

The complete chloroplast genome of *Lagochilus ilicifolius* Bunge ex Bentham, Labiat. Gen. 1834 (Lamiaceae) and its phylogenetic analysis

Wen-Juan Hou*, Wen-Xiao Men*, Yan-Ping Xing, Yan-yun Yang, Liang Xu, Ting-Guo Kang and Yue-Yue Song
School of Pharmacy, Liaoning University of Traditional Chinese Medicine, Dalian, China

ABSTRACT

Lagochilus ilicifolius Bunge ex Bentham, Labiat. Gen is a perennial herb with much-branched stems native to Nei Mongol, Ningxia, Gansu, N Shaanxi. It can be used clinically as a hemostatic agent. The chloroplast genome length is 151,466 bp. It contained two inverted repeat regions of 25,660 bp each, a large single-copy region of length 82,504 bp, and a small single-copy region of length 17,642 bp. Also, the GC content is 38.6%. There were 133 genes annotated, including 88 known protein-coding genes, 37 tRNAs, and eight rRNAs. The phylogenetic tree was constructed using Bayesian method for plastome data of 29 species. The entire chloroplast genome of *L. ilicifolius* within the Lamiaceae is the first to reveal genetic taxonomy at the molecular level, and the new phylogenetic tree data can be used for future evolutionary studies.

ARTICLE HISTORY

Received 2 November 2023
Accepted 18 March 2024

KEYWORDS

Assembly; annotation;
chloroplast genome;
phylogenetic analysis;
Lamiaceae

Introduction

Lagochilus ilicifolius is a perennial plant in the genus Lamiaceae of the family Lagochilus (Gen 1994). The aqueous extracts of the whole plant of *Lagochilus* have obvious *in vitro* anti-inflammatory activity, and can effectively scavenge DPPH (2,2-diphenyl-1-picrylhydrazyl, a stable free radical that can be used to measure the free radical scavenging activity of antioxidants) free radicals, showing certain *in vitro* antioxidant activity (Jiao et al. 2014). Analysis of the nonpolar constituents of the above-ground parts of *L. ilicifolius* allowed identification of 52 components (accounting for 92.61% of the total amount of nonpolar components), which include mainly fatty acid esters, long-chain alkanes, monoterpenes, sesquiterpenes, sterols, and aromatic compounds, and so on, mainly including fatty acid esters, long-chain alkanes, monoterpenes, sesquiterpenes, steroids, and aromatic compounds (Chen et al. 2012). Erythro-1-[(4-O-β-D-glucopyranosyl-3-methoxy)-phenyl]-2-[(5'-methoxy)-pinoresinol]-propane-1,3-diol (1), a newly identified lignan from the *L. ilicifolius*, showed relatively strong cytotoxicity against PC12 cell line which was derived from a rat adrenal medullary chromosomal tumor (Qian et al. 2015).


A study of *L. ilicifolius* from northern China, Mongolia and Russia using five cpDNAs spacers, confirmed that *L. ilicifolius* belonged to two lineages, one in northern China and the other in Inner Mongolia-Russia (Gao et al. 2014).



Figure 1. The morphological characteristics of *L. ilicifolius*. The photographs were taken by Xu Liang in Ordos, Inner Mongolia Autonomous Region, China (108°38'22.54" E, 39°49'13.61" N). The most characteristic feature of the specimen: leaves cuneate-rhombic, upwards, short awned spines at the ends of the teeth, glabrous on both sides; flowers whorled cyme, white tomentum on the outside of the corolla.

CONTACT Liang Xu  861364054@qq.com; Yan-yun Yang  99yyy@163.com  School of Pharmacy, Liaoning University of Traditional Chinese Medicine, Dalian, China

*Both authors contributed equally to this work; share co-first authorship.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2024.2334013>.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

