


## The complete chloroplast genome sequences of an endemic species of Urticaceae (*Debregeasia hekouensis*)

Zhao-Yang Sun<sup>a</sup>, Jie Liu<sup>b,c</sup>, Moses C. Wambulwa<sup>b,d</sup>, Zhong-Hu Li<sup>a</sup> and Zeng-Yuan Wu<sup>b</sup> 

<sup>a</sup>Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an, China; <sup>b</sup>Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China; <sup>c</sup>CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China; <sup>d</sup>Department of Life Sciences, South Eastern Kenya University, Kitui, Kenya

### ABSTRACT

*Debregeasia hekouensis*, which belongs to the nettle family (Urticaceae), is a local endemic species in Hekou County, Yunnan Province, China. To provide a basis for the development of effective molecular markers for its conservation, we sequenced the chloroplast (cp) genome of *D. hekouensis* in the present study. The total length of the chloroplast (cp) genome was 155,941 bp, and exhibited a typical quadripartite structure, with a pair of IRs (inverted repeats; 25,664 bp in length) being separated by a small single copy (SSC) region of 19,085 bp and a large single copy (LSC) region of 85,528 bp. The cp genome contained a total of 112 genes, including 78 protein-coding genes, 30 tRNA genes, and 4 rRNA genes. The GC content of the entire cp genome, LSC region, SSC region, and IR region was 36.3%, 34.0%, 29.4%, and 42.7%, respectively. Phylogenetic analysis indicated that *D. hekouensis* is evolutionarily closer to *Debregeasia orientalis* and *Debregeasia squamata*.

### ARTICLE HISTORY

Received 15 June 2021  
Accepted 9 October 2021

### KEYWORDS


*Debregeasia hekouensis*;  
complete chloroplast  
genome; phylogenetic  
analysis; Urticaceae

*Debregeasia* Gaudich. comprises eight species (*D. longifolia*, *D. elliptica*, *D. orientalis*, *D. saeneb*, *D. squamata*, *D. wallichiana*, *D. australis*, and *D. hekouensis*) with a worldwide distribution, but mainly found in subtropical or tropical areas of eastern Asia (Chen et al. 2003; Wilmot-Dear and Friis 2012; Wu et al. 2018). *Debregeasia hekouensis*, first described in 2016 (Wang 2016), is an endemic species narrowly distributed in the fragmented tropical rain forest along the Nanxi River of Hekou County, Yunnan Province, China. Although phylogenetic (Wu et al. 2013, 2018) and phylogenomic (Wang et al. 2020) relationships within *Debregeasia* have been tested in previous studies, *D. hekouensis* was not included. In the current study, we reported the complete chloroplast genome (cpDNA) sequence of *D. hekouensis*, and explored its phylogenetic relationships with related species in Rosales. Our results will significantly contribute to the molecular systematics and conservation genetics of this species.

Fresh, young, and clean leaf materials of *D. hekouensis* were collected from Nanxi Town, Hekou County, Honghe Prefecture, Yunnan Province, China (22.67°N; 103.94°E). Voucher specimens (No. HKD03) were deposited at the herbarium of the Kunming Institute of Botany, China (KUN) (contact person: Zeng-Yuan Wu, [wuzengyuan@mail.kib.ac.cn](mailto:wuzengyuan@mail.kib.ac.cn)). Total genomic DNA was extracted using the CTAB method

