matK* have been reported and studied (Chen et al. 2016). To date, the cp genome of *C. lipskyi* has not been reported and analyzed. Here, we sequenced the cp genomes of *C. lipskyi*, to conduct comprehensive research on the *Carpesium* cp genome, and serve as a reference for subsequent phylogenomic studies of the *Carpesium*.

Materials and methods

The healthy and fresh leaves of *C. lipskyi* (Figure 1) were collected from Chuanzhu Temple, Songpan County, Sichuan Province (103°34′43.12″E, 32°47′42.90″N, alt. 3101 m). The voucher specimens were deposited in the herbarium of Sichuan Normal University, China (SCNU) (https://bio.sicnu.edu.cn/; contact person: Dr. Zhixi Fu, email: fuzx2017@sicnu.edu.cn), under the voucher number: FZX1088. Total genomic DNA was obtained using the CTAB method (Doyle and Doyle 1987). We follow the Illumina DNA Library Construction Guide to complete paired-end DNA library construction (Allen et al. 2006). The complete cp genome was sequenced on the Illumina HiSeq XTen platform (San Diego, CA).

In this study, SPAdes v3.10.1 software was used to assemble high-quality data by default parameters (Bankevich et al.

Figure 1. The morphological characteristics of *C. lipskyi*. (A) Leaves and (B) capitula, respectively. Herbs, perennial; rhizome prostrate; leaves blade elliptic or spatulate; marginal florets narrowly funnelform, disk florets 3–3.5 mm. These photos were taken by Gang Fu in Chuanzhu Temple, Songpan County, Sichuan Province, September 2020 without any copyright issues.

2012). To check the accuracy of assembly, this study used BWA (version v0.7.17) to generate depth of coverage (https://www.protocols.io/view/generating-sequencing-depth-and-coverage-map-for-o-4r3l27jkxg1y/v1). Then, the cp genome sequence was annotated by PGA (Qu et al. 2019) and the result was checked by Geneious R11 (Kearse et al. 2012) and manually adjusted as needed. We further used CPGView (http://47.96.249.172:16085/cpgview/view) to improve annotation, visualize the structure of the cp genome, and identify gene structures, including cis-splicing and trans-splicing (Liu et al. 2023). The genome sequence of *C. lipskyi* has been deposited in GenBank (accession numbers: NC080873 and OR004472).

The phylogenetic analysis of the complete cp genomic dataset of 26 species was performed using the ML method implemented in RAxML. The species of *Anthriscus cerefolium* Hoffm. (GU456628) and *Kalopanax septemlobus* (Thunb.) Koidz. (NC022814) were selected as outgroups. These cp genome sequences were imported into MAFFT v7.520 software (Katoh and Standley 2013) for multiple comparisons, and the phylogenetic tree was constructed using CIPRES (https://www.phylo.org/), with maximum-likelihood (ML) method based on the GTRGAMMA model, and the bootstrap was set to 1000 (Stamatakis et al. 2008).

Results

The complete cp genome of *C. lipskyi* was 151,244 bp in size. The cp genome of *C. lipskyi* displayed a common quadripartite structure consisting of inverted repeats (IRs) (24,953 bp), large single-copy (LSC) region (82,908 bp), and small single-copy (SSC) region (18,430 bp), respectively (Figure 2). The entire coverage depth of *Carpesium lipskyi* cp genome is good (Figure S1). Overall, GC content of the *C. lipskyi* cp genome was 37.68%, and the GC content of the SSC and LSC regions was 31.34% and 35.87%. Each IR region was relatively rich in GC-rich rRNA genes and tRNA genes. The GC content of the IR region was 43.04%, which was significantly higher than that of the LSC and SSC regions. The species of *C. lipskyi* possessed 127 unique genes, including 83 protein-coding genes, 36 tRNA genes, and eight rRNA genes. Eleven genes including *rps16*, *rpoC1*, *atpF*, *petB*, *petD*, *rpl16*, *rpl2*, *ycf3*, *ndhB*,

clpP, and *ndhA* are cis-splicing genes; *ycf3* and *clpP* had two copies (Figure S2). The *rps12* is a trans-splicing gene (Figure S3).

In this study, 26 cp genome sequences and measured cp genome sequence of *C. lipskyi* were used to construct the ML evolutionary tree. The phylogenetic tree analysis revealed that *Carpesium* and *Inula* were sister groups (BS = 100%). *Carpesium* was divided into three main lineages, the new sequence *C. lipskyi* clustered with *C. faberi* C.Winkl. and *C. cernuum* L. (Figure 3).