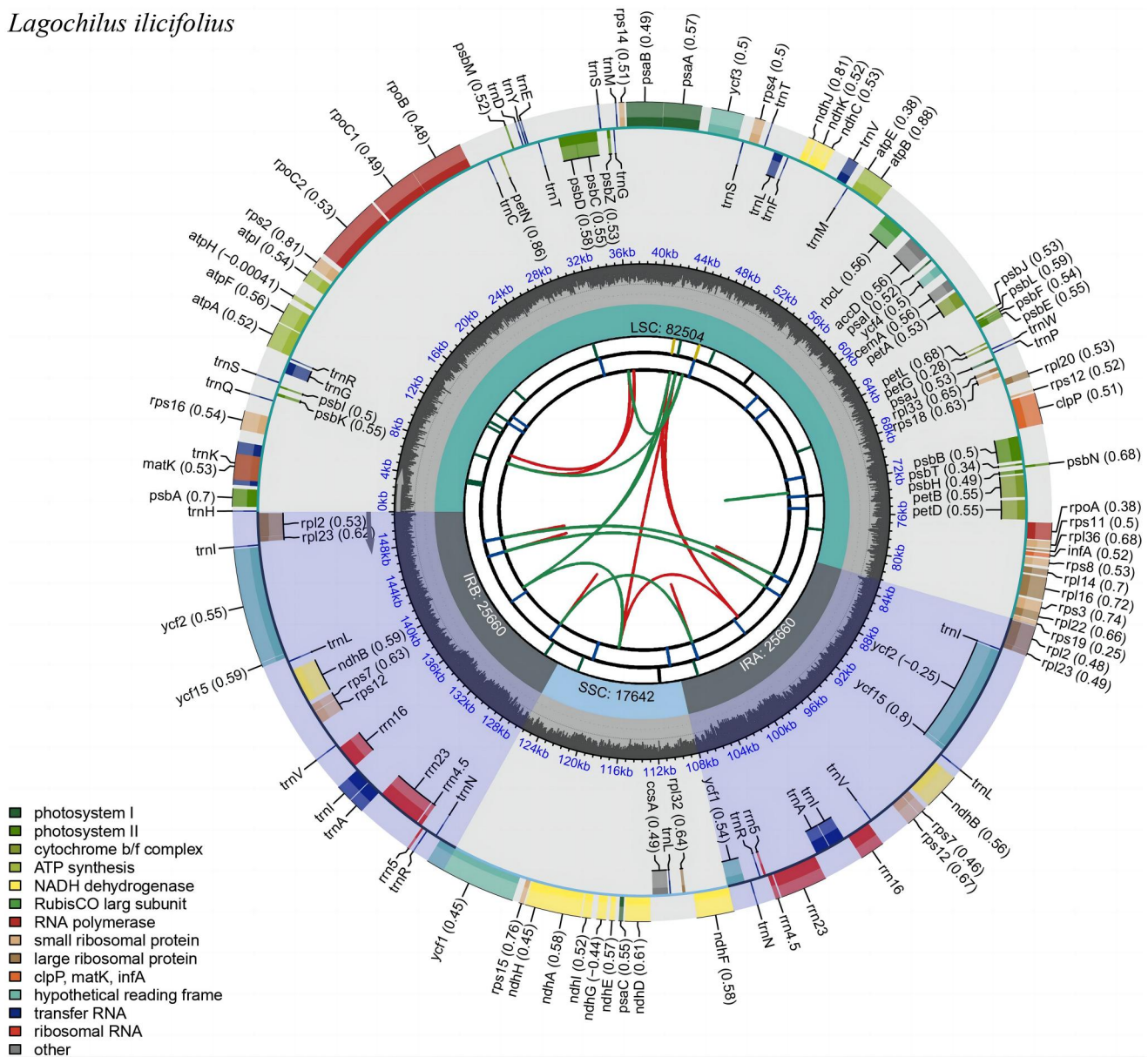
Lagochilus ilicifolius

Figure 2. A circular map of the entire chloroplast genome of *L. ilicifolius* using CPGview (Liu et al. 2023). The map encompasses six tracks that depict various features of the genome. Starting from the center, the first track displays dispersed repeats, consisting of both direct and palindromic repeats, which are delineated by red and green arcs. The second track highlights long tandem repeats represented by short blue bars. In the third track, short tandem repeats or microsatellite sequences are illustrated as colored bars. Each color corresponds to a specific type of repeat, and accompanying descriptions provide valuable information about the characteristics of each repeat type. The colors and their respective repeat types are as follows: 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 chloroplast genome contains an LSC region, an SSC region, and two IR regions, and they are shown on the fourth track. The GC content along the genome is shown on the fifth track. Genes within the genome visualization are meticulously color-coded based on their functional classification. The transcription directions of the inner genes are represented in a clockwise manner, while the outer genes are shown in an anticlockwise orientation. To assist with interpretation, the key for gene functional classification is provided in the bottom left corner of the visualization.