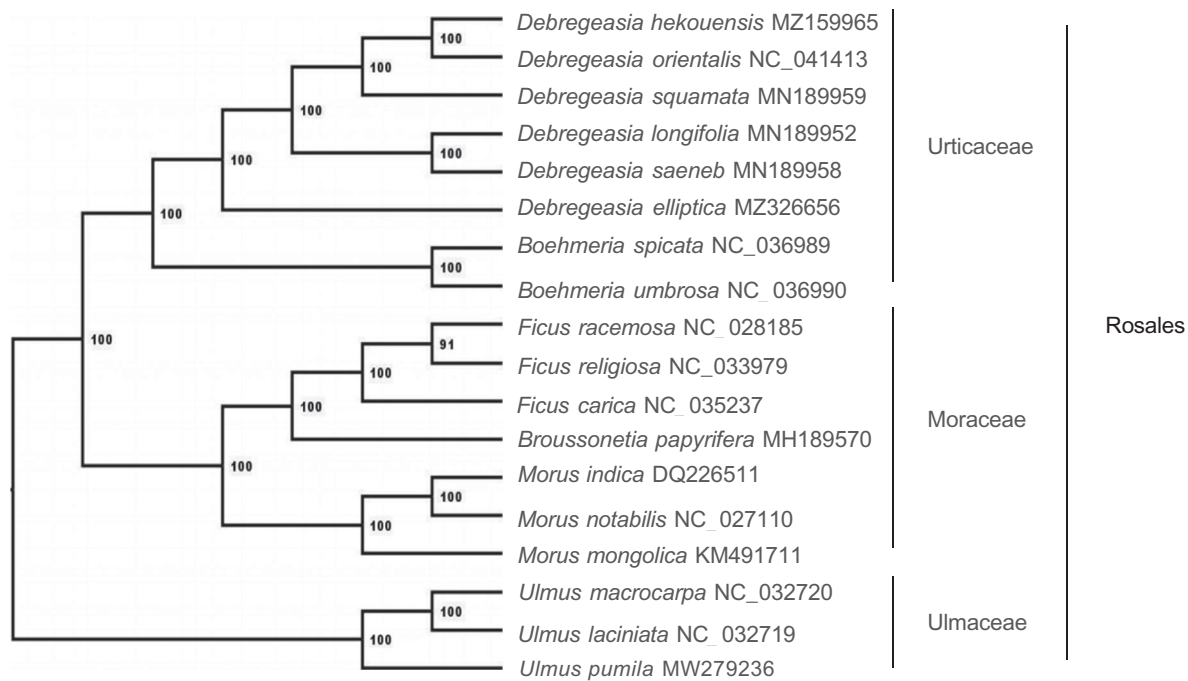
(Doyle and Doyle 1987). Library construction was performed with NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs) following the manufacturer's instructions. Sequencing was implemented using the Illumina HiSeq X Ten platform, which produced 150 bp paired-end reads. The expected sequencing quantity was ca. 4 Gigabyte. The raw reads were filtered to obtain clean reads, which were subsequently, assembled into circular contigs using GetOrganelle toolkit (Jin et al. 2020). Finally, the cpDNA was annotated using geneious v.9.0.2 software package (<https://www.geneious.com/>). The reference genomes used for assembly and annotation are the complete chloroplast genome of *D. orientalis* (GenBank accession: NC\_041413).

Raw reads of *D. hekouensis* were deposited in the NCBI Sequence Read Archive (SRA: SRX10910957) and the final annotated chloroplast genome sequence was deposited in NCBI GenBank (accession number: MZ159965). The total length of the chloroplast genome was 155,941 bp, and had a typical quadripartite structure, in which a pair of IRs (inverted repeats) of 25,664 bp was separated by a small single copy (SSC) region of 19,085 bp and a large single copy (LSC) region of 85,528 bp. The GC content of the entire cp genome, LSC region, SSC region, and IR region was 36.3%, 34.0%, 29.4%, and 42.7%, respectively. The assembled cp genome

**CONTACT** Zhong-Hu Li  [lizhonghu@nwu.edu.cn](mailto:lizhonghu@nwu.edu.cn)  Key Laboratory of Resource Biology and Biotechnology in Western China (Ministry of Education), College of Life Sciences, Northwest University, Xi'an 710069, China; Zeng-Yuan Wu  [wuzengyuan@mail.kib.ac.cn](mailto:wuzengyuan@mail.kib.ac.cn)  Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

© 2021 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 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** Maximum Likelihood (ML) tree of 18 species within Rosales based on the complete chloroplast genome sequences. Bootstrap support values are shown next to the nodes.

contained a total of 112 genes, including 78 protein-coding genes, 30 tRNA genes, and 4 rRNA genes. Of these, a total of 17 genes were duplicated in the inverted repeat regions, 16 contained one intron each, while three (*ycf3*, *rps12*, and *clpP*) had two introns each.

