MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 3, 352–356
https://doi.org/10.1080/23802359.2024.2329668

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

The complete chloroplast genome sequence of an invasive plant, *Tragopogon dubius* Scopoli (asteraceae)

Yue Huang[a,b], Jingjing Cao[b], Mengxin Zhao[b], Jianying Guo[b], Jiamei Li[a] and Rui Wang[b]

[a]College of Life Science, Henan Agricultural University, Zhengzhou, China; [b]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

ABSTRACT

Tragopogon dubius Scopoli is native to Europe and western Asia and is considered an invasive plant in China. In this study, the complete chloroplast genome of *T. dubius* was obtained using high-throughput next-generation sequencing technology. The whole chloroplast genome was 153,017 bp long with a GC content of 38% and comprised 130 genes (86 protein-coding genes, 36 tRNA genes, and 8 rRNA genes). Phylogenetic analysis based on the concatenated chloroplast protein-coding sequences showed that *T. dubius* is most closely related to *Tragopogon pratensis*. This study provides valuable genetic data for further phylogenetic analysis and molecular identification of species in the genus *Tragopogon*.

ARTICLE HISTORY
Received 3 January 2024
Accepted 7 March 2024

KEYWORDS
Tragopogon dubius; complete chloroplast genome; phylogenetic analysis

Introduction

Tragopogon dubius Scopoli 1772, is a biennial native to Europe and western Asia that grows primarily in the temperate biome (Soltis et al. 2022). *T. dubius* first invaded China in the 1930s (Zhang and Han 1997), where it is now considered an invasive plant (Lv et al. 2013; Wang et al. 2022). The seeds of *T. dubius* have large umbrella-like pappuses, which promote long distance dispersal by wind and human activities (Clements et al. 1999). So far, *T. dubius* has spread to multiple provinces in North and Northwest China, including Liaoning, Henan, Hebei, Shandong, and Beijing (Lin et al. 2022). This species is prone to forming a single dominant community, resulting in a reduction in species diversity within the areas it invades. Additionally, several studies have shown that hybridization often occurs between species of the *Tragopogon* (Cook and Soltis 1999; Li and Qiang 2012; Matthews et al. 2015). Therefore, there is a potential threat of hybridization between *T. dubius* with its native congeneric species *T. porrifolius*, *T. orientalis*, and *T. pratensis* in China and for the hybrid species to become invasive (Shan et al. 2020).

Chloroplast genomes, which contain a large amount of genetic information, have been used to trace the origins of species, including invasive plant *Lepidium draba* and *Mikania micrantha* (Gaskin et al. 2005; Daniell et al. 2016; Wang et al. 2016). The chloroplast genome of *T. dubius*, can thus provide us valuable information to explore its origin, evolution, and phylogenetic relationships with other species. There is also a critical need for a complete chloroplast genome to infer the invasion history of *T. dubius* in China, thereby providing a scientific basis for preventing new invasions. However, the complete chloroplast genome sequence of *T. dubius* has not yet been reported. In this study, the chloroplast genome was successfully assembled and annotated, and the relationship of *T. dubius* with closely related species was investigated.

Materials and methods

Fresh leaves were sampled from an individual *T. dubius* plant in Zhengzhou City, Henan Province, China (34°48′13″N, 113°49′42″E). The original plant was identified as *T. dubius* by Professor Jiamei Li (Figure 1). The voucher specimen is deposited at the Herbarium of Henan Agricultural University (Jiamei Li, jiamei_li@126.com) with specimen code 2204191 (HEAC).

Total genomic DNA was extracted from fresh leaves using a modified CTAB protocol (Allen et al. 2006), and the purified DNA was sent to Novogene (Beijing, China) for genomic library construction and sequencing using the Illumina HiSeq platform (Illumina, San Diego, CA). About 6.95 Gb of sequence data were generated and used for the assembly of the chloroplast genome with GetOrganelle v1.7.7.0 (Jin et al. 2020). The complete assembled genome was annotated

CONTACT Rui Wang wangrcaas@163.com State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, 100193, China; Jiamei Li jiamei_li@126.com College of Life Science, Henan Agricultural University, Zhengzhou, 450046, China

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2329668.