Materials and methods

The fresh leaves were collected in Ordos, Inner Mongolia Autonomous Region, China (108°38'22.54" E, 39°49'13.62" N) (Figure 1). Professor Tingguo Kang from the Liaoning University of Traditional Chinese Medicine in Shenyang, China, authenticated the sample vouchers. A specimen was deposited at the herbarium of Liaoning University of Traditional Chinese Medicine (Liang Xu 861364054@qq.com, *L. ilicifolius* number: 210711191006023LY). All operations were carried out in accordance with Specification on Good

Agriculture and Collection Practices for Medicinal Plants (GACP; Number: T/CCMHPIE 2.1-2018).

A total amount of 0.2 µg DNA per sample was used as input material for the DNA library preparations. Sequencing library was generated using NEB Next® Ultra™ DNA Library Prep Kit for Illumina (NEB, Ipswich, MA) following manufacturer's recommendations and index codes were added to each sample. Genomic DNA sample was fragmented by sonication to a size of 350 bp, and ligated with the full-length adapter for Illumina sequencing, followed by further PCR

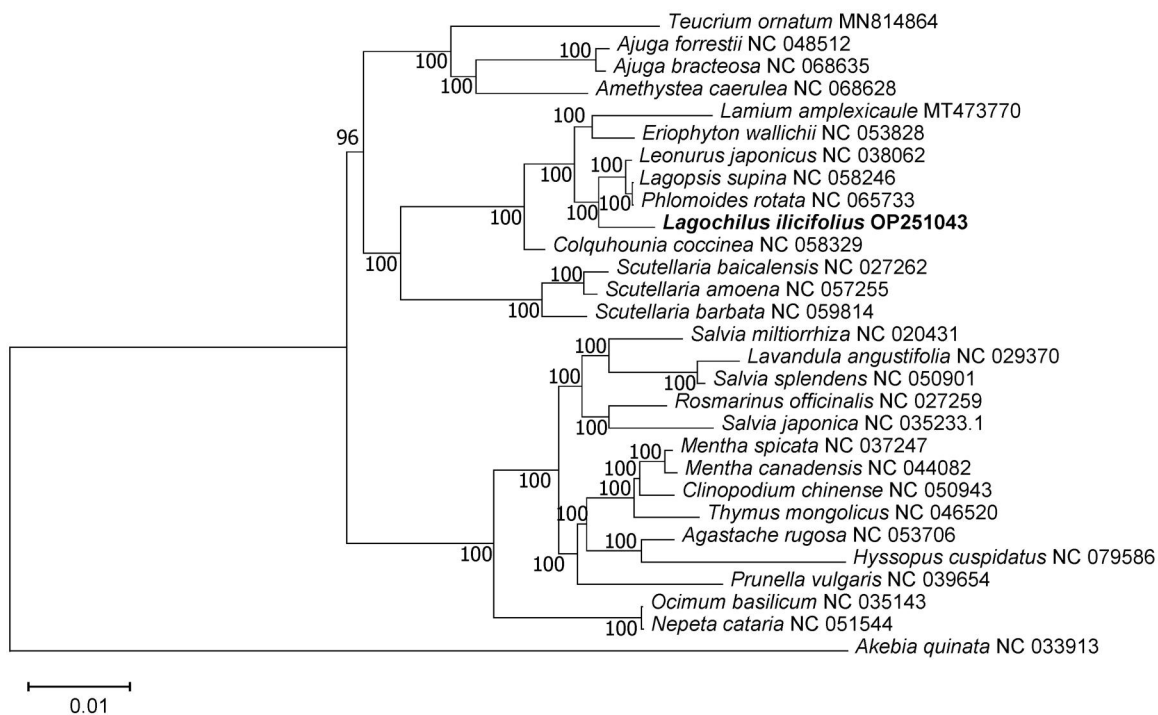


Figure 3. Maximum-likelihood (ML) phylogenetic tree with evolutionary distances for *L. ilicifolius* and 28 other species. The numbers above the branches show the bootstrap support values. The following sequences were used: *Teucrium ornatum* (MN814864), *Ajuga forrestii* (NC048512), *Ajuga bracteosa* (NC068635), *Amethystea caerulea* (NC068628), *Lamium amplexicaule* (MT473770) (Zhao et al. 2021), *Eriophyton wallichii* (NC053828), *Leonurus japonicus* (NC038062), *Lavandula angustifolia* (NC029370), *Phlomoides rotata* (NC065733), *Lagochilus ilicifolius* (OP251043) (this study), *Colquhounia coccinea* (NC058329) (Zhao et al. 2021), *Scutellaria baicalensis* (NC027262), *Scutellaria amoena* (NC057255) (Chen and Zhang 2019), *Scutellaria barbata* (NC059814), *Salvia miltiorrhiza* (NC020431) (Qian et al. 2013), *Salvia splendens* (NC050901) (Zhao et al. 2020), *Rosmarinus officinalis* (NC027259), *Salvia japonica* (NC035233) (Wang et al. 2017), *Mentha canadensis* (NC044082), *Clinopodium chinense* (NC050943) (Zhang et al. 2020), *Thymus mongolicus* (NC046520), *Agastache rugosa* (NC053706), *Hyssopus cuspidatus* (NC079586), *Prunella vulgaris* (NC039654) (Han and Zheng 2018), *Ocimum basilicum* (NC035143), *Nepeta cataria* (NC051544) (Zhou et al. 2020), and *Akebia quinata* (NC033913) (Li et al. 2016).

