Journal of Asia-Pacific Biodiversity 17 (2024) 586–589

Contents lists available at ScienceDirect

Journal of Asia-Pacific Biodiversity

journal homepage: http://www.elsevier.com/locate/japb

Short Communication

Complete chloroplast genome sequences of *Dracocephalum argunense* and *D. integrifolium* (Lamiaceae: Nepetinae)

Rong-Qing Zhang [a,b], Xiao-Lei Ma [a], Ya-Ping Chen [a], Chun-Lei Xiang [a,*]

[a] Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201, China
[b] University of Chinese Academy of Sciences, Beijing, 100049, China

ARTICLE INFO

Article history:
Received 12 January 2024
Received in revised form
24 March 2024
Accepted 7 April 2024
Available online 21 May 2024

Keywords:
Dracocephalum
medicinal plants
Nepetoideae
phylogenomics
plastome

ABSTRACT

Dracocephalum is a large genus comprises approximately 80 species of perennial herbs, mainly distributed in high elevation regions of temperate and Central Asia, with about half native to China. Within the genus, *D. argunense* and *D. integrifolium* are important herbs with great medicinal value. We sequenced plastomes of these two species for the first time. These plastomes showed a typical quadripartite structure and the length varied from 149,978 bp (*D. integrifolium*) to 150,802 bp (*D. argunense*). Plastomes of *D. argunense* and *D. integrifolium* included 127 and 133 genes, respectively. Phylogenetic analyses supported the monophyly of the redefined *Dracocephalum* and systematic relationships within the sampled species were well resolved. Our study added new genetic information about *D. argunense* and *D. integrifolium* and has great potential for further diversity studies, phylogenetic studies, and understanding of evolutionary history in *Dracocephalum*.

Introduction

Dracocephalum L. (Lamiaceae) is the second-largest genus within the subtribe Nepetinae, comprising ca. 80 species (Budantsev 1987; Li and Hedge 1994; Zhao et al. 2021; Chen et al. 2022a; Rose et al. 2023). Geographically, species of *Dracocephalum* are widely distributed in various habitats (steppes, semi-deserts, deserts, and alpine regions) across temperate Eurasia (Chen et al. 2022a), and one species each distribute in North America and North Africa (Budantsev 1987, 1993; Li and Hedge 1994; Harley et al. 2004; Mabberley 2017). Morphologically, *Dracocephalum* is most similar to *Nepeta* L. but can be readily distinguished by its calyces with a thickened sinus-like fold between the bases of the calyx lobes (Li and Hedge 1994; Harley et al. 2004).

Dracocephalum is utilized throughout its range for medicinal purposes, and many species are of economic importance. For example, *D. integrifolium* Bunge ('Marzan Juxi' in Uyghur) is a traditional medicine that is widely used to treat coughs and asthma (Zhou et al. 2019). *Dracocephalum officinalis* (L.) Y.P. Chen & B.T. Drew is one of the most frequently consumed herbal remedies in the world (Fathiazad and Hamedeyazdan 2011). In addition, some species are widely cultivated as ornamental plants (i.e. *D. argunense* Fisch. Ex Rchb.; Kim et al. 2006).

As above-mentioned, both *D. argunense* and *D. integrifolium* are socio-economically important plant species, but their chloroplast genomes have never been reported. This study sequenced their plastomes for the first time and aimed to provide genetic data which will shed new light on the plastid phylogenomics of the genus and provide new insight into plastome evolution across the subfamily Nepetoideae.

Material and methods

Plant material and DNA extraction

Fresh and young leaves of *D. argunense* and *D. integrifolium* were collected in the field from Tonghua City, Jilin province (*Chen YP et al. EM1448*) and Zhaosu County, Xinjiang province, China, respectively (*Chen YP et al. EM814*). Vouchers were deposited in Kunming Institute of Botany, Chinese Academy of Sciences.