Discussion and conclusions

The complete cp genome of *C. lipskyi* is analyzed based on second-generation high-throughput sequencing technology and bioinformatics analysis. The results showed that the complete cp genomes of *C. lipskyi* possessed a typical quadripartite structure. The cp genome of *C. lipskyi* is found to be a typical circular DNA molecule (151,244 bp) long, which is consistent with the currently reported cp genomes varying in size between 140,000 bp and 160,000 bp in angiosperms (Bock et al. 2014; Vargas et al. 2017; Smith 2015). The species of *C. lipskyi* is clustered with other species of Inuleae to form a highly supported clade (BS = 100%), which is consistent with previous results (He et al. 2022; Shi, Hou, et al. 2022; Shi, Xu, et al. 2022).

The study provides genomic information of *C. lipskyi* and provides new insights into the phylogenetics and evolution of *Carpesium* and also shed the lights on the genetic diversity of wild germplasm resources of *Carpesium*.

Author contributions

Ethical approval

Carpesium lipskyi does not belong to the national key protected wild plants category, and collecting it does not violate the Regulations of the

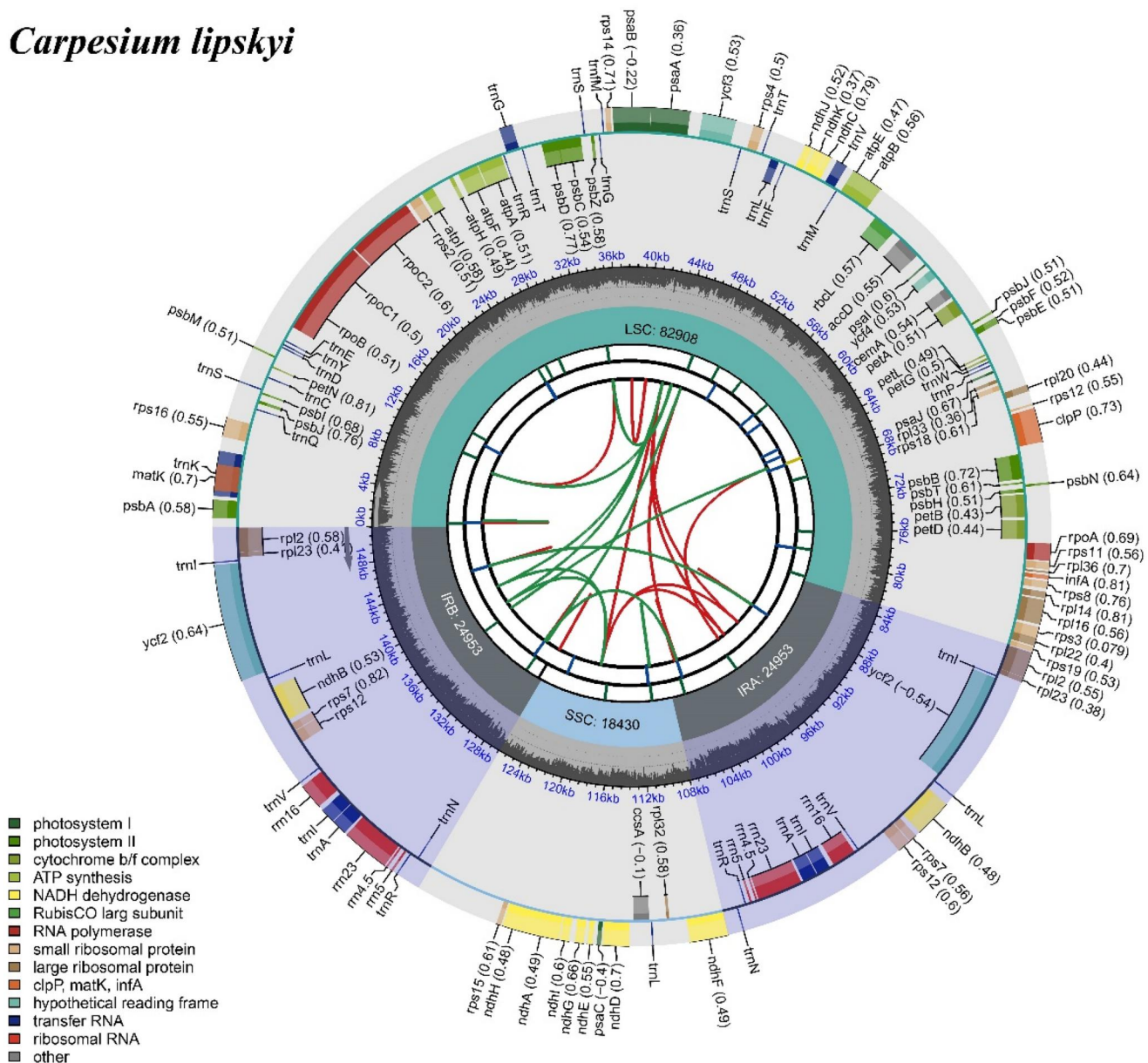**Figure 2.** Circular map of *C. lipskyi*. The map of complete chloroplast genome was generated using CPGView (http://www.1kmpg.cn/cpgview). The map consists of six tracks. From the center to the outer, the first track shows dispersed repeats connected by red and green arcs indicating the direction (forward and reverse, respectively). The second track shows long tandem repeats as blue bands, and the third track shows short tandem repeats or microsatellites as green bands. The fourth track represents the GC content along the plastome. Finally, the sixth track represents the genes as colored boxes, the inner boxes present clockwise transcription, and the outer boxes present counterclockwise transcribed genes. The optional codon usage bias is displayed in the parenthesis after the gene.