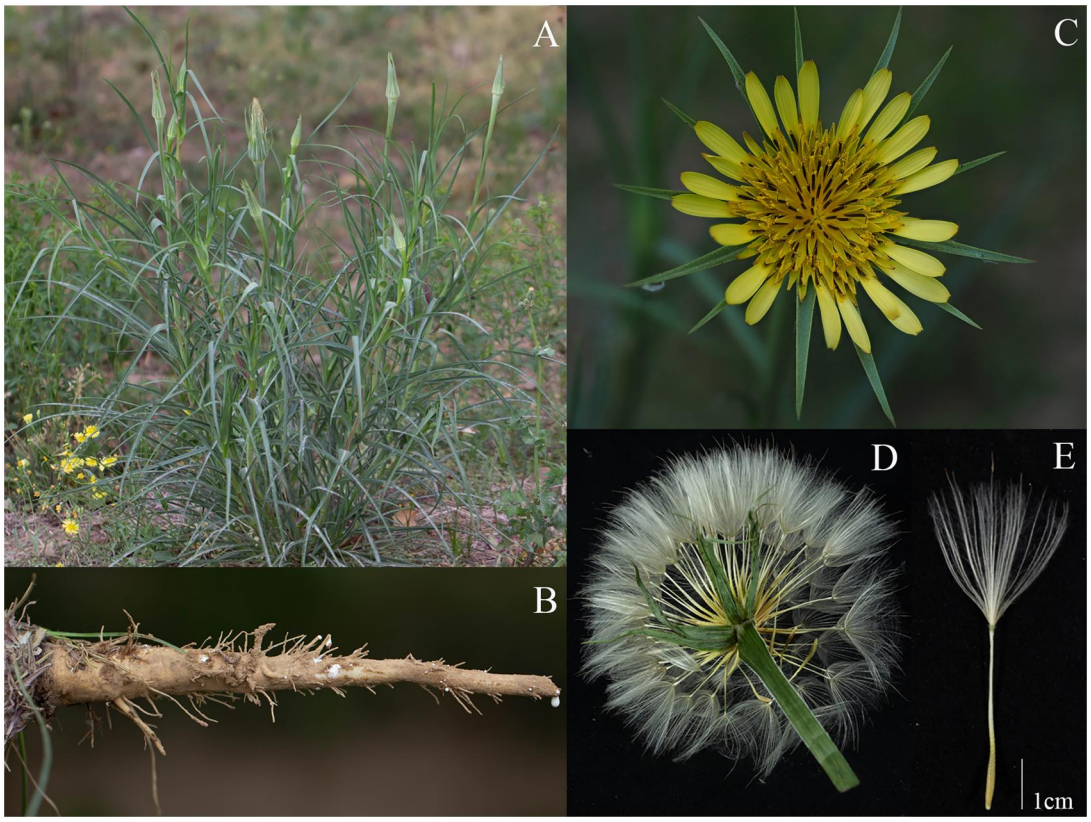

Figure 1. Photographs of *T. dubius*, taken by Yue Huang in Zhengzhou City, China, provide visual documentation of the plant's characteristics. (A) Vegetative body. *T. dubius* is a biennial herb, typically reaching a height of 40-80(-100) cm, with an erect stem that is either simple or branched from the lower or middle third, and is glabrous. Basal and lower cauline leaves are lanceolate to linear, measuring 15-40 × 0.3-0.5 cm. (B) Root. The plant features vertical roots with a well-developed caudex. (C) Flower. The flower has 8-12(-14) phyllaries, which are longer than the florets and are equal to or longer than the achenes with pappus. The ligules are yellow. (D) Infructescence. The infructescence is terminal and solitary, with numerous scabrid bristles in the pappus, which are dirty white. (E) Seed. The seeds of the outer achenes measure 2.2-3 cm; the body of seed is pale brown, ± curviform, 1.4-1.7 mm in diam, with 5 fairly well-differentiated ribs and tuberculate. The beak of seed is whitish, slender, and measures 1.2-1.6 cm.

using Plastid Genome Annotator (PGA) (Qu et al. 2019; Guo 2021), and Geneious was used to manually correct codons and gene boundaries (Kearse et al. 2012; Yu et al. 2023). A map of the whole chloroplast genome of *T. dubius* was drawn using CPGView (Liu et al. 2023) (http://www.1kmpg.cn/cpgview/).

To identify the phylogenetic position of *T. dubius*, a maximum-likelihood (ML) phylogenetic tree was constructed based on the protein-coding sequences (CDS) of *T. dubius*, 17 other species within the family Asteraceae, and *Arabidopsis thaliana* as an outgroup (Fu et al. 2016). The ML tree was constructed using MEGA11 with 1000 bootstrap replicates (Tamura et al. 2021) and JTT + G + F model.

