

The complete chloroplast genome sequence of *Turpinia montana* (Staphyleaceae)

Wei-Hong Sun^{a,b,c}, Yi-Fan Wang^{a,b}, De-Qiang Chen^{a,b,c} and Shuang-Quan Zou^{a,b}

^aCollege of Forestry, Fujian Agriculture and Forestry University, Fuzhou, P. R. China; ^bFujian Colleges and Universities Engineering Research Institute of Conservation and Utilization of Natural Bioresources, College of Forestry, Fujian Agriculture and Forestry University, Fuzhou, P. R. China; ^cKey Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization at College of Forestry, Fujian Agriculture and Forestry University, Fuzhou, P. R. China

ABSTRACT

The Rosids are characterized by remarkable morphological and ecological diversity. Here, we provide the completed plasmid genome of *Turpinia montana*. The complete chloroplast size of *T. montana* is 160,111 bp, including a large single-copy (LSC) region of 89,631 bp, a small single-copy (SSC) region of 18,247 bp, a pair of invert repeats (IR) regions of 26,120 bp. Plastid genome contains 131 genes, 86 protein-coding genes, 37 *tRNA* genes, and eight *rRNA* genes. Phylogenetic analysis based on 23 plastid genomes indicates that *T. montana* is clustered with the plants of the *Euscaphis japonica* and *Staphylea bumalda*.

ARTICLE HISTORY

Received 2 August 2020
Accepted 4 September 2020

KEYWORDS

Rosids; *Turpinia montana*; plastid genome; phylogeny

The Rosids are characterized by remarkable morphological and ecological diversity, with a total of 17 orders, containing 140 families and about 70,000 species (APG II 2003; APG III 2009; APG IV 2016). Their life forms include herbs, trees, aquatic plants, and succulents. Some of them are important economic crops (e.g. Rosaceae, Fabaceae, and Brassicaceae), and others are important trees (e.g. Betulaceae, Sapindaceae, and Fagaceae). The Rosids mainly divided into Fabids and Malvids, among which Fabids are divided into COM clade (Celastrales, Oxalidales, and Malpighiales) and the nitrogen-fixing clade (Cucurbitales, Fagales, Fabales, and Rosales; Qiu et al. 2010; APG IV 2016). Although many species of Rosids have completed chloroplast sequencing, the phylogenetic positions of the COM branch, Nitrogen-fixing branch, and the Malvids branch are still controversial (Zhao et al. 2016). The basal group is an ideal group for study the origin and evolution of species taxa. Here, we report the complete plastid genome of *Turpinia montana*, a species of the base taxa of the Malvids. *Turpinia montana* mainly distributed in South of China, Java and Sumatra in Indonesia (Li et al. 2008). The leaves and branches of *T. montana* are rich in flavonoids, triterpenes, and other secondary metabolites, which are widely used in medicinal materials (Lei et al. 2019; Xiao et al. 2019). Thus, the complete plastid genome of *T. montana* will be helpful to study the origin, evolution, and diversification within the Rosids and provide theoretical basis for effective conservation strategies for this important plant.

Since the leaves of *T. montana* are rich in secondary metabolism, we collected freshly germinated leaves in

Qingyun Mountain, Yongtai County, Fuzhou City, Fujian province, China (118°57'37"E, 25°46'47"N) for total genomic DNA extraction. The voucher specimens of branches, leaves, flowers, and fruits are kept in the Herbarium of the College of Forestry, Fujian Agriculture, and Forestry University (specimen code FAFU08031). A modified acetyl trimethyl ammonium bromide method was used to extract genomic DNA (Porebski et al. 1997) from *T. montana* leaves. The concentration and quality of the recovered DNA were measured by NanoDrop™ spectrophotometry (Thermo) at 260 and 280 nm, and the purity of the DNA was verified by 0.8% (w/v) agarose gel. The high-quality DNA was randomly interrupted 400 bp using the Covaris ultrasonic breaker to construct a library. Repair the end of the small fragment, and then add an A to the end of the small fragment 3' to connect to the adapter because there is a T at the adapter 3'/end. PCR enriches the target fragments, and finally detects the quality of the library. The library was sequenced on Illumina HiSeq Xten platform for PE150, generating about approximately 6.89 GB of raw data. The raw data was filtered by script in the cluster with default parameters of $-l5$, $-p0.5$, $-n 0.1$. With *Staphylea trifolia* complete plastid genome (GeneBank accession: MK488092) as reference, we assembled the plastid genome of *T. montana* by using GetOrganelle pipe-line (<https://github.com/Kinggerm/GetOrganelle>; Jin et al. 2019). The Bandage (Wick et al. 2015) was used to view and edit assembled sequences, and the Geneious version 11.1.5 (America) was used to annotation the plastid genome of *T. montana* (Kearse et al. 2012). Finally, the annotation result