People's Republic of China on Wild Plants Protection. According to Article 5 of the regulation, the state encourages and supports scientific research on wild plants and on-site and *ex situ* protection of wild plants. No ethical approval or specific permission was needed in this research.

Disclosure statement

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

Funding

Data availability statement

After uploading the data, the NCBI database releases two accession numbers of *C. lipskyi*. These two accession numbers contain exactly the same sequence information and author. Data are available in the NCBI GenBank at https://www.ncbi.nlm.nih.gov (accession numbers: NC080873 and OR004472). The associated BioProject, SRA, and BioSample numbers are PRJNA1091223, SRR28435025, and SAMN40590854, respectively.
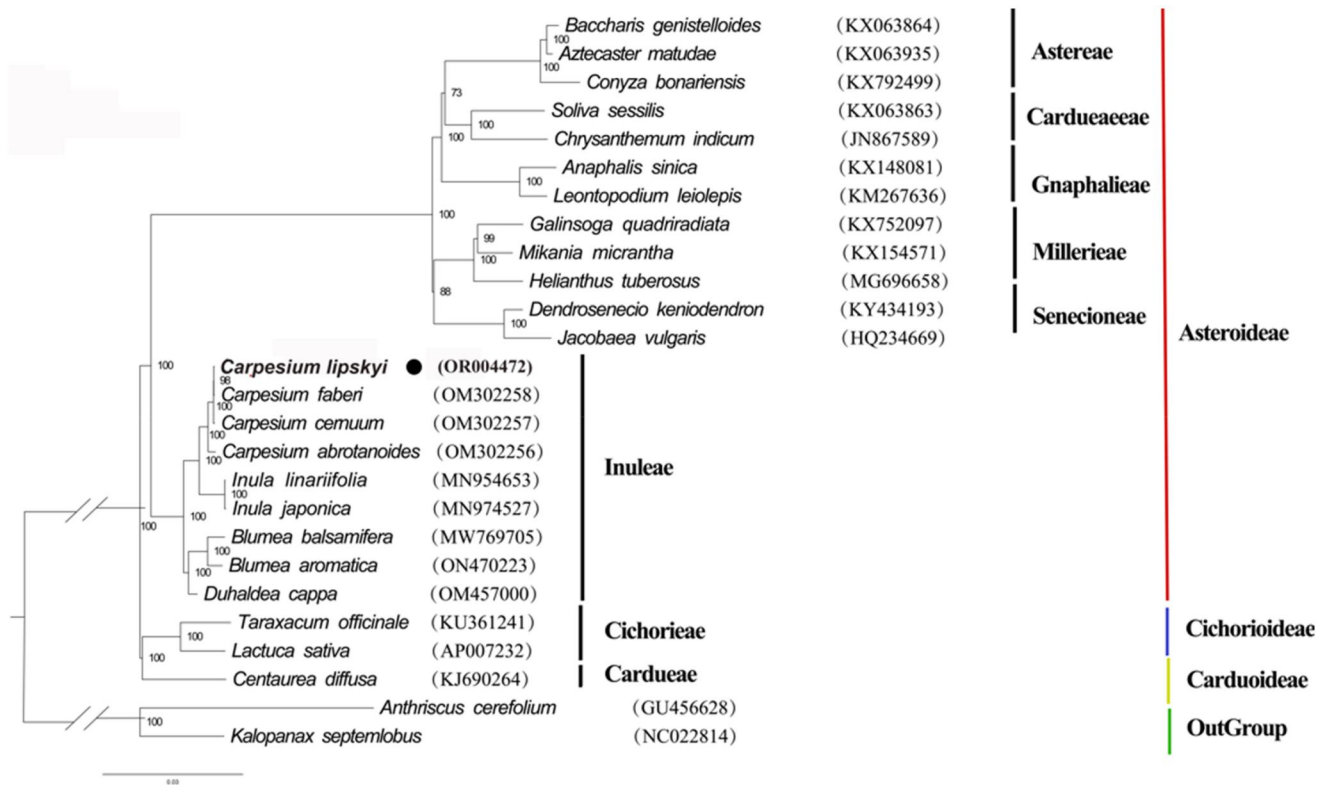
Figure 3. The best maximum-likelihood (ML) phylogram inferred from 26 chloroplast genomes (bootstrap values are indicated on the branches). The circle represents newly sequenced species. The following sequences were used: *Baccharis genistelloides* (Lam.) Pers. KX063864 (Vargas et al. 2017), *Aztecaster matudae* (Rzed.) G.L.Nesom KX063935 (Vargas et al. 2017), *Conyza bonariensis* (L.) Cronquist KX792499 (Wang et al. 2018), *Soliva sessilis* Ruiz & Pav. KX063863 (Vargas et al. 2017), *Chrysanthemum indicum* L. JN867589 (unpublished), *Anaphalis sinica* Hance KX148081 (unpublished), *Leontopodium leiolepis* Nakai KM267636 (unpublished), *Galinsoga quadriradiata* Ruiz & Pav. KX752097 (Wang et al. 2018), *Mikania micrantha* Kunth KX154571 (Huang et al. 2016), *Helianthus tuberosus* Parry MG696658 (unpublished), *Dendrosenecio keniodendron* (R.E.Fr. & T.C.E.Fr.) B.Nord. KY434193 (unpublished), *Jacobaea vulgaris* Gaertn. HQ234669 (Doorduin et al. 2011), *Carpesium lipskyi* OR004472 (the newly sequenced specie), *Carpesium faberi* C.Winkl. OM302258 (unpublished), *Carpesium