Total genomic DNA was extracted using the CTAB protocol (Doyle and Doyle 1987). The library construction followed the method of Chen et al. (2022b). Sequencing with 2 × 150 bp paired-end reads was conducted to generate approximately 10 Gb of data for each accession using an Illumina HiSeq 2000 platform (Illumina,

* Corresponding author. Tel.: +86 182 13838796; ORCID.: 0000-0001-8775-6967
E-mail address: xiangchunlei@mail.kib.ac.cn (C.-L. Xiang).
Peer review under responsibility of National Science Museum of Korea (NSMK) and Korea National Arboretum (KNA).

https://doi.org/10.1016/j.japb.2024.04.001

San Diego, CA, United States) at BGI Genomics (BGI-Shenzhen, China).

Plastome assembly and annotation

Quality control of raw reads was carried out using FastQC v0.12.0 programs (Andrews 2010), with the parameter set as Q ≥ 25 to acquire high-quality, clean reads for *De novo* assembling. Assembly of the plastomes was implemented in the GetOrganelle v.1.7.6.1 (Jin et al. 2020), and the resulting contigs were visualized and edited using Bandage v.0.8.0 (Wick et al. 2015). Annotation was first performed using Plastid Genome Annotator (PGA) (Qu et al. 2019) with the published plastome of *D. tanguticum* Maxim. (MT457746; Yao et al. 2020) as a reference, the start and stop codons were manually adjusted in Geneious v.11.0.3 (Kearse et al. 2012). A circular map of chloroplast genomes were drawn by OGDRAW (Greiner et al. 2019). The script "get_annotated_regions_from_gb.py" developed by Zhang et al. (2020) was employed to extract coding regions (CDS) for phylogenetic analysis.

Phylogenetic analysis

Plastomes of eleven species were downloaded from the GenBank, with *Nepeta hemsleyana* Oliver ex Prain (MZ618620; Bautista et al. 2022) and *Schizonepeta tenuifolia* Briq. (MW013765; Wang et al. 2021) were selected as outgroups. Alignment was originally performed using MAFFT v7.017 (Katoh and Standley 2013) and then trimmed by trimAl (Capella-Gutiérrez et al. 2009). The resulting processed alignment is used to reconstruct a phylogenetic tree using the maximum likelihood (ML) method. ML analyses were conducted using RAxML v8.2.12 (Stamatakis 2014), with 1000 bootstrap iterations were selected with all other parameters set to default. The phylogenetic tree was edited using Figtree v.1.4.1 (http://tree.bio.ed.ac.uk/software/figtree/).

Results and discussion

Characterization of plastomes

In this study, chloroplast genomes of *Dracocephalum argunense* and *D. integrifolium* were newly sequenced and assembled. The annotated plastid genome sequences have been deposited into the GenBank with the accession number of PP379900 (*D. integrifolium*) and PP379901 (*D. argunense*). The complete chloroplast genome length of two species varied from 149,978 bp (*D. integrifolium*) to 150,802 bp (*D. argunense*). Plastomes were conserved in structure (Figure 1), displaying a typical quadripartite structure with a large single-copy (LSC) region (82,228 bp and 81,764 bp), a small single-copy (SSC) region (17,192 bp and 17,042 bp), and a pair of inverted repeat (IR) regions (25,691 bp and 25,586 bp).

When duplicated genes in IR regions were counted only once, the plastome of *D. argunense* included 107 unique genes, with 37 tRNA genes and eight rRNA genes, as well as 82 protein-coding genes, 20 genes were duplicated in the IR regions, including nine protein-coding genes (*ndhF*, *rps7*, *rps19*, *rps12*, *ndhB*, *ycf2*, *ycf15*, *rpl2*, and *rpl23*), four ribosomal RNA genes (*rrn4.5*, *rrn5*, *rrn16*, and *rrn23*), and seven transfer RNA genes (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnI-CAA*, *trnN-GUU*, *trnR-ACG*, *and trnV-GAC*). Seven protein-coding genes (*rpl2*, *atpF*, *ndhA*, *ndhB*, *rpl2*, *rpoC1*, and *rps16*) and six tRNA genes (*trnA-UGC*, *trnG-UCC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, and *trnV-UAC*) contained one intron, and three genes (*rps12*, *ycf3*, *clpP*) contained two introns. Plastome of *D. integrifolium* contained 112 unique genes, including 37 tRNA genes and eight rRNA genes, as well as 88 protein-coding genes. 21 genes were duplicated in the IR regions, including 10 protein-coding genes (*ndhF*, *rps7*, *rps19*, *rps12*, *ndhB*, *ycf1*, *ycf2*, *ycf15*, *rpl2*, and *rpl23*), four ribosomal RNA genes (*rrn4.5*, *rrn5*, *rrn16*, and *rrn23*), and seven transfer RNA genes (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnI-CAA*, *trnN-GUU*, *trnR-ACG*, *and trnV-*

