

MITOGENOME ANNOUNCEMENT



The complete chloroplast genome of *Hylotelephium erythrostictum* (Miq.) H. Ohba (Crassulaceae)

Jing Chen^a, Hui Zhou^b and Wei Gong^b

^aCollege of Life Science, Zhejiang Chinese Medical University, Hangzhou, PR China; ^bCollege of Life Sciences, South China Agricultural University, Guangzhou, PR China

ABSTRACT

The species *Hylotelephium erythrostictum* is well known as an ornamental plant in China. Here, we report the complete chloroplast genome (cp) of *H. erythrostictum* using the next-generation sequencing. It shows a total length of 151,707 bp with typical quadripartite structure. The genome is composed of the large single copy region (LSC) of 83,070 bp, the small single copy region (SSC) of 17,018 bp, and two separated inverted regions (IRs) of 25,793 bp, respectively. It contains 134 genes, including 80 protein-coding genes (PCGs), 8 rRNA genes, and 37 tRNA genes. The overall GC content of the cp genome is 37.7%. Maximum likelihood (ML) tree based on ten complete chloroplast genomes of Crassulaceae and one outgroup species suggested a monophyly formed by *H. erythrostictum* together with *Hylotelephium ewersii*, which demonstrate a comparably closer phylogenetic relationship. The complete chloroplast genome of this *H. erythrostictum* provides valuable information and further phylogenetic reconstruction of the Crassulaceae family.

ARTICLE HISTORY

Received 27 July 2021
Accepted 27 January 2022

KEYWORDS

Hylotelephium erythrostictum; chloroplast genome; ornamental plant; Crassulaceae

The genus *Hylotelephium*, belonging to the family of Crassulaceae, shows a wide geographical distribution range in Euro-Asia and North America. *Hylotelephium erythrostictum* (Miq.) H. Ohba 1977 is a perennial succulent herb in this genus. This species has been widely used in urban greening and ecological landscaping in China due to its strong stress-resistance and highly ornamental value. In this study, we report the first complete chloroplast (cp) genome of *H. erythrostictum*, on the purpose of further phylogenetic reconstruction of the Crassulaceae family.

Before we carried out the project, we got ethical approval and permission from the Natural Reserve of Tianmu Mountain. Field studies have been carried out in accordance with guidelines and comply with local legislation. One individual was collected from Tianmu Mountain, Zhejiang Province (N23.08°, E113.22°). A specimen was deposited at the herbarium of College of Life Science, Zhejiang Chinese Medical University (Jing Chen; cj00123@zcmu.edu.cn) under the voucher number 20150720-051-021. Tender leaves were sampled and instantly put into the silica gel for drying and preservation (Jing Chen; cj00123@zcmu.edu.cn). Total genomic DNA was extracted by using a modified CTAB method (Doyle and Doyle 1987). The paired-end (2 × 150 bp) library was sequenced by Illumina PE150 at Novogene Co. Ltd (Beijing, China). A total of 2.19 Gb clean reads were obtained after removing low quality reads and adaptor sequences. Using GetOrganelle (Jin et al. 2020) and SPAdes (Bankevich et al. 2012), we assembled the complete cp genome of

H. erythrostictum followed by manual adjustment and annotated it by the aid of Geneious prime 2019.2.1 (Kearse et al. 2012), using default parameters to predict protein-coding genes (PCGs) and rRNA genes. The tRNA genes were annotated on ARAGORN (Laslett and Canback 2004). The cp genomes of nine species of Crassulaceae were downloaded from GenBank database, including *Aeonium arboreum* (MW206792), *Cotyledon tomentosa* (MW206793), *Crassula perforata* (MW206794), *Graptopetalum amethystinum* (MW206795), *Hylotelephium ewersii* (MN794014), *Kalanchoe fedtschenkoi* (MW206796), *Kalanchoe tomentosa* (MN794319), *Rosularia alpestris* (MN794333), and *Sedum lineare* (MT755626). *Eucommia ulmoides* (KU204775) of Eucommiaceae was used as an outgroup. Totally, 11 cp genomes were aligned with online software MAFFT on CIPRES (<https://www.phylo.org>, Miller et al. 2015). Maximum likelihood (ML) analyses were performed by using RAXML-HPC version 8.2.10 on XSEDE (<https://www.phylo.org>) with 1000 bootstrap replicates and the substitution GTR + I + G model (Stamatakis 2014), the latter of which was determined by the Bayesian information criterion (BIC) in jModeltest version 2.1.10 (Darriba et al. 2012).

The *H. erythrostictum* cp genome has been deposited in GenBank (Accession No.: MZ519882). The total length is 151,707 bp with the typical quadripartite structure. It consists of a pair of inverted regions (IRs) of 25,793 bp separated by large single copy region (LSC) of 83,070 bp and small single copy region (SSC) of 17,018 bp, respectively. The overall GC

CONTACT Wei Gong ✉ wong@scau.edu.cn College of Life Sciences, South China Agricultural University, Guangzhou, PR China

© 2022 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.