Results

The complete chloroplast genome sequence of *T. dubius* (GenBank accession number: OR840963) was 153,017 bp with a GC content of 38% (Figure 2). The minimal and average read mapping depths for assembled genomes were 20× and 1654× (Figure S1), respectively. Similar to the chloroplast genomes of most angiosperms, the *T. dubius* chloroplast genome had a typical quadripartite structure (Zhao et al. 2020), with a large single copy (LSC) region (84,239 bp in length), a small single copy (SSC) region (18,408 bp in length), and two inverted repeat (IR) regions (each 25,185 bp in length). The chloroplast genome harbors 130 genes including 86 protein-coding genes, 36 tRNA genes, and 4 ribosomal genes. Additionally, 11 cis-splicing genes including *rps16*, *rpoC1*, *atpF*, *ycf3*, *clpP*, *petB*, *petD*, *rpl16*, *rpl2*, *ndhB*, and *ndhA* (Figure S2), and one trans-splicing gene *rps12* (Figure S3) were detected. Phylogenetic analysis showed that *T. dubius* and *T. pratensis* are most closely related to each other among species in the family Asteraceae (Figure 3).

Discussion and conclusions

Sequencing of the complete chloroplast genome of *T. dubius* revealed that its size (153,017 bp) and gene content are not significantly different from those of most chloroplast genomes or plastomes in the genus *Tragopogon*. Although there are approximately 150 species in this genus, the chloroplast genome data in NCBI are relatively limited. Therefore, additional complete chloroplast genome sequences for