cernuum* L. OM302257 (Shi, Hou, et al. 2022; Shi, Xu, et al. 2022), *Carpesium abrotanoides* L. OM302256 (Shi, Hou, et al. 2022; Shi, Xu, et al. 2022), *Inula linariifolia* Turcz. MN954653 (unpublished), *Inula japonica* Thunb. MN974527 (unpublished), *Blumea balsamifera* (L.) DC. MW769705 (Zhao et al. 2021), *Blumea aromatica* DC. ON470223 (unpublished), *Duhaldea cappa* (Buchanan-Hamilton ex D. Don) Pruski & Anderberg OM457000 (unpublished), *Taraxacum officinale* F. H. Wigg. KU361241 (unpublished), *Lactuca sativa* L. AP007232 (unpublished), *Centaurea diffusa* Lam. KJ690264 (unpublished), *Anthriscus cerefolium* Hoffm. GU456628 (Downie and Jansen 2015), and *Kalopanax septemlobus* (Thunb.) Koidz. NC022814 (Li et al. 2013).

References

Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. 2006. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat Protoc. 1(5):2320–2325. doi:10.1038/nprot.2006.384.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477. doi:10.1089/cmb.2012.0021.

Bock DG, Kane NC, Ebert DP, Rieseberg LH. 2014. Genome skimming reveals the origin of the *Jerusalem Artichoke* tuber crop species: neither from Jerusalem nor an artichoke. New Phytol. 201(3):1021–1030. doi:10.1111/nph.12560.

Chen YS, Anderberg AA. 2011. *Carpesium*. In: Wu ZY, Raven PH, Hong DY, editors. Flora of China. Vol. 20–21. Beijing/St. Louis: Science Press/Missouri Botanical Garden Press; p. 801–822.

Chen ZD, Yang T, Lin L, Lu LM, Li HL, Sun M, Liu B, Chen M, Niu YT, Ye JF, et al. 2016. Tree of life for the genera of Chinese vascular plants. J Syst Evol. 54(4):277–306. doi:10.1111/jse.12219.

Doorduin L, Gravendeel B, Lammers Y, Ariyurek Y, Chin-A-Woeng T, Vrieling K. 2011. The complete chloroplast genome of 17 individuals of pest species *Jacobaea vulgaris*: SNPs, microsatellites and barcoding markers for population and phylogenetic studies. DNA Res. 18(2):93–105. doi:10.1093/dnares/dsr002.