amplification. After PCR products were purified by AMPure XP system (Beckman Coulter, Beverly, MA), DNA concentration was measured by Qubit[®]3.0 Fluorometer (Invitrogen, Waltham, MA). The clustering of the index-coded samples was performed on a cBot Cluster Generation System using Illumina PE Cluster Kit (Illumina, San Diego, CA) according to the manufacturer's instructions. High-quality reads were assembled into chloroplast genome using a *de novo* assembler SPAdes v3.14.1 (Bankevich et al. 2012). Finally, it was annotated by PGA (Qu et al. 2019) with *Lagopsis supina* (NC058246) as reference genome. The Illumina short sequences were compared to the chloroplast genome sequences using BWA software and finally the coverage was calculated using samtools depth (high coverage of over $\times 100$). The horizontal coordinate is the chloroplast length and the vertical coordinate is the coverage depth (Figure S1) (Li et al. 2009; Li 2013). The maps of the chloroplast genome, cis-splicing genes, and trans-splicing genes of *L. ilicifolius* were processed by CPGview (Liu et al. 2023).

To analyze the relationship between *L. ilicifolius* and other species in the family Lamiaceae, the whole chloroplast genomes of 28 species and an outgroup taxon (*Akebia quinata*) were selected from NCBI. Using mafft-7.037 software (Kato and Standley 2013), 70 shared protein genes were selected for comparison. After sequence alignment using MAFFT, the sequences were then aligned using Gblocks

(Xiao et al. 2021) 0.91b, which is used to extract conserved sites from multiple sequence alignment results for the next evolutionary analysis. The maximum-likelihood evolutionary tree was then constructed using the iqtree-1.6.12 software (Xue et al. 2022) with a self-expansion value of 1000. The selection of the optimal substitution model (TVM + F + I + G4) was performed according to Bayesian information criterion (BIC) method implemented in ModelFinder (Kalyaanamoorthy et al. 2017).