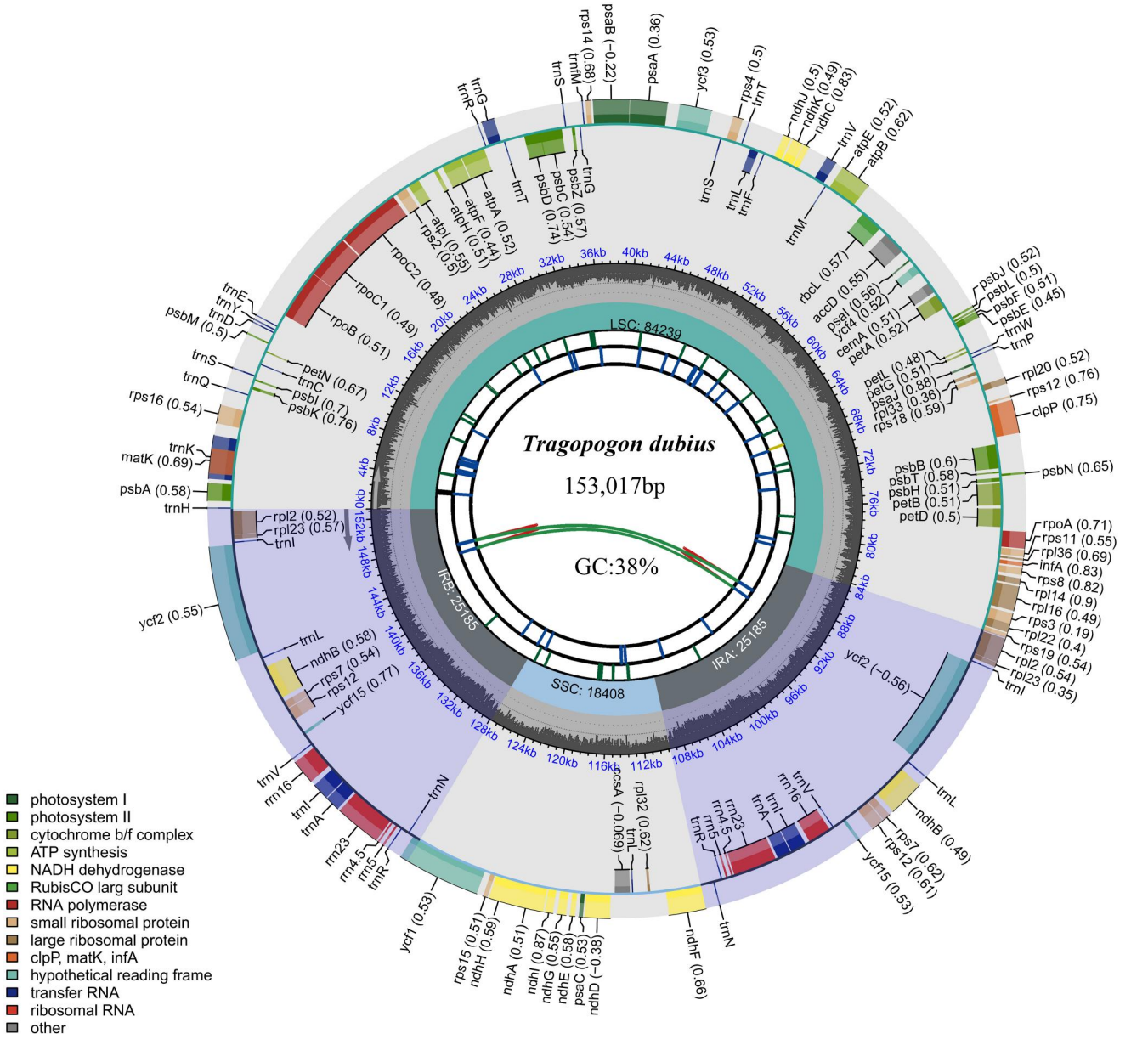

Figure 2. Map of the chloroplast genome of *T. dubius*. The map contains six tracks. From the center outward, the first track shows the dispersed repeats, consisting of direct and palindromic repeats, connected with red and green arcs. Second track: long tandem repeats (short blue bars). Third track: short tandem repeats or microsatellite sequences (short bars) with different colors representing the type of repeat. Black: complex repeat; green: repeat unit size 1; yellow: repeat unit size 2; purple: repeat unit size 3; blue: repeat unit size 4; orange: repeat unit size 5; red: repeat unit size 6. Fourth track: small single-copy (SSC), inverted repeat (IRa and IRb), and large single-copy (LSC) regions. Fifth track: GC content. Sixth track: annotated genes color-coded by functional classification (bottom left corner). Codon usage bias is displayed in parentheses after the gene name. The transcription directions for the inner and outer genes are clockwise and anticlockwise, respectively.

Tragopogon species are required to further study the evolutionary history of this genus. Our research results provide valuable genetic data for further study of the phylogeny and molecular identification of species in the genus *Tragopogon*.

The identification of the dispersal source is often a challenge when studying the diffusion process of invasive plants. Chloroplast molecular markers can be used trace the dispersal source through analysis of the haplotype composition within and between populations. The acquisition of the chloroplast genome of *T. dubius* serves as a prerequisite for the development of chloroplast molecular markers, providing a possibility for tracing the dispersal origin.

Ethical approval

The authors declare no ethical or legal violations when obtaining the study materials and performing the research.

Authors' contributions

Rui Wang, Jiamei Li, Mengxin Zhao, Jianying Guo, and Yue Huang collectively conceived and designed this study. Jiamei Li, Yue Huang, and Rui Wang collected plant samples. Yue Huang, Jingjing Cao, Jiamei Li, Mengxin Zhao, and Jianying Guo collaborated to assemble and annotate the chloroplast genome. Yue Huang, Jingjing Cao, Rui Wang, and Jiamei Li drafted the initial manuscript. All authors contributed to revising and approving the final manuscript.