A total of 18 cp genome sequences of Rosales species were downloaded from the NCBI database, and subjected to phylogenetic analysis. After using MAFFT for aligning (Katoh and Standley 2013), the ModelFinder (Subha et al. 2017) within the IQTREE (Nguyen et al. 2015) was used to determine the best-fitting model for Maximum likelihood (ML) analysis. The bootstrap value based on 1,000 replicates for constructing the phylogenetic tree. Based on the analysis of the 18 species, our results showed that *D. orientalis* and *D. squamata* are the closest relatives of *D. hekouensis* (Figure 1).

The newly reported chloroplast genome here would be beneficial to developing molecular markers in *Debregeasia*, which, in turn, will lay a foundation for conservation genetic studies of this species.

## Acknowledgment

Special thanks go to Mr. Qi Chen for assistance with data analysis. We would like to also thank the Laboratory of Molecular Biology at the Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, for providing the laboratory bench space necessary for the molecular work.

## Disclosure statement

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

## Funding

This study was supported by the National Natural Science Foundation of China [41971071, 31970356], the Key Research Program of Frontier Sciences, CAS [ZDBS-LY-7001], the Top-notch Young Talents Project of Yunnan Provincial 'Ten Thousand Talents Program' [YNWR-QNBJ-2018-46]. Zeng-Yuan Wu was supported by CAS' Youth Innovation Promotion Association [2019385], and Top-notch Young Talents Project of Yunnan Provincial 'Ten Thousand Talents Program'. Moses Wambulwa was supported by the Postdoctoral International Exchange Program of the Office of China Postdoctoral Council, and the Postdoctoral Targeted Funding and Postdoctoral Research Fund of Yunnan Province.

## ORCID

Zeng-Yuan Wu  <http://orcid.org/0000-0003-4652-0194>

## Data availability statement

The chloroplast genome sequences data of *D. hekouensis* that support the findings of this study are deposited in GenBank at <https://www.ncbi.nlm.nih.gov/> under the accession no. MZ159965. The associated BioProject, SRA, and Bio-Sample numbers of *D. hekouensis* are PRJNA730369, SRX10910957, and SAMN19229652, respectively.

## References

- Chen CJ, Lin Q, Friis I, Wilmot-Dear CM, Monro AK. 2003. Urticaceae. In: Wu ZY, Raven PH, editors. Flora of China. Beijing (China): Science Press; St Louis (MO): Missouri Botanical Garden Press; p. 76–189.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin*. 19:11–15.
- Jin JJ, Yu WB, Yang JB, Song Y, Claude W, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol*. 21(1):241–231.

- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Nguyen LT, Heiko AS, Arndt H, Bui QM. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Subha K, Bui QM, Thomas KFW, Arndt H, Lars SJ. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 14: 587–589.
- Wang RN, Milne RI, Du XY, Liu J, Wu ZY. 2020. Characteristics and mutational hotspots of plastomes in *Debregeasia* (Urticaceae). *Front Genet.* 11(729):729–713.
- Wang WT. 2016. Two new species of Urticaceae from China. *Bull Bot Res.* 36(4):481–483.
- Wilmot-Dear CM, Friis I. 2012. *Debregeasia australis* Sp. Nov. (Urticaceae), with a new synopsis of and a new key to the genus. *Edinburgh J Bot.* 69(2):301–311.
- Wu ZY, Liu J, Provan J, Wang H, Chen CJ, Cadotte MW, Luo YH, Amorim BS, Li DZ, Milne RI. 2018. Testing Darwin's transoceanic dispersal hypothesis for the inland nettle family (Urticaceae). *Ecol Lett.* 21(10):1515–1529.
- Wu ZY, Monro AK, Milne RI, Wang H, Yi TS, Liu J, Li DZ. 2013. Molecular phylogeny of the nettle family (Urticaceae) inferred from multiple loci of three genomes and extensive generic sampling. *Mol Phylogenet Evol.* 69(3):814–827.