

MITOGENOME ANNOUNCEMENT



## Characterization of the complete chloroplast genome of the *Helleborus atrorubens* Waldst. & Kit. (Ranunculaceae)

Xiaohua Shi, Lihui Mao , Liang Jin and Guangying Ma

Zhejiang Academy of Agricultural Sciences, Zhejiang Institute of Landscape Plants and Flowers, Hangzhou, China

### ABSTRACT

*Helleborus atrorubens* is an economically important perennial garden plant with medicinal value. Here, we sequenced the complete chloroplast genome of *H. atrorubens*. The results revealed the chloroplast genome to be 166,695 bp in length. It possesses a typical quadripartite structure containing one large single copy (LSC) region (84,994 bp), one small single copy (SSC) region (17,825 bp), and a pair of inverted repeat (IR) regions (31,938 bp). This chloroplast genome encoded 130 genes, out of which 85 code for proteins, 37 for transfer RNAs, and 8 for ribosomal RNAs. Simple sequence repeat (SSR) markers and the top variable coding regions were identified. Our study lays a foundation for further research, such as species differentiation and phylogenetic reconstruction of the *Helleborus* genus.

### ARTICLE HISTORY

Received 13 March 2022  
Accepted 25 August 2022

### KEYWORDS

*Helleborus*; plastome;  
genome structure;  
phylogenomics



*Helleborus atrorubens* is a perennial species native to the Balkan Peninsula with ornamental and medicinal value. While the diversity of phytochemicals produced by *H. atrorubens* has been well investigated, there is a dearth of genomic resources for this species (Male and Medic-Saric 2001). The genetic relationships in *Helleborus* are complex although there are only approximately 20 species in this genus. The English botanist Brian Mathew classified the genus into six groups, species within same group can hybridize with each other while species from different groups are hard to hybridize (Rice and Strangman 1993). It is necessary to find SSR and potential polymorphic gene fragments from the chloroplast genome to better analyze the genetic relationships among the species and varieties in *Helleborus*. *Helleborus atrorubens* were purchased from Het Wilgenbroeks farm (Oostkamp, Belgium, 51.17° N, 3.23° E). The plants were cultivated in Wangcun Village, Linpu Town, Xiaoshan District, Hangzhou, China (30.069° N, 120.232° E). A voucher specimen was deposited in the Herbarium of Northwest A&F University (NWFC) under voucher number Shi001 (contact person: Linjuan Du, [dulingjuan@nwafu.edu.cn](mailto:dulingjuan@nwafu.edu.cn)). The genomic DNA was extracted with a plant genomic DNA kit (Tiangen Biotech, Beijing, China) and sequenced using the Illumina NovaSeq platform according to the manufacturer's recommendations.

Pair-end Illumina raw reads were cleaned from adaptors and barcodes and then quality filtered using Trimmomatic (Bolger et al. 2014). After the quality filter, reads were mapped to the plastome of *H. thibetanus*, the closest species with an available plastome, using Bowtie2 v.2.2.6 (Langmead

et al. 2009) to exclude reads of nuclear and mitochondrial origins. Next, all putative plastid reads mapped to the reference sequence above were used for *de novo* assembly to reconstruct the plastome using GetOrganelle 1.7.5 (Jin et al. 2020). The clean reads were then mapped to the complete plastome again to examine and correct mis-assemblies. Automatic annotation of the plastome was generated by CpGAVAS2 (Shi et al. 2019) and then manually corrected according to previously published plastomes using Geneious V8.1 (Kearse et al. 2012). A circular representation of both sequences was plotted using the online tool OGDRAW (Greiner et al. 2019).