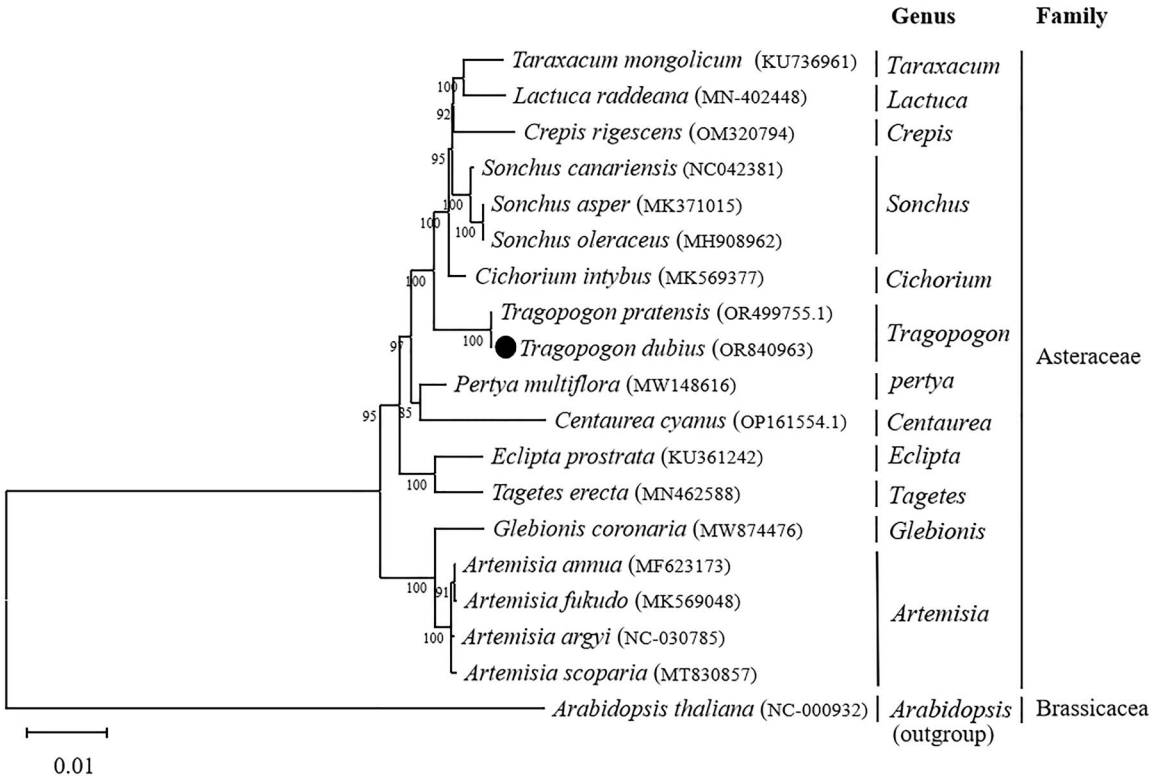

Figure 3. Phylogenetic tree based on the CDS sequences of *T. dubius*, 17 species in the asteraceae family, and one outgroup species in the brassicaceae family. The tree was generated using the ML method in MEGA11 with 1000 bootstrap replicates. Numbers on the nodes indicate bootstrap values. The following sequences were used: *Taraxacum mongolicum* (KU736961, Kim et al. 2016), *lactuca raddeana* (MN-402448, Jiang et al. 2021), *crepis rigescens* (OM320794, Li et al. 2022), *sonchus canariensis* (NC042381, Kim et al. 2019), *sonchus asper* (MK371015, Abdalla et al. 2020), *sonchus oleraceus* (MH908962, Cho et al. 2019), *cichorium intybus* (MK569377, Yang et al. 2019), *tragopogon pratensis* (OR499755.1, direct submission), *pertya multiflora* (MW148616, Liu et al. 2021), *centaurea cyanus* (OP161554.1, direct submission), *eclipta prostrata* (KU361242, Park et al. 2016), *tagetes erecta* (MN462588, direct submission), *glebionis coronaria* (MW874476, Li et al. 2021), *artemisia annua* (MF623173, Shen et al. 2017), *artemisia fukudo* (MK569048, Min et al. 2019), *artemisia argyi* (NC-030785, Chen et al. 2019), *artemisia scoparia* (MT830857, Li et al. 2020), and *Arabidopsis thaliana* (NC-000932, Sato et al. 1999).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by grants from the National Key R&D Program of China (2021YFC2600400, 2022YFC2601000), the National Nature Science Foundation of China (32272569, 31471827) and Xinyang Academy of Ecological Research Open Foundation (2023XYMS05).

ORCID

Jiamei Li http://orcid.org/0000-0003-3434-6141

Data availability statement

The genome sequence data of this study are available in the NCBI GenBank database at https://www.ncbi.nlm.nih.gov/ under accession no. OR840963. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA1050040, SRR27154907, SAMN38725175, respectively.

References

Abdalla MS, Kim SH, Kim SC. 2020. Characterization of the complete plastome sequence of perennial sowthistle, *Sonchus arvensis* (Asteraceae). Mitochondrial DNA B Resour. 5(3):2652–2654. doi:10.1080/23802359.2020.1778569.

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.

Chen HM, Jiang M, Wang LQ, You JW, Liu C. 2019. Complete plastome of *Leucanthemum maximum*, the first in genus Leucanthemum. Mitochondrial DNA B Resour. 5(1):19–20. doi:10.1080/23802359.2019.1693922.

