

MITOGENOME ANNOUNCEMENT



# The complete mitochondrial genome of a Chinese endemic species, *Paraprenanthes diversifolia* (Vaniot) N. Kilian (Cichorieae; Asteraceae)

Shuoming Fu, Ran Chu, Xin Xu, Haoyuan Chen, Yuhua Shi, Shixin Zhu and Zhen Wei

School of Life Sciences, Zhengzhou University, Zhengzhou, China

## ABSTRACT

The first complete mitochondrial genome (mt) of *Paraprenanthes diversifolia* (Vaniot) N. Kilian (Cichorieae; Asteraceae) was sequenced and successfully assembled in this study. The full length of the mt genome is 360,751 bp, containing 73 genes (33 protein-coding genes, 29 tRNA genes, 6 rRNA genes, and 5 protein-coding genes containing internal stop codons). There are two pairs of long (over 1000 bp) repeat regions in the mt genome of *P. diversifolia*. The phylogenetic analysis indicated that *P. diversifolia* has a close relationship with other Lactucinae species.

## ARTICLE HISTORY

Received 10 December 2019  
Accepted 1 February 2020

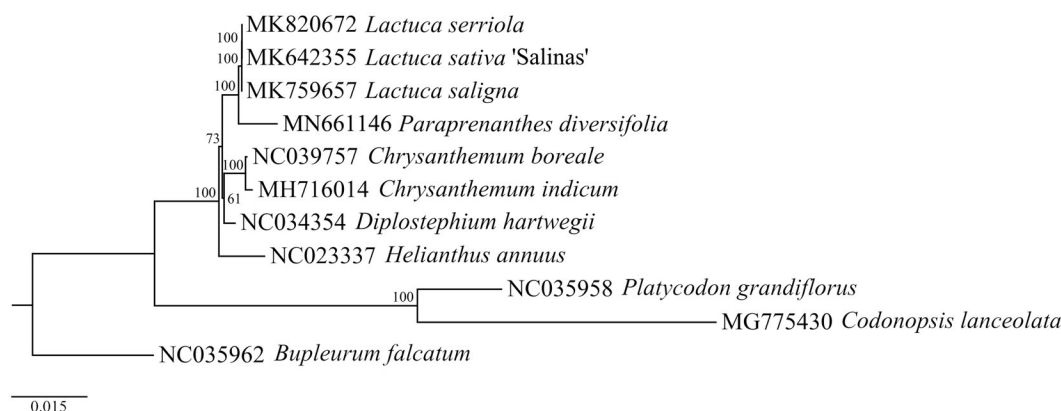
## KEYWORDS

Mitochondrial genome;  
*Paraprenanthes diversifolia*;  
Asteraceae; Cichorieae;  
phylogenetics

*Paraprenanthes diversifolia* (Vaniot) N. Kilian belongs to Lactucinae (Cichorieae, Asteraceae) and is an endemic herbaceous plant mainly distributed in south-central and south-east of China (Shih 1988). In order to identify the taxonomic position of *P. diversifolia* and provide phylogenetic evidence for further study of Lactucinae based on mt sequences, we sequenced and assembled the complete mitochondrial (mt) genome of *P. diversifolia*.

The *P. diversifolia* specimen (ZZU2012-1845) was collected from Chongqing, China (N29°2'122", E107°8'255") and stored at the herbarium of School of Life Sciences, Zhengzhou University, Zhengzhou, China. The total DNA of *P. diversifolia* was extracted from the leaves of specimen using modified CTAB method (Wei et al. 2017) and then was purified by Wizard DNA clean-up system (Promega Corp., Madison, USA).

The DNA sample was sequenced using Illumina HiSeq X Ten platform at Novogene Biotech Co. (Beijing, China). About 10.7 GB paired-end clean reads were generated and assembled using NOVOPlasty (Dierckxsens et al. 2016) software with reference sequences *Diplostephium hartwegii* Hieron. (NCBI accession no. NC\_034354), *Chrysanthemum boreale* B. Fedtsch. (NC\_039757) and *Helianthus annuus* L. (NC\_023337). The Kmer number was estimated and selected as 35 using KmerGenie (Chikhi and Medvedev 2014). Other parameters were default settings. The assembled contigs were annotated and adjusted manually in Geneious Prime 2019.1.1 (<https://www.geneious.com>) with previously mentioned references plus *Lactuca saligna* L. (MK759657), *L. serriola* L. (MK820672) and *L. sativa* 'Salinas' L. (MK642355). All borders of 11 contigs were verified by PCR amplification and