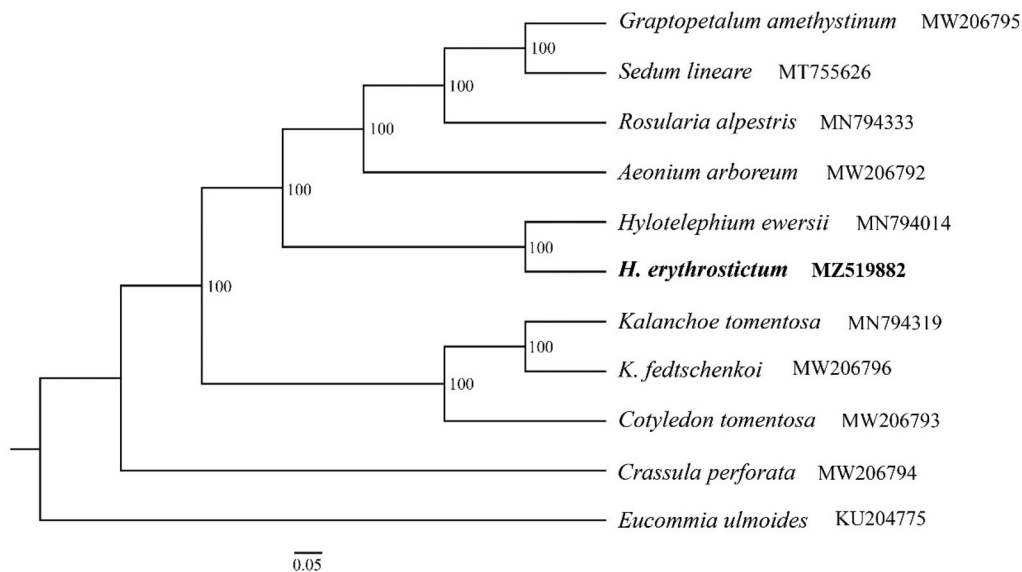


Figure 1. ML tree based on 10 complete chloroplast genomes of Crassulaceae and one outgroup species. Numbers at the nodes are bootstrap support values based on 1000 replicates. GenBank accession Numbers are listed beside the species. The species *H. erythrostictum* is highlighted in bold.

content of the cp genome is 37.7%. The whole cp genome of *H. erythrostictum* contains 134 genes with 80 PCGs, 37 tRNA genes, and 8 rRNA genes. Among these genes, 60 PCGs and 22 tRNA genes are located in the LSC region, while 12 PCGs and one tRNA gene occur in the SSC region. All these eight rRNA genes are duplicated in the IR regions. IR regions contain six PCGs and seven tRNA genes, if counting only once. Among the annotated genes, a total of 15 genes contain one intron, which are *trnK-UUU*, *trnL-UAA*, *rps16*, *trnV-UAC*, *rpl2*, *ndhB*, *trnG-GCC*, *trnI-GAU*, *trnA-UGC*, *atpF*, *ndhA*, *rpoC1*, *petD*, *petB*, and *rpl16*, while three genes including *clpP*, *ycf3*, and *rps12* possess two introns. The numbers and genes that contain one or two introns are exactly the same as *Sedum tricar-pum*, which is largely due to the low variation based on cpDNA genome and the relatively close phylogenetic relationship between *H. erythrostictum* and *Sedum tricar-pum* in the Crassulaceae family. A well supported phylogenetic tree is reconstructed, suggesting a monophyly formed by *H. erythrostictum* together with *H. ewersii* (Figure 1). The two species demonstrate a comparably closer phylogenetic relationship with the clade consisting of *Graptopetalum amethystinum*, *Sedum lineare*, *Rosularia alpestris*, and *Aeonium arboreum*.

Author contributions

Wei Gong and Jing Chen conceived and designed the project. Wei Gong and Hui Zhou analyzed and interpreted the data. Wei Gong and Jing Chen wrote the paper and revised it for intellectual content. All authors agree to be accountable for all aspects of the work.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the National Natural Science Foundation of China [82174272, 31470312], the Research Project of Zhejiang Chinese Medical University [2020ZG33].

Data availability statement

The complete chloroplast genome sequence data that support the findings of this study are openly available in GenBank of NCBI at [<https://www.ncbi.nlm.nih.gov>] (<https://www.ncbi.nlm.nih.gov/>) under the accession no. MZ519882. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA749471, SRR15239599, and SAMN20371389, respectively.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods.* 9(8):772.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19:11–15.
- 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.
- 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.
- Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32(1):11–16.
- Miller MA, Schwartz T, Pickett BE, He S, Klem EB, Scheuermann RH, Passarotti M, Kaufman S, O'Leary MA. 2015. A RESTful API for access to phylogenetic tools via the CIPRES science gateway. *Evol Bioinform Online.* 11:43–48.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9):1312–1313.