Cho MS, Kim JH, Kim CS, Mejías JA, Kim SC. 2019. Sow thistle chloroplast genomes: insights into the plastome evolution and relationship of two weedy species, *Sonchus asper* and *Sonchus oleraceus* (Asteraceae). Genes (Basel). 10(11):881. doi:10.3390/genes10110881.

Clements DR, Upadhyaya MK, J. Bos S. 1999. The biology of canadian weeds. 110. *Tragopogon dubius Scop., Tragopogon pratensis L.,* and *Tragopogon porrifolius L.* Can J Plant Sci. 79(1):153–163. doi:10.4141/P98-007.

Cook LM, Soltis PS. 1999. Mating systems of diploid and allotetraploid populations of Tragopogon (Asteraceae). II. Artificial populations. Heredity (Edinb). 84 (Pt 4)(3):410–415. doi:10.1046/j.1365-2540.2000.00654.x.

Daniell H, Lin CS, Yu M, Chang WJ. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol. 17(1):134. doi:10.1186/s13059-016-1004-2.

Fu ZX, Jiao BH, Nie B, Zhang GJ, Gao TG. 2016. A comprehensive generic-level phylogeny of the sunflower family: implications for the systematics of Chinese Asteraceae. J Syt Evolut. 54(4):416–437. doi:10.1111/jse.12216.

Gaskin JF, Zhang DY, Bon MC. 2005. Invasion of *Lepidium draba* (Brassicaceae) in the western United States: distributions and origins of chloroplast DNA haplotypes. Mol Ecol. 14(8):2331–2341. doi:10.1111/j.1365-294X.2005.02589.x.

Guo YH. 2021. The complete chloroplast genome sequence of *Prunus discadenia* (Rosaceae), a species with great ornamental value. Mitochondrial DNA B Resour. 6(3):936–937. doi:10.1080/23802359.2021.1888336.

Jiang M, Li YN, Chen HM, Wang B, Liu C. 2021. Comparative and phylogenetic analysis of the complete chloroplast genome sequences of *Lactuca raddeana* and *Lactuca sativa*. Mitochondrial DNA B Resour. 6(4):1498–1506. doi:10.1080/23802359.2021.1911700.

Jin JJ, Yu WB, Yang JB, Song Y, DePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21(1):241. doi:10.1186/s13059-020-02154-5.

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.

Kim HB, Cho MS, Kim SH, Yang J, Kim SC. 2019. Characterization of the complete chloroplast genome of the woody sow-thistle, *Sonchus leptocephalus* (Asteraceae) endemic to the Canary Islands. Mitochondrial DNA B Resour. 4(2):3388–3389. doi:10.1080/23802359.2019.1674719.

Kim JK, Park JY, Lee YS, Woo SM, Park HS, Lee SC, Kang JH, Lee TJ, Sung SH, Yang TJ. 2016. The complete chloroplast genomes of two Taraxacum species, *T. platycarpum* Dahlst. and *T. mongolicum* Hand.-Mazz. (Asteraceae). Mitochondrial DNA B Resour. 1(1):412–413. doi:10.1080/23802359.2016.1176881.

Li J, Qiang S. 2012. Polyploidization, one of the driving forces for weed origin and evolution. J Nanj Agricult Univ. 35(5):64–76.

Li JL, Jiang M, Chen HM, Yu J, Liu C. 2020. Intraspecific variations among the chloroplast genomes of *Artemisia scoparia* (asteraceae) from Pakistan and China. Mitochondrial DNA Part B Resources. 5(3):3182–3184. doi:10.1080/23802359.2020.1810167.

Li P, Fang X, Liang X. 2022. The complete chloroplast genome of *Crepis rigescens* (Cichorioideae), a traditional Chinese medicinal plant. Mitochondrial DNA B Resour. 7(5):729–731. doi:10.1080/23802359.2022.2068978.

Li WQ, Jiang SF, Wang J, Yu YJ, Zhu ZJ. 2021. Complete chloroplast genome and phylogenetic analysis of *Glebionis coronaria* (L.) Cass. ex Spach (Asteraceae). Mitochondrial DNA B Resour. 6(9):2693–2694. doi:10.1080/23802359.2021.1966331.

Lin QW, Xiao C, Ma JS. 2022. A dataset on catalogue of alien plants in China. Biodiver Sci. 30(5):22127. doi:10.17520/biods.2022127.

