The complete chloroplast genome of *Leptodermis scabrida* (Rubiaceae): an endemic shrub in Himalaya-Hengduan Mountains

Ying Zhang\textsuperscript{a,b}, Sheng Chen\textsuperscript{c} and Ruijiang Wang\textsuperscript{a}

\textsuperscript{a}Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China; \textsuperscript{b}College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China; \textsuperscript{c}College of Life Sciences, South China Agricultural University, Guangzhou, China

**ABSTRACT**

*Leptodermis scabrida* is an important endemic species in Himalaya-Hengduan Mountains. Here, we report the complete chloroplast genome of *L. scabrida*. The cp genome was determined to be 154783 bp in length and the GC content was 37.5%. The sequence included a large single copy (LSC) region of 84190 bp, a small single copy (SSC) region of 17183 bp and two separated inverted regions of 26705 bp, respectively. It contained 132 genes, including 84 protein-coding genes, 38 tRNA genes, and 8 rRNA genes. This complete genome of *L. scabrida* will provide valuable information to resolve the complex phylogeny relationship and to elucidate the mechanism of speciation of *Leptodermis*, as well as for the phylogenetic studies of Rubiaceae.

*Leptodermis scabrida* Hook. is a deciduous shrub with a height of up to 1 meter, which belongs to *Leptodermis* Wall., Rubiaceae (Luo 1999b). *Leptodermis scabrida* is a kind of wild native plant with strong resistance that is mainly distributed in the Himalaya-Hengduan Mountains (Li 1985), which is the biodiversity hotspot. Until now, no genome studies of the genus *Leptodermis* have been reported. Therefore, additional more genomic data are urgently needed to resolve the complex phylogeny relationship and to elucidate the mechanism of speciation of *Leptodermis*, which may also provide new insight into other research in Himalaya-Mountains. Here, we report and characterize the complete chloroplast genome (cp genome) sequence of *L. scabrida* (GenBank accession number: MN686284) for the first time.

The fresh, young and healthy leaves of *L. scabrida* were collected from Bomi County, Linzhi City, Tibet, China (29°53'54.7360''N, 95°39'41.1760''E, 2680 m). The voucher specimen (voucher No.: 888) was deposited in the South China Botanical Garden Herbarium (IBSC), Guangzhou, China. The total genomic DNA was extracted from the silica-gel dried leaves following the modified CTAB method (Doyle and Doyle 1987). Then the genomic library (paired-end, PE = 150 bp) was sequenced on an Illumina Hiseq X Ten platform at Beijing Genomics Institute (Shenzhen, China). Totally 2 Gb sequence reads were obtained and used to assemble the cp genome after filtering and trimming the low-quality reads and adaptor sequences. The complete cp genome assembly was executed on NOVOPlasty 2.6.3 (Dierckxsens et al. 2017) with the default k-mer of 39–59, while Geneious version 11.0.3 (Kearse et al. 2012) was used to annotate the genome.

*Galium mollugo* (GenBank accession number: NC_036970) was used as the reference plastid genome for assembling and annotation. The tRNA genes were annotated on ARAGORN (Laslett and Canback 2004). The complete cp genome of *L. scabrida* was 154783 bp with the typical quadripartite structure of angiosperms, including a small single-copy region (SSC) of 17183 bp, a large single-copy region (LSC) of 84190 bp and a pair of inverted repeats (IRs) of 26705 bp. The genome harbored 132 genes, including 38 tRNA genes, 8 rRNA genes, and 84 protein-coding genes. The overall GC content in the cp genome of *L. scabrida* was 37.50%, which the corresponding value of the SSC, LSC, and IR region were 31.40, 35.40, and 42.80%, respectively.

To identify the phylogenetic position of *L. scabrida* in the Rubiaceae, a maximum likelihood (ML) tree was reconstructed by using RAxML (Stamatakis 2006) with 1000 bootstrap replicates under the GTRGAMMA substitution model. The matrix, used for phylogenetic analysis, included 14 whole cp genome sequences of Rubiaceae species (13 of them were downloaded from GenBank) and was aligned in MAFFT (Katoh and Standley 2013), using *Buddleja colvillei* Hook. f. & Thoms (Loganiaceae) (GenBank accession number: NC_042766) as an outgroup (Figure 1). The phylogenetic analysis was divided into three parts, which was consistent with the subfamily classified (Rubioidae, Cinchonoidae, and Ixoroideae). Most nodes in the chloroplast ML tree were strongly supported. *Leptodermis scabrida* was closely related to *Galium mollugo* L. formed a monophyletic clade with strong bootstrap support, belonging to the subfamily Rubioidae.
Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The study was financially supported by the National Natural Science Foundation of China [Grant No. 30770156], and the “Strategic Priority Research Program” of the Chinese Academy of Sciences [Grant No. XDA13020602].

References

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bullet. 19:11–15.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

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.

Li H. 1985. Flora of Xizang. Vol 4. Beijing: Science Press; p. 456.

Luo XR. 1999b. Leptodermis. In: Chen WQ, editor. Flora of China. Vol. 19. Beijing: Science Press; p. 120–153.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22(21):2688–2690.