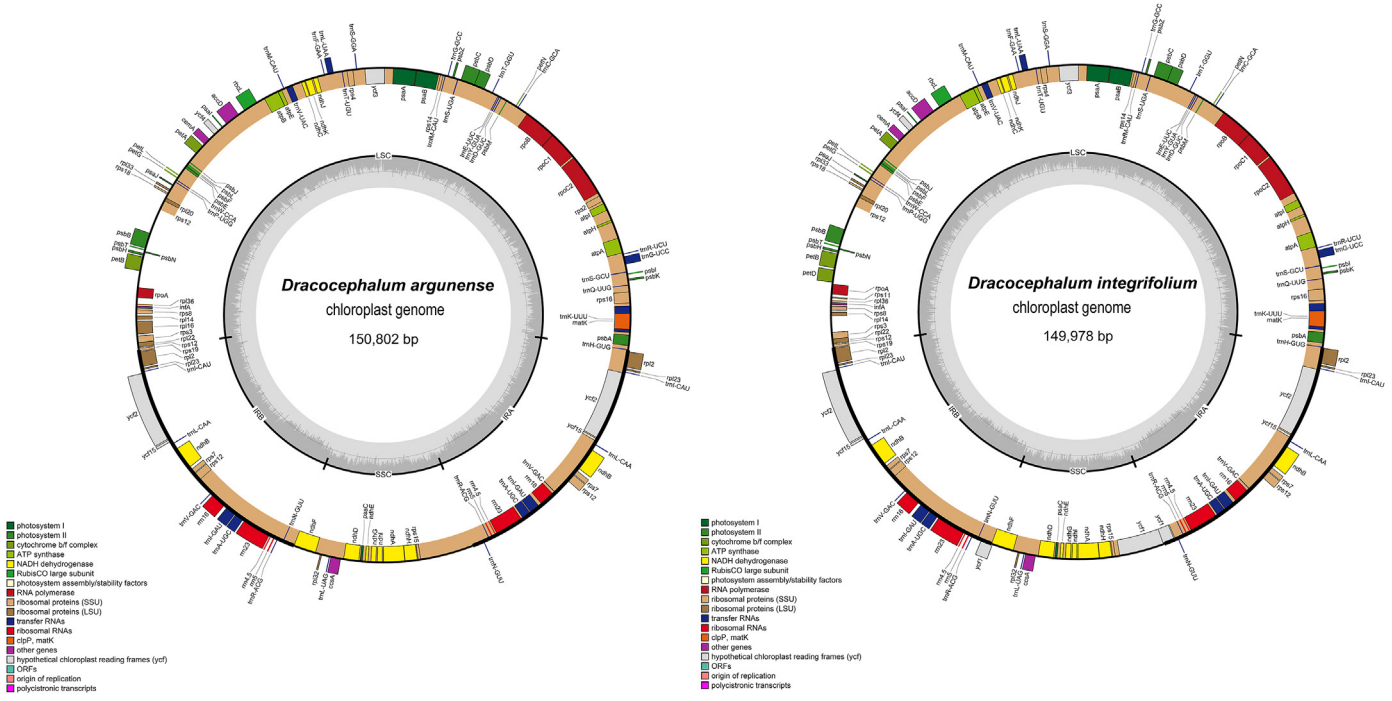

Figure 1. Complete plastome maps of *Dracocephalum argunense* and *D. integrifolium*. The genes inside and outside of the outer circle are transcribed in clockwise and counterclockwise directions, respectively. Genes belonging to different functional groups are shown in different colors. The inner circle represents different regions of the cp genome. IRA, inverted repeat region A; IRB, invert repeat region B; LSC, large single-copy region; SSC, small single-copy region. The line-chart in gray shows the GC content along the genome.