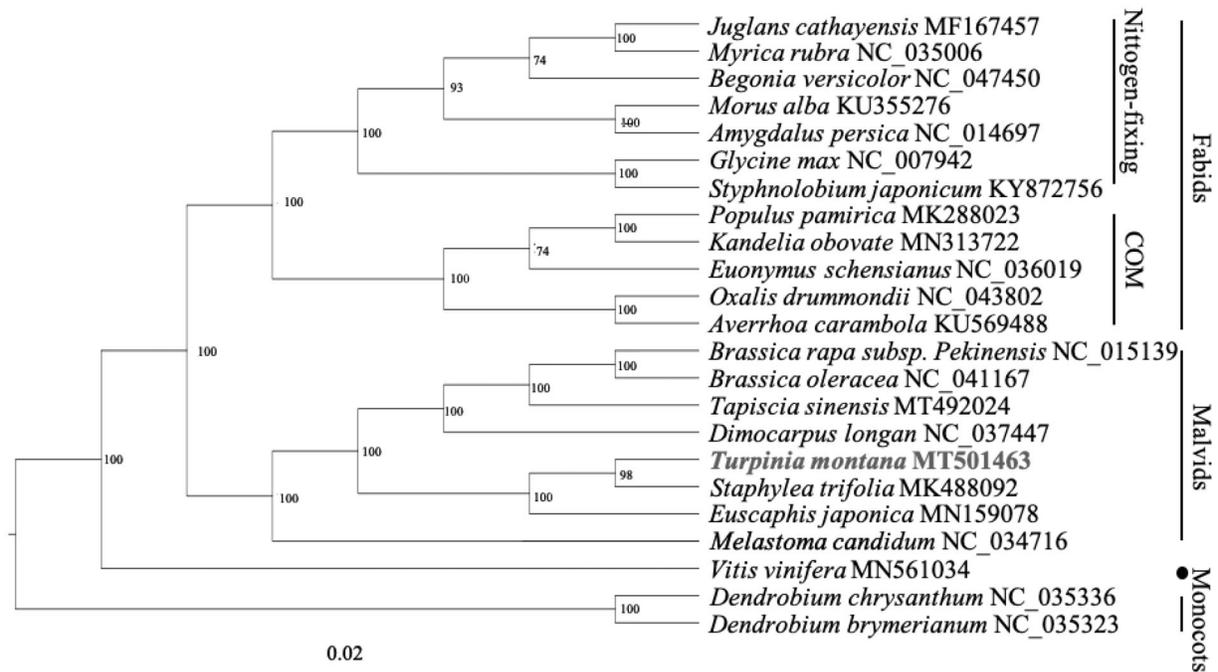


Figure 1. The phylogenetic tree of *Turpinia montana* with other species.

was visualized by using the online tool OGDRAW (<http://ogdraw.mpimp-golm.mpg.de/>; Lohse et al. 2013). The complete plastid genome sequence of *T. montana* has been submitted to GenBank with the accession number MT501463.

