

## Characterization of the complete chloroplast genome of the Taiwan alder *Alnus formosana* (Betulaceae) based on next-generation sequencing technology

Xiaoning Zhang<sup>a</sup> , Mimi Li<sup>b,c</sup> , Qiulan Wei<sup>a</sup>, Yufei Xiao<sup>a</sup>, Yufeng Qin<sup>a</sup>, Lianxiang Zhong<sup>a</sup> and Zihai Qin<sup>a</sup>

<sup>a</sup>Guangxi Forestry Research Institute, Nanning, China; <sup>b</sup>Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China; <sup>c</sup>The Jiangsu Provincial Platform for Conservation and Utilization of Agricultural Germplasm, Nanjing, China

### ABSTRACT

*Alnus formosana* (Betulaceae) is an important ecological and economic deciduous tree species widely distributed throughout subtropical regions of Taiwan province, China. At the present study, the complete chloroplast genome of *A. formosana* was assembled using next-generation sequencing technology. The complete chloroplast sequence is 161,029 bp in length, which consisted of a large single copy (LSC, 89,720 bp) and a small single copy (SSC; 19,205 bp) separated a pair of inverted repeats (IRs; 26,052 bp). The overall guanine-cytosine (GC) content was 36.4%. A total of 131 genes were annotated, including 85 protein-coding genes, 37 tRNAs, eight rRNAs and one pseudogene (*ψycf1*). The phylogenetic analysis fully resolved *A. formosana* in a clade with *A. japonica*. The plastome of *A. formosana* will provide informative genomic resources for further phylogenetic application and genetic improvement.

### ARTICLE HISTORY

Received 8 April 2021  
Accepted 12 August 2021

### KEYWORDS

*Alnus formosana*;  
Betulaceae; chloroplast  
genome; phylogeny



*Alnus formosana* (Burkill) Makino 1912, classified in the birch family (Betulaceae), is an important ecological and economic deciduous tree species (Li and Skvortsov 1999). It has a wide natural distribution throughout the subtropical regions in the Taiwan province, China (Liao and Weng 2002). It is characterized by fast-growing, nitrogen-fixing and wide ranging adaptability to adverse environmental conditions, and is, therefore, widely used as pioneer tree of afforestation (Pan et al. 2008). This species is also a source of excellent quality wood for furniture and industrial production. However, the phylogenetic position of *A. formosana* within *Alnus* has not yet been sufficiently resolved (Ren et al. 2010). Here, we report the complete chloroplast (cp) genome sequence of *A. formosana* using next-generation sequencing, which will lay the foundation for further phylogenetic application and genetic resolution.

Total genomic DNA was extracted from the fresh leaves of *A. formosana* cultivated in Guangxi Forestry Research Institute, Guangxi Province, China (22°55'30"N, 108°21'E) by modified hexadecyl trimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987). The specimen was deposited at Guangxi Forestry Research Institute (contact: Zihai Qin, [75455621@qq.com](mailto:75455621@qq.com)) under voucher number: 20210303002. The isolated genomic DNA was subsequently sequenced on an Illumina HiSeq X-ten platform (San Diego, USA) at Novogene (Beijing, China). The raw Paired-end (PE) reads were *de novo* assembled into the chloroplast genome using the perl script NOVOPlasty 4.3.1 (Dierckxsens et al. 2017) with default settings. The assembled genome was

annotated via GeSeq (Tillich et al. 2017) and adjusted manually in Geneious 11.1.5 (Kearse et al. 2012).

The chloroplast genome of *A. formosana* has a circular structure with a length of 161,029 bp, comprising four parts: a LSC (89,720 bp), a SSC (19,205 bp) and a pair of IRs (26,052 bp). The overall guanine-cytosine (GC) content was 36.4%, with the GC contents of LSC, SSC and IRs at 34.1%, 30.1%, and 42.6%, respectively. The complete sequence data was submitted to National Center for Biotechnology Information (NCBI) (accession number MW865380). A total of 131 genes were annotated, including 85 protein-coding, 37 transfer RNA (tRNA), eight ribosomal RNA (rRNA) and one pseudogene (*ψycf1*). Among them, four rRNA (4.5S, 5S, 16S and 23S rRNA), seven tRNA (*trnI-GAU*, *trnA-UGC*, *trnL-CAA*, *trnI-CAT*, *trnR-ACG*, *trnV-GAC*, *trnN-GTT*), and seven protein-coding genes (*rpl2*, *rpl23*, *ycf2*, *ndhB*, *rps7*, *rps12*, and *ycf15*) were duplicated in the IR regions. The plastome of *A. formosana* had similar in the terms of structure, gene content and order, and GC content compared to those of other published *Alnus*.

To reveal the phylogenetic relationship of *A. formosana* within *Alnus*, we downloaded additional plastomes from NCBI including fifteen *Alnus* and one *Betula* species. The whole plastome sequences were aligned by MAFFT 7.409 (Katoh and Standley 2013) using default settings. The phylogenetic analysis was generated by the GTR+GAMMA model in RAxML with 1,000 bootstrap replicates (Stamatakis 2014). The maximum likelihood (ML) phylogenetic tree indicated that *Alnus* is a monophyletic group which is consistent with the previous analysis by Chen and Li (2004) but with higher bootstrap support (100%) (Figure 1). Within genus *Alnus*,

**CONTACT** Zihai Qin  [75455621@qq.com](mailto:75455621@qq.com)  Guangxi Forestry Research Institute, Nanning, China

© 2021 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.

