

The complete chloroplast genome of *Betula costata* (Betulaceae)

Min Woo Lee, Sang-Chul Kim, Jei-Wan Lee and Ji-Young Ahn

Division of Forest Bioinformation, National Institute of Forest Science, Suwon, Republic of Korea

ABSTRACT

In the present study, we analyzed the complete chloroplast genome sequence of *Betula costata* using the Ion Torrent platform. The chloroplast genome of *B. costata* was found to be 160,547 bp in length, with a large single-copy (LSC) region of 89,385 bp, a small single-copy (SSC) region of 19,038 bp, and a pair of inverted repeats (IRs) of 26,062 bp each. The overall GC content of the chloroplast genome was 36.1%. It contained 132 genes, including 87 protein-coding genes, 8 rRNA genes, and 37 tRNA genes. The phylogenetic analysis indicated that *B. costata* is closely related to *Betula nana* and *Betula pubescens*.

ARTICLE HISTORY

Received 31 December 2019
Accepted 19 January 2020

KEYWORDS

Betula costata; Betulaceae; complete chloroplast genome; phylogenetic tree

The genus *Betula* (Betulaceae) is mainly distributed in the Northern Hemisphere and consists of about 60 species (Chang 2000). Five species of *Betula* (*Betula chinensis* Maxim., *Betula costata* Trautv., *Betula davurica* Pall., *Betula ermanii* Cham., and *Betula schmidtii* Regel.) are distributed in Northeast Asia. Among these, *B. costata* is distributed in the subalpine regions of Korea and Northeast China (Chen 2002; Han et al. 2012). It is used as a wooden board and its sap is used for medical purposes, such as in constipation, gout, and neuralgia (Kim et al. 1991; Choi et al. 2006). *Betula costata* was reported as one of the species used for making printing woodblocks of the Tripitaka Koreana, which was registered as a Memory of the World by UNESCO (Cho 2011). Previous studies on this species included determination of the phylogenetic relationships based on AFLP markers (Schenk et al. 2008) and sequencing of the nuclear ribosomal DNA (Li et al. 2005). However, no studies have been reported on the complete chloroplast genome sequence information. In the present study, we characterized the chloroplast DNA information in *B. costata* and analyzed its phylogenetic relationships with eight other species of Betulaceae.

Plant materials were collected from Mt. Gyebangsan, South Korea (N37°43'9.3", E128°26'48.5"). Total DNA was extracted using a DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) and stored in a DNA bank (Division of Forest Bioinformation DNA bank, No. 0335133002). The whole-gen-

ome sequencing data were generated using the Ion Torrent platform (Life Technologies Corporation, Carlsbad, CA). The sequenced fragments were assembled using Geneious 10.2.3. The tRNAs were confirmed using the web-based tool, tRNAscan-SE (Lowe and Eddy 1997). A maximum-likelihood (ML) tree was constructed using the RAxML Blackbox web server, which used the rapid bootstrap analysis (Kozlov et al. 2019). The phylogenetic analysis was conducted using 100 bootstrap replicates. The complete chloroplast genome of *B. costata* (GenBank: MN830400) was found to have 160,547 bp, with a large single-copy (LSC) region of 89,385 bp, a small single-copy (SSC) region of 19,038 bp, and two inverted repeat regions (IRa and IRb) of 26,062 bp each. The overall GC content was 36.1% (LSC, 33.7%; SSC, 29.7%; IRs, 42.5%). The genome contained 132 genes, including 87 protein-coding genes, 8 rRNA genes, and 37 tRNA genes. Seven of the protein-coding genes, four rRNA genes, and seven tRNA genes, were duplicated in the IR regions. The other 15 protein-coding genes contained a single intron, and 2 genes (*ycf3* and *clpP*) had two introns. Phylogenetic analysis of *B. costata* with eight other species of *Betula*, and *Carpinus laxiflora* (using the outgroup). *Betula costata* was found to be closely related to *Betula nana* and *Betula pubescens* (Figure 1). This study provides basic information for phylogenetic studies on *B. costata* as well as on other *Betula* species.