Downie SR, Jansen RK. 2015. A comparative analysis of whole plastid genomes from the Apiales: expansion and contraction of the inverted repeat, mitochondrial to plastid transfer of DNA, and identification of highly divergent noncoding regions. Syst. Bot. 40(1):336–351. doi:10.1600/036364415X686620.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19:11–15.

Gitzendanner MA, Soltis PS, Wong GKS, Ruhfel BR, Soltis DE. 2018. Plastid phylogenomic analysis of green plants: a billion years of evolutionary history. Am J Bot. 105(3):291–301. doi:10.1002/ajb2.1048.

He XY, Dong SJ, Gao CS, Wang QR, Zhou MJ, Cheng RB. 2022. The complete chloroplast genome of *Carpesium abrotanoides* L. (Asteraceae): structural organization, comparative analysis, mutational hotspots and phylogenetic implications within the tribe Inuleae. Biologia. 77(7):1861–1876. doi:10.1007/s11756-022-01038-2.

Huang L, Wang Z, Wang T, Su YJ. 2016. The complete chloroplast genome sequence of *Mikania micrantha* (Asteraceae), a noxious invasive weed to South China. Mitochondrial DNA B Resour. 1(1):603–604. doi:10.1080/23802359.2016.1209090.

Jansen RK, Cai Z, Raubeson LA, Daniell H, Depamphilis CW, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, 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(49):19369–19374. doi:10.1073/pnas.0709121104.

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.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the

organization and analysis of sequence data. Bioinformatics. 28(12): 1647–1649. doi:10.1093/bioinformatics/bts199.

Li R, Ma PF, Wen J, Yi TS. 2013. Complete sequencing of five Araliaceae chloroplast genomes and the phylogenetic implications. PLOS One. 8(10):e78568. doi:10.1371/journal.pone.0078568.

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.

Liu XF, Luo JJ, Zhang MK, Wang Q, Liu J, Wu D, Fu ZX. 2023. Phylogenomic analysis of two species of *Parasenecio* and comparative analysis within tribe Senecioneae (Asteraceae). Diversity. 15(4):563. doi:10.3390/d15040563.

Parks M, Cronn R, Liston A. 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol. 7(1):84. doi:10.1186/1741-7007-7-84.

Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods. 15(1):50. doi:10.1186/s13007-019-0435-7.

Shi NN, Hou CC, Liu Y, Li KY, Mi SD, Tong BL, Zhang ML. 2022. Chemical constituents of plants from the genus *Carpesium*. Heterocycl Commun. 28(1):95–123. doi:10.1515/hc-2022-0012.

Shi X, Xu W, Wan M, Sun Q, Chen Q, Zhao C, Sun KF, Shu YX. 2022. Comparative analysis of chloroplast genomes of three medicinal *Carpesium* species: genome structures and phylogenetic relationships. PLOS One. 17(8):e0272563. doi:10.1371/journal.pone.0272563.

Smith DR. 2015. Mutation rates in plastid genomes: they are lower than you might think. Genome Biol Evol. 7(5):1227–1234. doi:10.1093/gbe/evv069.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol. 57(5):758–771. doi:10.1080/10635150802429642.

Vargas OM, Ortiz EM, Simpson BB. 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). New Phytol. 214(4):1736–1750. doi:10.1111/nph.14530.

Wang A, Wu H, Zhu X, Lin J. 2018. Species identification of *Conyza bonariensis* assisted by chloroplast genome sequencing. Front Genet. 9:374. doi:10.3389/fgene.2018.00374.

Zhang X, Jiang PP, Fan SJ. 2021. Characterization of the complete plastome of *Aster pekinensis* (Asteraceae), a perennial herb. Mitochondrial DNA B Resour. 6(3):1064–1065. doi:10.1080/23802359.2021.1899081.

Zhao C, Xu W, Huang Y, Sun Q, Wang B, Chen C, Chen Q. 2021. Chloroplast genome characteristics and phylogenetic analysis of the medicinal plant *Blumea balsamifera* (L.) DC. Genet Mol Biol. 44(4):e20210095. doi:10.1590/1678-4685-GMB-2021-0095.

Zhong WH, Li M, Gu R, Ding R, Cao L, Mu ZJ, Du XL, Feng YL, Wei RR, Zhan LJ, et al. 2022. Lipskynoids A–G, new acyclic diterpenes from the flowers of *Carpesium lipskyi*. Chem Biodivers. 19(12):e202200898. doi:10.1002/cbdv.202200898.