Results

The chloroplast genome of *L. ilicifolius* is 151,466 bp, including a large single-copy region (LSC) of 82,504 bp, a small single-copy region (SSC) of 17,642 bp, and a pair of inverted repeats of 25,660 bp. The overall GC content of the chloroplast genome is 38.6%. The average sequencing depth was $\times 1099$ (Figure S1) (Li et al. 2009; Li 2013). The 133 genes encoding are 88 protein coding genes, eight rRNA genes, and 37 tRNAs. *trnK-UUU*, *rps16*, *trnG-UCC*, *atpF*, *rpoC1*, *trnL-UAA*, *trnV-UAC*, *petB*, *petD*, *rpl16*, *rpl2*, *ndhB*, *trnI-GAU*, *trnA-UGC*, and *ndhA* genes each contain one intron, *clpP*, *ycf3* genes contain two introns, whereas *rps12* appeared as a trans-spliced gene with its 5' end located in the LSC region and the duplicated 3' end in both IR regions. The *L. ilicifolius* chloroplast genome and an annotated map of cis- and trans-

splicing genes (Figure 2, Figures S2 and S3, respectively) were processed with CPGview (Liu et al. 2023).

The topology of the phylogenetic tree revealed the close relationship between *L. ilicifolius* and *Phlomooides rotata*, *Lagopsis supina* and *Leonurus japonicus*, while the phylogenetic relationships between other representatives of the Lamiaceae family were consistent with previous studies. The presented results will help to carry out further studies on Lamiaceae. The complete chloroplast genome sequence of *L. ilicifolius* could provide new directions for the evolutionary history of Lamiaceae and other species (Figure 3).

Discussion and conclusions

The topology of the phylogenetic tree is in agreement with the results of the previous studies and provides new genetic information for the analyses reconstructing the position of *L. ilicifolius* within the family Lamiaceae (Pema et al. 2021). The relationship between *L. ilicifolius* and other species in the genus *Lagochilus* is currently unclear. A fully resolved phylogenetic tree was obtained in this study, enriching our understanding of the genetic composition, genome evolution, and genetic diversity of species in the genus *Lagochilus*. This finding facilitates future comparative and evolutionary analyses of chloroplast genome sequences within the genus *Lagochilus*.

Ethical approval

Lagochilus ilicifolius does not belong to the national key protected wild plants category, and collecting it does not violate the Regulations of the 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.

Author contributions

T. G. K. revised critically important intellectual content and was involved in validation and supervision; W. J. H. and W. X. M. were responsible for conceptualizing, writing, and revising; Y. Y. Y. and L. X. were responsible for identification and collection, preservation of plant specimens, and conceiving the work; Y. P. X. and Y. Y. S. were responsible for acquisition and interpretation of the data; all authors discussed the results and revised the manuscript.

Disclosure statement

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

Funding

This research was funded by Liaoning BaiQianWan Talents Program [No. 2021A039], 2019 Liaoning Provincial Department of Education Scientific Research Project [No. L201942], Major Special Fund for Science and Technology of Inner Mongolia Autonomous Region [No. 2019ZD004], Natural Science Fund Project of Liaoning Province [No. 2020-MS-224], National Natural Science Foundation of China [No. 82373999], Key Project at Central Government Level: The Ability Establishment of

Sustainable Use for Valuable Chinese Medicine Resources [No. 2060302], and 2023 Liaoning University of Traditional Chinese Medicine Seedling Breeding Project.