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 Y, Zhang HJ, Wu Q, Huang WX, Wu T, Guan M, Jiang M. 2021. Characterization and phylogenetic analysis of the complete chloroplast genome sequence of *Pertya multiflora* (Asteraceae), a rare and endangered wild plant species in China. Mitochondrial DNA B Resour. 6(7):1934–1936. doi:10.1080/23802359.2021.1936671.

Lv YF, Zhang JL, Bian Y, Liu RS. 2013. A new record of invasive weed in North China-*Tragopogon dubius* Scop. J Beij Univ Agricult. 28(4):3–4.

Matthews A, Emelianova K, Hatimy AA, Chester M, Pellicer J, Ahmad KS, Guignard MS, Rouhan G, Soltis DE, Soltis PS, et al. 2015. 250 years of hybridization between two biennial herb species without speciation. AoB Plants. 7:plv081. doi:10.1093/aobpla/plv081.

Min JH, Park JS, Kim YS, Kwon WC. 2019. The complete chloroplast genome of *Artemisia fukudo* Makino (Asteraceae): providing insight of intraspecies variations. Mitochondrial DNA Part B Resources. 4(1):1510–1512. doi:10.1080/23802359.2019.1601044.

Park JY, Lee YS, Kim JK, Lee HO, Park HS, Lee SC, Kang JH, Lee TJ, Sung SH, Yang TJ. 2016. The complete chloroplast genome of *Eclipta prostrata* L. (Asteraceae). Mitochondrial DNA B Resour. 1(1):414–415. doi:10.1080/23802359.2016.1176882.

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.

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. 1999. Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res. 6(5):283–290. doi:10.1093/dnares/6.5.283.

Shan SC, Boatwright JL, Liu XX, Chanderbali AS, Fu CN, Soltis PS, Soltis DE. 2020. Transcriptome dynamics of the inflorescence in reciprocally formed allopolyploid *Tragopogon miscellus* (Asteraceae). Front Genet. 11:888. doi:10.3389/fgene.2020.00888.

Shen XF, Wu ML, Liao BS, Liu ZX, Bai R, Xiao SM, Li XW, Zhang BL, Xu J, Chen SL. 2017. Complete Chloroplast Genome Sequence and Phylogenetic Analysis of the Medicinal Plant *Artemisia annua*. Molecules. 22(8):1330. doi:10.3390/molecules22081330.

Soltis DE, Mavrodiev EV, Gitzendanner MA, Alexeev YE, Godden GT, Soltis PS. 2022. *Tragopogon dubius*: multiple introductions to North America and the formation of the New World tetraploids. J Int Associat Plant Taxon. 71(6):1287–1298. doi:10.1002/tax.12743.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 38(7):3022–3027. doi:10.1093/molbev/msab120.

Wang FQ, Mi ZL, Huang Y, Shi ZY, Zeng XN, Wang PH, Zhang MF, Li JM. 2022. Four newly recorded invasive species of Asteraceae in Henan Province. J Henan Agricult Univ. 56(5):816–821. doi:10.16445/j.cnki.1000-2340.20220830.001.

Wang T, Wang Z, Chen GP, Wang CB, Su YJ. 2016. Invasive chloroplast population genetics of *Mikania micrantha* in China: no Local adaptation and negative correlation between diversity and geographic distance. Front Plant Sci. 7:1426. doi:10.3389/fpls.2016.01426.

Yang SP, Sun XM, Wang LH, Jiang XT, Zhong QW. 2019. The complete chloroplast genome sequence of Chicory (*Cichorium intybus* L.). Mitochondrial DNA Part B Resources. 4(1):1533–1534. doi:10.1080/23802359.2019.1601524.

Yu X, Song YR, Zhao ZN. 2023. The complete chloroplast genome of *Elsholtzia fruticosa* (D. Don) Rehd. (Labiatae), an ornamental plant with high medicinal value. Mitochondrial DNA B Resour. 8(3):336–341. doi:10.1080/23802359.2023.2183069.

Zhang SM, Han QZ. 1997. A preliminary study on exotic plants in Dalian area. J Liaon Normal Univ (Nat Sci Ed). 20(4):62–69.

Zhao KH, Li LQ, Lu YZ, Yang JB, Zhang ZR, Zhao FY, Quan H, Ma XJ, Liao ZH, Lan XZ. 2020. Characterization and comparative analysis of two rheum complete chloroplast genomes. Biomed Res Int. 2020:6490164. doi:10.1155/2020/6490164.