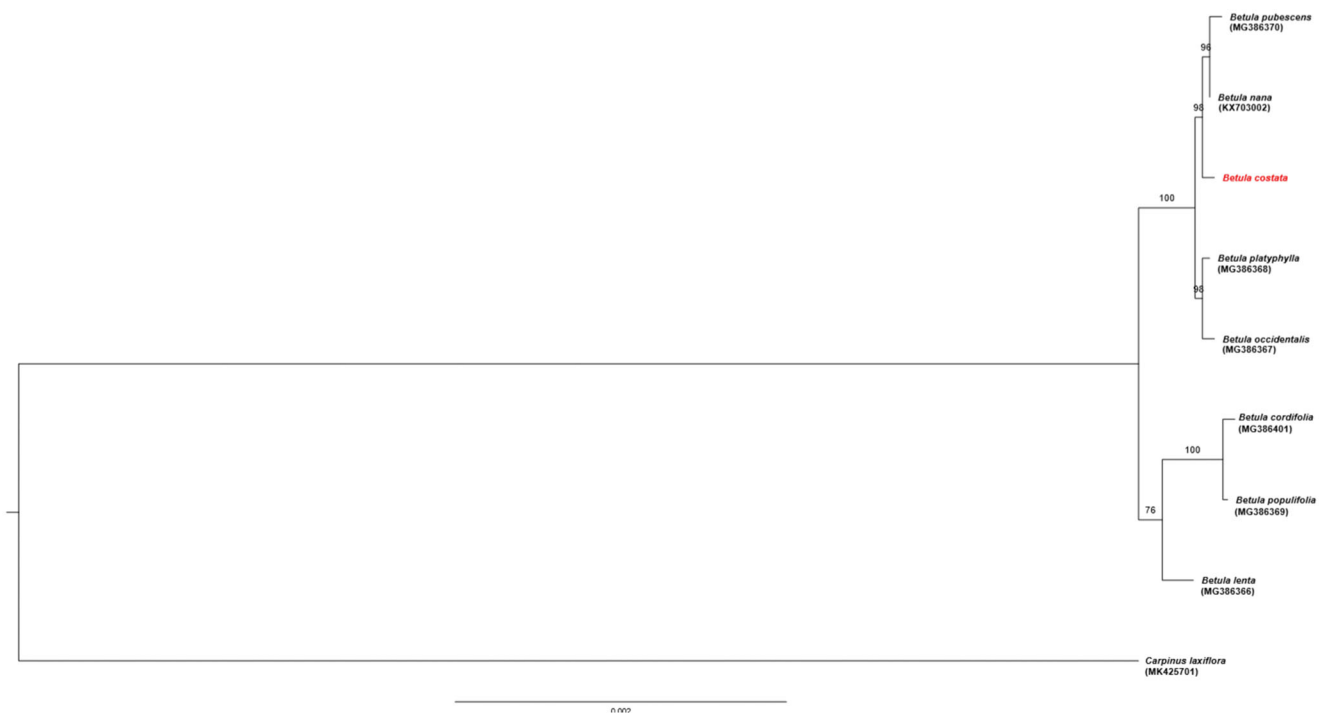


Figure 1. The phylogenetic tree of *B. costata* with eight species belonging to the Betulaceae based on chloroplast protein-coding sequences. Numbers in the nodes are the bootstrap values from 100 replicates.

Disclosure statement

The authors report no declaration of interest. The authors alone are responsible for the content and writing of this article.

References

- Chang CS. 2000. Foliar flavonoids of eastern Asian birch (*Betula*) - with respect to Korean plants. *Korean J PI Taxon.* 30(1):75–91.
- Chen X. 2002. Simulation of shift of range for *Betula costata* Trautv. and *Juglans mandshurica* Maxim. along Northeast China Transect (NECT). *Polish J Ecol.* 50(3):397–402.
- Cho HH. 2011. Awareness on the Goryeo Daejanggyeong and its Wood Plates [Master's thesis]. Seoul (South Korea): University of Seoul; p. 44.
- Choi SY, Sung NJ, Kim HJ. 2006. Physicochemical analysis and sensory evaluation of fermented soy sauce from Gorosoe (*Acer mono* Max.) and Kojesu (*Betula costata* T.) saps. *The Kor J Food Nutr.* 19(3):318–326.
- Han SS, Kim DH, Kim GN, Lee JC, Yun CW. 2012. Changes on Initial Growth and Physiological Characteristics of *Larix kaempferi* and *Betula costata* Seedlings under Elevated Temperature. *Kor J Agricult Forest Meteorol.* 14(2):63–70.
- Kim CM, Jung DL, Sheo HJ. 1991. A Study on the Ingredients in the Sap of *Acer mono* MAX. and *Betula costata* T. in Mt. Jiri Area - On the components of mineral and sugar. *J Kor Soc Food Sci Nutr.* 20(5):479–482.
- Kozlov A, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable, and user-friendly tool for maximum-likelihood phylogenetic inference. *Bioinformatics.* 35(21):4453–4455.
- Li J, Shoup S, Chen Z. 2005. Phylogenetics of *Betula* (Betulaceae) inferred from sequences of nuclear ribosomal DNA. *Rhodora.* 107(929):69–87.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25(5): 955–964.
- Schenk MF, Thienpont CN, Koopman WJM, Gilissen L, Smulders M. 2008. Phylogenetic relationships in *Betula* (Betulaceae) based on AFLP markers. *Tree Genet Genomes.* 4:911–924.