Data availability statement

The chloroplast genome sequence data that support the findings of this study are openly available in GenBank of NCBI (<https://www.ncbi.nlm.nih.gov/>) under the accession no. OP251043. The associated BioProject, SRA, and BioSample numbers are PRJNA870254, SRX17131821, and SAMN30355605, respectively.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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.
- Chen LG, Mishig X-Z, Yi D, Zang J-H, Luo G-L, Ying-gang. 2012. GC–MS analysis of nonpolar constituents from aerial parts of *Lagochilus ilicifolius*. *Nat Prod Res Dev.* 18(6):924–927.
- Chen Q, Zhang D. 2019. Characterization of the complete chloroplast genome of *Scutellaria amoena* C. H. Wright (Lamiaceae), a medicinal plant in southwest China. *Mitochondrial DNA B Resour.* 4(2):3057–3059. doi:10.1080/23802359.2019.1666677.
- Gao X, Meng H, Zhang M. 2014. Diversification and vicariance of desert plants: evidence inferred from chloroplast DNA sequence variation of *Lagochilus ilicifolius* (Lamiaceae). *Biochem Syst Ecol.* 55:93–100. doi:10.1016/j.bse.2014.03.002.
- Gen. 1994. Lamiaceae. In: Symb Sin, editor. *Flora of China*. Vol. 17. Beijing (China)/St. Louis (MI): Science Press/Missouri Botanical Garden Press; p. 166–169.
- Han Y, Zheng T. 2018. The complete chloroplast genome of the common self-heal, *Prunella vulgaris* (Lamiaceae). *Mitochondrial DNA B Resour.* 3(1):125–126. doi:10.1080/23802359.2018.1424587.
- Jiao Y, Zhang CG, Zhang T, Chou GX, Xu H, Liu Q. 2014. Anti-inflammatory effects of aqueous extracts of whole herbs from five *Lagochilus* species in vitro. *Chin J New Drugs Clin.* 33(3):204–211.
- Kalyanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 14(6):587–589. doi:10.1038/nmeth.4285.
- 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.
- Li B, Li Y, Cai Q, Lin F, Huang P, Zheng Y. 2016. Development of chloroplast genomic resources for *Akebia quinata* (Lardizabalaceae). *Conserv Genet Resour.* 8(4):447–449. doi:10.1007/s12686-016-0593-0.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 25(16):2078–2079. doi:10.1093/bioinformatics/btp352.
- Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Prepr.* 3.
- 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.
- Pema Y, Ma C, Bautista MAC, Chen T. 2021. The complete chloroplast genome sequence of the Tibetan herb *Phlomooides rotata* (Benth. ex Hook.f.) Mathiesen (Lamiaceae). *Mitochondrial DNA B Resour.* 6(11): 3261–3262. doi:10.1080/23802359.2021.1993459.
- Qian J, Song J, Gao H, Zhu Y, Xu J, Pang X, Yao H, Sun C, Li X, Li C, et al. 2013. The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. *PLOS One.* 8(2):e57607. doi:10.1371/journal.pone.0057607.
- Qian J, Zhang C, Wei W, Ting Z, Hong X, Chou G-X. 2015. A new lignan glucoside from *Lagochilus ilicifolius*. *Pharmacogn Mag.* 11(41):191–195.

- Qu X-J, Moore MJ, Li D-Z, Yi T-S. 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.
- Wang K, Li L, Hua Y, Zhao M, Li S, Sun H, Lv Y, Wang Y. 2017. The complete chloroplast genome of *Mentha spicata*, an endangered species native to South Europe. *Mitochondrial DNA B Resour*. 2(2):907–909. doi:10.1080/23802359.2017.1413311.
- Xiao J, Wei X, Zhou Y, Xin Z, Miao Y, Hou H, Li J, Zhao D, Liu J, Chen R, et al. 2021. Genomes of 12 fig wasps provide insights into the adaptation of pollinators to fig syconia. *J Genet Genomics*. 48(3):225–236. doi:10.1016/j.jgg.2021.02.010.
- Xue H, Song Y, Yang Y, Bian C, Xu L, Kang T. 2022. The complete chloroplast genome of *Pulsatilla campanella* Fischer ex Krylov. (Ranunculaceae, *Pulsatilla* Miller). *Mitochondrial DNA B Resour*. 7(6): 1126–1128.
- Zhang L, Zhao F, Tao A. 2020. Characterization of the complete chloroplast genome of *Clinopodium chinense* (Labiatae), a medicinal plant in southwest of China. *Mitochondrial DNA B*. 5(2):1491–1493. doi:10.1080/23802359.2020.1742595.
- Zhao F, Chen Y-P, Salmaki Y, Drew BT, Wilson TC, Scheen A-C, Celep F, Bräuchler C, Bendiksby M, Wang Q, et al. 2021. An updated tribal classification of Lamiaceae based on plastome phylogenomics. *BMC Biol*. 19(1):2. doi:10.1186/s12915-020-00931-z.
- Zhao F, Drew BT, Chen Y-P, Hu G-X, Li B, Xiang C-L. 2020. The chloroplast genome of *Salvia*: genomic characterization and phylogenetic analysis. *Int J Plant Sci*. 181(8):812–830. doi:10.1086/710083.
- Zhou H, Du L, Fang Z. 2020. The complete chloroplast genome sequence of *Nepeta cataria* L. (Labiatae). *Mitochondrial DNA B Resour*. 5(3): 3280–3281. doi:10.1080/23802359.2020.1813058.