Table 1. Gene contents and functions of *Dracocephalum argunense* and *D. integrifolium*.

Category of genes	Group of genes	Name of genes
Photosynthetic	Subunits of photosystem I	*psaA*, *psaB*, *psaC*, *psaI*, *psaJ*
	Submits of photosystem II	*psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbN*, *psbT*, *psbZ*
	Subunits of NADH dehydrogenase	*ndhA*†, *ndhB*† (2x), *ndhC*, *ndhD*, *ndhE*, *ndhF* (2x), *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*
	Subunits of cytochrome b/f complex	*petA*, *petB** †, *petD** †, *petG*, *petL*, *petN*
	Subunits of ATP synthase	*atpA*, *atpB*, *atpE*, *atpH*†, *atpI*
	Large subunit of rubisco	*rbcL*
Self-replication	Proteins of large ribosomal subunit	*rpl2*† (2x), *rpl14*, *rpl16** †, *rpl20*, *rpl22*, *rpl23* (2x), *rpl32*, *rpl33*, *rpl36*
	Proteins of small ribosomal subunit	*rps2*, *rps3*, *rps4*, *rps7* (2x), *rps8*, *rps11**, *rps12*‡ (2x), *rps14*, *rps15*, *rps16*†, *rps18*, *rps19* (2x)
	Subunits of RNA polymerase	*rpoA*, *rpoB*, *rpoC1*†, *rpoC2*
	Ribosomal RNAs	*rrn23S* (2x), *rrn16S* (2x), *rrn5S* (2x), *rrn4.5S* (2x)
	Transfer RNAs	*trnA-UGC*† (2x), *trnC-GCA*, *trnD-GUC*, *trnE-UUC*, *trnF-GAA*, *trnfM-CAU*, *trnG-GCC*, *trnG-UCC*†, *trnH-GUG*, *trnI-CAU* (2x), *trnI-GAU*† (2x), *trnK-UUU*†, *trnL-CAA* (2x), *trnL-UAA*†, *trnL-UAG*, *trnM-CAU*, *trnN-GUU* (2x), *trnP-UGG*, *trnQ-UUG*, *trnR-ACG* (2x), *trnR-UCU*, *trnS-GCU*, *trnS-GGA*, *trnS-UGA*, *trnT-GGU*, *trnV-GAC* (2x), *trnV-UAC*†, *trnW-CCA*, *trnY-GUA*
Biosynthesis	Maturase	*matK*
	Protease	*clpP*‡
	Envelope membrane protein	*cemA*
	Acetyl-CoA carboxylase	*accD*
	c-type cytochrome synthesis gene	*ccsA*
	Translation initiation factor	*infA*
Unknown function	Conserved hypothetical chloroplast Reading frames	*ycf1** (2x), *ycf2* (2x), *ycf3*‡, *ycf4*, *ycf15* (2x)

* Represents the gene unique to D. integrifolium.
† gene with a single intron.
‡ gene with two introns, (2x) duplicated gene.

GAC). Nine protein-coding genes (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl16*, *rpl2*, *rpoC1*, and *rps16*) and six tRNA genes (*trnA-UGC*, *trnG-UCC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, and *trnV-UAC*) contained one intron, and three genes (*rps12*, *ycf3*, *clpP*) contained two introns (Table 1). Compared with *D. integrifolium*, five genes (*petB*, *petD*, *rpl16*, *rps11*, and *ycf1*) were absent in *D. argunense*.

Within those sampled sequences, protein-coding regions accounted for 47.4%–53.1% of the whole genome, while tRNA and rRNA regions accounted for 7.85%–7.89%. The overall guanine-cytosine (GC) content of the two plastomes was 37.7% and 37.9%, respectively.