The complete plastid genome sequence length of *T. montana* was 160,111 bp, in which a pair of inverted repeats (IR) regions were 26,120 bp, each, a large single-copy (LSC) region was 89,631 bp, and small single-copy (SSC) region was 18,247 bp. The complete genome GC content was 37.4%, and the complete plastid genome encodes 131 genes, containing 86 protein-coding genes, 37 transfer RNA (tRNA) genes, and eight ribosomal (*rRNA*) genes. To reveal the phylogenetic position within the Rosids, a phylogenetic tree was conducted based on complete plastid genomes of 23 species, with two monocotyledons, *Dendrobium chrysanthum* and *Dendrobium brymerianum*, as the outer group. These genome data were downloaded from NCBI GenBank (<https://www.ncbi.nlm.nih.gov/search/all/?term=GenBank>). The chloroplast genome of *T. montanas* was aligned with other 22 chloroplast genome sequence, using RAxML to construct the maximum likelihood tree (Stamatakis 2014). The phylogenetic tree showed that the Staphyleaceae species *T. montanas*, *Euscaphis japonica*, and *Staphylea bumalda* cluster together (Xiang et al. 2019; Jing-Yao et al. 2020), but they are not the base group of the Malvids. This result is inconsistent with the results of previous studies that *T. montana* is located at the base of Malvids (APG IV 2016). The completed plastid genome sequence of *T. montana* can help reveal its phylogenetic position, and provide data for future conservation efforts and biological research (Figure 1).

Disclosure statement

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

Funding

This work was supported by the Special Funds for Leading Scientific and Technological Innovation Talents of Fujian Province [118/KRC16006A] and the Fujian Agriculture and Forestry University Excellent Postgraduate Thesis Funding Fund [324-1122yb062].

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 MT501463.

References

- APG II. 2003. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. *Bot J Linn Soc.* 141:399–436.
- APG III. 2009. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. *Narnia.* 161:105–121.
- APG IV. 2016. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc.* 181:1–20.
- Jin JJ, Yu WB, Yang JB, Song YW, de Pamphilis C, Yi TS, Li DZ. 2019. Get organelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Gen Biol.* 21:241.
- Jing-Yao Z, De-Qiang C, Shuang X, Xi W, Yi-Fan W, Yi Xun L, Wei-Hong S, Shuang-Quan Z, Yun-Wei Z, Xiao-Xing Z. 2020. The complete chloroplast genome sequence of *Tapiscia sinensis* (Staphyleaceae). *Mitochondrial DNA Part B.* 5(3):2658–2660.
- 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.
- Lei K, Xiao CR, Tu LF, Wu YM, Zhang RZ, Liu DP, Luo YM. 2019. Chemical constituents from the leaves of *Turpinia arguta*. *J Chin Med Mater.* 42(11): 2570–2573.

- Li DZ, Cai J, Wen J. 2008. Staphyleaceae. In: Wu ZY, Raven PH, Hong DY, editors. Flora of China. Vol. 11. Beijing (China): Science Press; St. Louis (MO): Missouri Botanical Garden Press; p. 498–504.
- Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 41:W575–W581.
- Porebski S, Bailey LG, Baum BR. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Rep.* 15(1):8–15.
- Qiu YL, Li LB, Wang B, Xue JY, Hendry TA, Li RQ, Brown JW, Liu Y, Hudson GT, Chen ZD. 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *J Syst E.* 48:391–425.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9):1312–1313.
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics.* 31(20):3350–3352.
- Xiang S, Liu XD, Sun WY, Lan SR. 2019. The complete chloroplast genome sequence of *Euscaphis japonica* (Staphyleaceae). *Mitochondrial DNA Part B.* 4(2):3484–3485.
- Xiao CR, Tu LF, Zhang RZ, Liu DP, Luo YM. 2019. Study on chemical constituents of flavonoids in *Turpinia arguta*. *Acta Sin Sin.* 54(09):1620–1626.
- Zhao L, Li X, Zhang N, Zhang SD, Yi TS, Ma H, Zhen H. 2016. Phylogenomic analyses of large-scale nuclear genes provide new insights into the evolutionary relationships within the rosids. *Mol Phylogenet Evol.* 105:166–176.