**Figure 1.** Phylogenetic tree of *P. diversifolia* and other 7 Asteraceae species based on 23 mt protein-coding genes using RAxML method. *Bupleurum falcatum* L. (Bupleureae, Apiaceae), *Platycodon grandiflorus* (Jacq.) A. DC. (Campanuloideae, Campanulaceae) and *Codonopsis lanceolata* (Siebold & Zucc.) Benth. & Hook.f. ex Trautv. (Campanuloideae, Campanulaceae) were set as outgroups. The numbers above the branches are bootstrap values.

**CONTACT** Zhen Wei ✉ [weizhen@zzu.edu.cn](mailto:weizhen@zzu.edu.cn) School of Life Sciences, Zhengzhou University, Zhengzhou, China

© 2020 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.

Sanger sequencing. Repeat regions (over 100 bp) were detected using Repeat Finder v1.0 in Geneious Prime.

The complete mt genome of *P. diversifolia* (NCBI accession no. MN661146) is circular and 360,751 bp in length with a GC content of 45.2%. The mt genome contains 73 genes, including 33 protein-coding genes, 29 tRNA genes, 6 rRNA genes and 5 protein-coding genes containing internal stop codons. The protein-coding genes consist of 9 NADH dehydrogenase genes, 1 cytochrome b gene, 3 cytochrome c oxidase genes, 4 cytochrome c biogenesis genes, 5 ATP synthase genes, 8 ribosomal proteins genes and 3 other genes. Additionally, there are four protein-coding genes (*nad4*, *ccmFc*, *rps3* and *cox2*) having one intron and three (*nad1*, *nad2* and *nad7*) having four introns. There are 26 repeat regions (over 100 bp) in the mt genome of *P. diversifolia*, including two pairs of repeat regions over 1000 bp. The sequence (containing *dpo* gene, position 173,401–180,661 bp) between two inverted repeat regions was found also having its reverse complementary sequence existing in the mt genome.

Phylogenetic relationships among *P. diversifolia* and other 7 Asteraceae species were studied based on 23 mt protein-coding genes. The phylogenetic tree was constructed using RAxML-HPC2 on XSEDE from CIPRES Science Gateway website (Miller et al. 2010). The results showed that *P. diversifolia* is closely related to *Lactuca* species (Cichorieae). Other Asteraceae species from Anthemideae, Astereae and Heliantheae are sister groups of the species of Cichorieae. The phylogenetic analysis supported previous research based on the nrITS and chloroplast DNA (Gao et al. 2009; Kilian et al. 2017) (Figure 1).

## Disclosure statement

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

## Funding

This work was supported by [grant 31700175] of The National Natural Science Foundation of China (to ZW).

## References

- Chikhi R, Medvedev P. 2014. Informed and automated k-mer size selection for genome assembly. *Bioinformatics*. 30(1):31–37.
- Dierckxsens N, Mardulyn P, Smits G. 2016. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res*. 45(4):e18–e18.
- Gao TG, Wang W, Bayer RJ, Li DZ. 2009. Systematic position of the enigmatic genus *Sheareria* (Asteraceae)-evidence from molecular, morphological and cytological data. *Taxon*. 58(3):769–780.
- Kilian N, Sennikov A, Wang ZH, Gemeinholzer B, Zhang JW. 2017. Sub-paratethyan origin and middle to late miocene principal diversification of the *Lactucinae* (Compositae: Cichorieae) inferred from molecular phylogenetics, divergence-dating and biogeographic analysis. *Taxon*. 66(3):675–703.
- Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 14 Nov. 2010, New Orleans, LA. p. 1–8.
- Shih C. 1988. Revision of *Lactuca* L. and two new genera of tribe Lactuceae (Compositae) on the mainland of Asia. *Acta Phytotaxon Sin*. 26:382–393.
- Wei Z, Zhu SX, Van den Berg RG, Bakker FT, Schranz ME. 2017. Phylogenetic relationships within *Lactuca* L. (Asteraceae), including African species, based on chloroplast DNA sequence comparisons. *Genet Resour Crop Evol*. 64(1):55–71.