Phylogenetic analysis

As currently defined, *Dracocephalum* includes two traditionally defined genera, i.e. *Hyssopus* L. and *Lallemantia* Fisch. & C.A. Mey. (Chen et al. 2022a). Our study highly supported the monophyly of the *Dracocephalum* (100%), and two clades are recognized

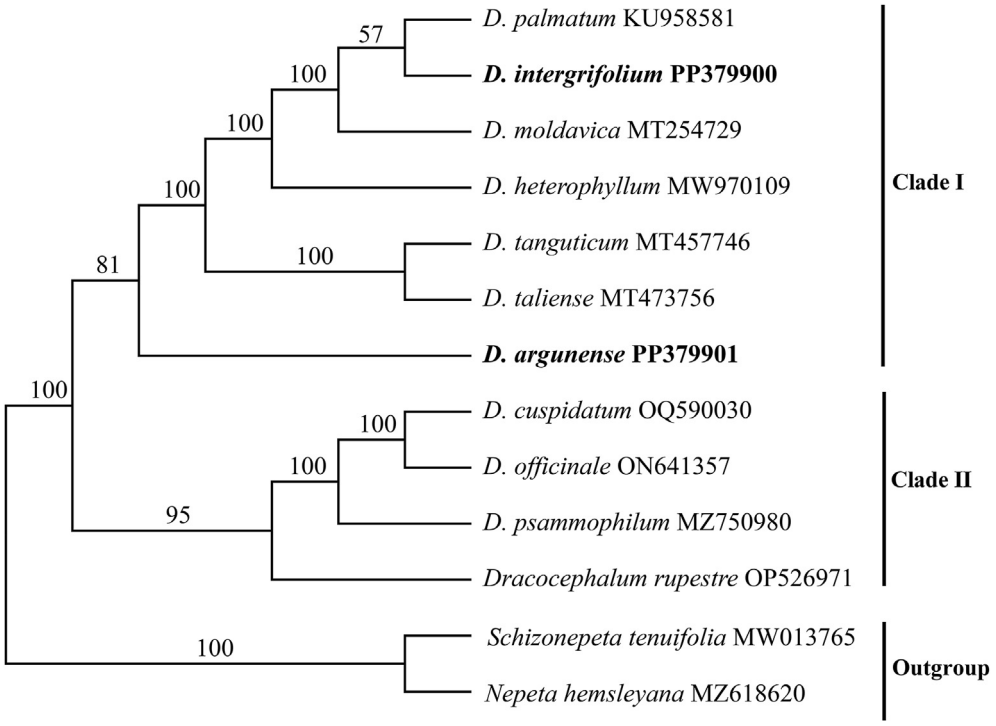

Figure 2. The best-score tree from maximum likelihood analysis of Dracocephalum based on protein-coding regions.

(Figure 2). Clade I (81%) consists of seven species, and *D. argunense* appears as a sister to the remaining six species. Another focal species, *D. integrifolium* is also nested within Clade I and is sister to *D. palmatum* Stephan ex Willd. Clade II receives high support (95%), comprising four species, *D. cuspidatum* (Boriss.) Y.P. Chen and B.T. Drew, *D. officinalis*, *D. psammophilum* C.Y. Wu & W.T. Wang, and *D. rupestre* Hance. Although only 11 speices were included for phylogenetic analysis, the relationships among them are comparable to those in the previous study (Chen et al. 2022a). In addition, the tree based on plastome sequences generally has higher support values than the previous study using four plastid DNA markers (i.e. *rpl32-trnL*, *trnL-trnF*, *ycf1*, and *ycf1-rps15*) and nuclear genes.

Dracocephalum is an economically important group, although several plastomes of the genus have been reported using chloroplast genomes to resolve phylogenetic questions within the genus have been rare. Our study presents a bettersupported phylogeny of *Dracocephalum*, demonstrating that complete plastome sequences can improve phylogenetic resolution. Now, with increasing rapid and less expensive next generation sequencing (NGS) technology continually developing, we can expect complete chloroplast genomes to least largely resolve infrageneric relationships within *Dracocephalum*.