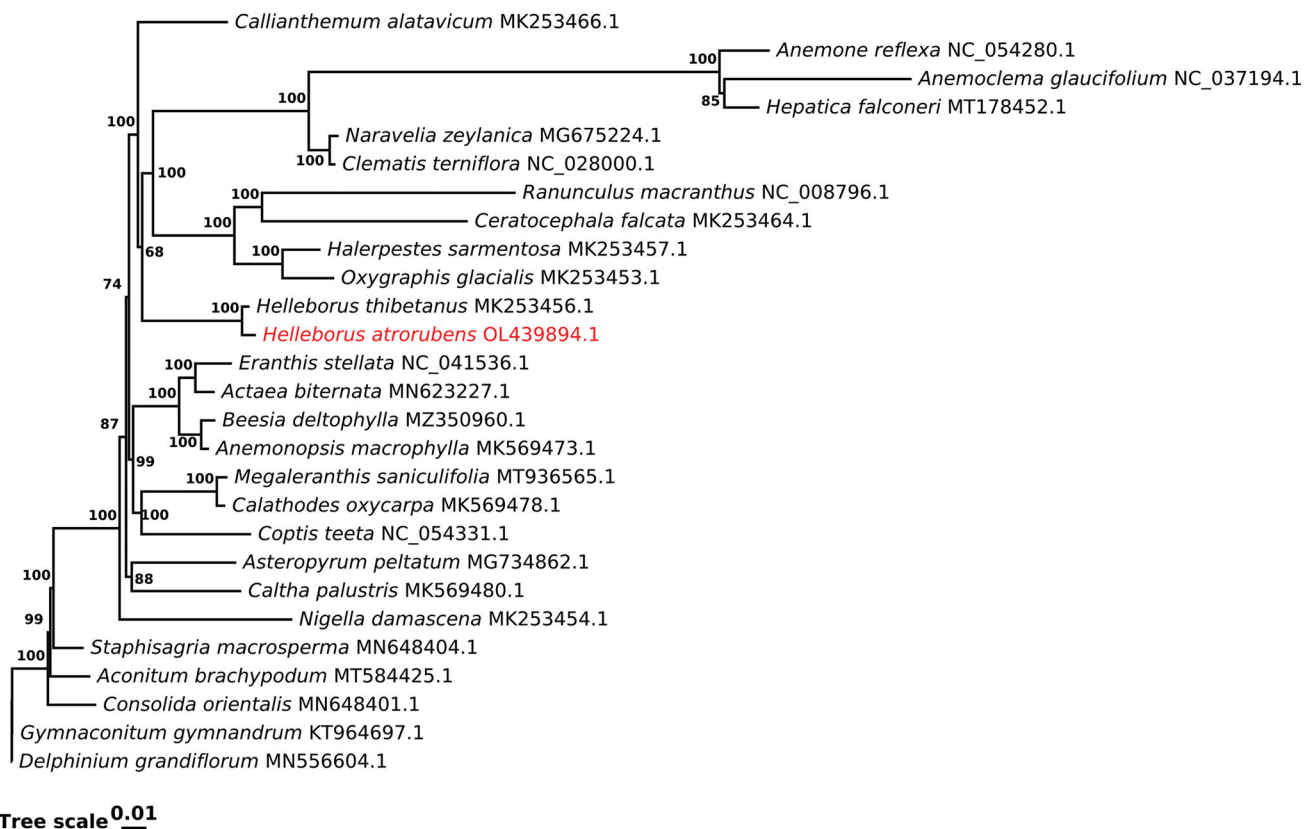
The complete plastome spans 166,695 bp and contains 130 genes, including 85 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. A maximum-likelihood analysis was performed by RAxML v8.2.10 (Stamatakis 2014) using the GTR model and 1000 bootstrap replicates to confirm the phylogenetic position of *H. atrorubens*. Twenty-seven species in Ranunculaceae family were used in the phylogenetic analysis, including *Anemone reflexa*, *Anemone glaucifolia*, *Clematis terniflora*, *Oxygraphis glacialis*, *Ranunculus macranthus*, and another 21 species. The analysis revealed *H. atrorubens* to have the closest relationship with *H. thibetanus* (Figure 1). The result indicates that the phylogenetic analysis of Ranunculaceae based on the chloroplast genome is credible.

In this study, 106 chloroplast SSRs (cpSSRs) were identified for *H. atrorubens*, ranging in length from 10 to 18 bp. These SSR loci could be used to investigate the genetic diversity and genetic structure of natural populations and cultivars of this species. The nucleotide diversity ( $\pi$ ) values were

**CONTACT** Lihui Mao  [mlh@mail.zaas.ac.cn](mailto:mlh@mail.zaas.ac.cn); 316194304@qq.com  No. 508, Wangcun Village, Linpu Town, Xiaoshan District, Hangzhou City, Zhejiang Province 311251, 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.** Molecular phylogenetic reconstruction of *Helleborus atrorubens* and 26 other Ranunculaceae species using maximum likelihood. Bootstrap support values ( $N = 1000$ ) are indicated at each node.

calculated using each gene to determine the hotspots of divergence. A total of 25 genes including *psbA*, *psbD*, *psbC* and others with a high Pi value ( $>0.01$ ) and suitable length ( $>500$  and  $<1500$  bp) for PCR amplification can be used as candidate molecular markers to study interspecific relationships (Table S1).

In this study, we elucidated the complete plastome of *H. atrorubens* for the first time. The gene order and cp genome arrangement of *H. atrorubens* were similar to those of related species. The current study provides a valuable contribution that could help species identification, crossbreeding and genetic diversity research.

### Ethical approval

Our plant collection and experiment passed the ethical review of our Institution.

### Author contributions

Xiaohua Shi and Guangying Ma were involved in data collection and drafting of the manuscript, Liang Jin provided data analysis, data interpretation, and article revising, Xiaohua Shi and Lihui Mao contributed to the data interpretation and writing the article. All authors final approved the final version of the study to be published and agreed to be accountable for all aspects of the work.

### Disclosure statement

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

### Funding

This work was supported by the 2020 Major Scientific and Technological Projects for Social Development of Xiaoshan District [2020222]. This work was also supported by Key Research and Development Plan of Zhejiang Province [2019c02025].

### ORCID

Lihui Mao  <http://orcid.org/0000-0002-9556-6642>

### Data availability statement

The data that support the findings of this study are openly available in GenBank of NCBI at <https://www.ncbi.nlm.nih.gov>, reference number OL439894. The associated \*\*BioProject\*\*, \*\*SRA\*\*, and \*\*Bio-Sample\*\* numbers are SAMN26535971, SRS12222224, and *Helleborus atrorubens* respectively.

### References

- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina Sequence Data. *Bioinformatics*. 30(15):2114–2120.
- Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res*. 47(W1):W59–W64.
- 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.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 10(3):R25.
- Male E, Medic-Saric M. 2001. Optimization of TLC analysis of flavonoids and phenolic acids of *Helleborus atrorubens* Waldst. et Kit. *J Pharm Biomed Anal*. 24(3):353–359.
- Rice G, Strangman E. 1993. *Gardener's guide to growing Hellebores*. Portland: Timber Press.
- Shi LC, Chen HM, Jiang M, Wang LQ, Wu X, Huang LF, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res*. 47(W1):W65–W73.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9): 1312–1313.