Data availability

Plastomes of two newly sequenced species can be available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the “Ten Thousand Talents Program of Yunnan” (Top-notch Young Talents Project, No. YNWR-QNBJ-2018-279), the National Natural Science Foundation of China (Grant No. 31800168), and Yunnan Revitalization Talent Support Program “Innovation Team” Project.

References

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. [Accessed 6 March 2023].

Bautista MAC, Wen X, Ma C, et al. 2022. Complete chloroplast genome sequence of the Tibetan catnip Nepeta hemsleyana Oliver ex Prain (Lamiaceae). *Mitochondrial DNA Part B* 7:552–553.

Budantsev AL. 1987. The system of the genus *Dracocephalum* (Lamiaceae). *Botanicheskii Zhurnal* 72:260–267.

Budantsev AL, Lobova TA. 1993. A synopsis of the tribe Nepeteae (Lamiaceae). The genera *Lophanthus, Dracocephalum, Cedronella, Schizonepeta* and *Agastache*. *Botanicheskii Zhurnal* 78:106–115.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973.

Chen YP, Turdimatovich TO, Nuraliev MS, et al. 2022a. Phylogeny and biogeography of the northern temperate genus *Dracocephalum* s.l. (Lamiaceae). *Cladistics* 38: 429–451.

Chen YP, Zhao F, Paton AJ, et al. 2022b. Plastome sequences fail to resolve shallow level relationships within the rapidly radiated genus *Isodon* (Lamiaceae). *Frontiers in Plant Science* 13:985488.

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

Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1. 3. 1: Ex-panded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* 47:W59–W64.

Fathiazad F, Hamedeyazdan S. 2011. A review on *Hyssopus officinalis* L.: Composition and biological activities. *African Journal of Pharmacy and Pharmacology* 5:1959–1966.

Harley RM, Atkins S, Budantsev AL, et al. 2004. Labiatae. In: Kubitzki K, Kadereit JW, editors. *The Families and Genera of Vascular Plants*, vol. 7. Berlin, Heidelberg: Springer. pp. 167–275.

Jin JJ, Yu WB, Yang JB, et al. 2020. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology* 21:1–31.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30:772–780.

Kearse M, Moir R, Wilson A, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649.

Kim SH, Kim SH, Kim SH, et al. 2006. Action of *Dracocephalum argunense* on mast cell-mediated allergy model. *Biological & Pharmaceutical Bulletin* 29:494–498.

Li HW, Hedge IC. 1994. Lamiaceae. In: Wu CY, Raven PH, editors. *Flora of China*, vol. 17. St. Louis, MO, Beijing: Science Press, Missouri Botanical Garden Press. pp. 269–291.

Mabberley DJ. 2017. *Mabberley's Plant-Book, a Portable Dictionary of Plants, their classification and uses*. 4th ed. Cambridge: Cambridge University Press.

Qu XJ, Moore MJ, Li DZ, et al. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods* 15:1–12.

Rose JP, Wiese J, Pauley N, et al. 2023. East Asian-North American disjunctions and phylogenetic relationships within subtribe Nepetinae (Lamiaceae). *Molecular Phylogenetics and Evolution* 187:107873.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.

Wang H, Wang Y, Cheng H, et al. 2021. The complete chloroplast genome of *Schizonepeta tenuifolia* (Benth.) Briq., a traditional Chinese herb. *Mitochondrial DNA Part B* 6:907–908.

Wick RR, Schultz MB, Zobel J, et al. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 31:3350–3352.

Yao J, Zhao F, Xu Y, et al. 2020. Complete chloroplast genome sequencing and phylogenetic analysis of two *Dracocephalum* plants. *BioMed Research International* 2020:4374801.

Zhao F, Chen YP, Salmaki Y, et al. 2021. An updated tribal classification of Lamiaceae based on plastome phylogenomics. *BMC Biology* 19:2.

Zhang R, Wang YH, Jin JJ, et al. 2020. Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. *Systematic Biology* 69:613–622.

Zhou S, Wei C, Zhang C, et al. 2019. Chemical composition, phytotoxic, antimicrobial and insecticidal activity of the essential oils of *Dracocephalum integrifolium*. *Toxins* 11:598.