Korean Journal of Plant Taxonomy
52(2): 114–117, 2022

SHORT COMMUNICATION
https://doi.org/10.11110/kjpt.2022.52.2.114

Characteristics of the complete plastid genome sequence of *Lindera angustifolia* (Lauraceae) in the geographically separated northern edge

Amarsanaa GANTSETSEG, Jung-Hyun KIM[1], Chang Woo HYUN[2], Eun-Kyeong HAN and Jung-Hyun LEE*

Department of Biology Education, Chonnam National University, Gwangju 61186, Korea
[1]Korean Plant Diversity Institute, Yecheon 36849, Korea
[2]Plant Resources Division, National Institute of Biological Resources, Incheon 22689, Korea

Corresponding author
Jung-Hyun LEE
E-mail: quercus@jnu.ac.kr

ABSTRACT: *Lindera angustifolia* is mainly distributed in the temperate climate zone of China but shows an extraordinary distribution, disjunctively isolated on the western coastal islands of Korea. We therefore present the complete chloroplast genome of Korean *L. angustifolia*. The complete plastome was 152,836 bp in length, with an overall GC content of 39.2%. A large single copy (93,726 bp) and a small single copy (18,946 bp) of the genome were separated by a pair of inverted repeats (20,082 bp). The genome consists of 125 genes, including 81 protein-coding, eight ribosomal RNA, and 36 transfer RNA genes. While five RNA editing genes (*psb*L, *rpl*2, *ndh*B×2, and *ndh*D) were identified in *L. angustifolia* from China, the "*ndh*D" gene was not recognized as an RNA editing site in the corresponding Korean individual. A phylogenetic analysis revealed that Korean *L. angustifolia* is most closely related to the Chinese *L. angustifolia* with strong bootstrap support, forming a sister group of *L. glauca*.

Keywords: chloroplast, Lauraceae, *Lindera angustifolia*, medicinal plant, variations

RECEIVED 7 March 2022; **REVISED** 31 May 2022; **ACCEPTED** 16 June 2022

INTRODUCTION

Lindera angustifolia W. C. Cheng is a traditional medicinal plant and mainly distributed in the temperate zone of the China (Zhao et al., 2006). However, the species exhibits an extraordinary distribution, disjunctively isolated in western coastal islands of Korea, about 800 km northeast from China. In Korea, the plants are sometimes regarded as a variety, *L. angustifolia* var. *glabra* (Nakai) J. M. H. Shaw (Shaw, 2013). Therefore, the complete plastid genome sequence is an important basis for improving our understanding of the evolutionary process and solving taxonomic problems in inter/intra species within the genus *Lindera*.

MATERIALS AND METHODS

Plant material

A fresh leaf sample of *L. angustifolia* was collected from Daeyeonpyeongdo Island (37°39'44.0"N, 125°41'23.3"E), Yeonpyeong-ri, Ongjin-gun, Incheon. The voucher specimen of *L. angustifolia* was deposited at the Herbarium of the National Institute of Biological Resources in Korea (voucher number NIBRVP0000463020).

Chloroplast genome determination

The DNA library was constructed using Illumina Hi-seq platform (Macrogen, Seoul, Korea) and generated 56,426,712 raw reads (150 bp paired-end). The plastome was assembled using NOVOPlasty 4.3.1 (Dierckxsens et al., 2017), with the complete plastome of Chinese *L. angustifolia* (MG581438) (Jo et al., 2019) as the seed. The assembled plastome was checked using Geneious 11.0.9 (Kearse et al., 2012) by mapping 198,594 reads, resulting in a coverage of 150×. The annotation was conducted using Geneious 11.0.9 with the Chinese *L. angustifolia* (GenBank accession number MG581438) and *L. glauca* (GenBank accession number MG581443) (Jo

et al., 2019) chloroplast complete genome as a reference. We manually corrected the start and stop codons as well as the intron/exon boundaries.

Phylogenetic analysis

To reconstruct the phylogenetic tree, we downloaded 29 complete plastomes of Lauraceae (*Lindera* 18 taxa, *Litsea* 4 species, *Neolitsea* 3 species, and *Laurus*, *Actinodaphne* one species each, and two outgroups *Cinnamomum* and *Machilus* separately one species) from the NCBI database, including recently reported genome from Korea (Park and Cheon, 2021). The alignments were performed using MAFFT (Katoh and Toh, 2010). The maximum likelihood tree analysis was performed on a data set that includes 75 common genes extracted from 30 species with RAxML v.8.0 (Stamatakis, 2014) using default parameters and 1,000 bootstrap replicates. For the RAxML tree, the general time-reversible model of nucleotide substation was used with the Gamma model of rate heterogeneity.

RESULTS AND DISCUSSION

The novel annotated complete plastid genome of *L. angustifolia* is openly available in the NCBI GenBank at https://www.ncbi.nlm.nih.gov/ under the accession number OK328173. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA790663, SRR17287206, and SAMN24219956, respectively.

The complete chloroplast genome of *L. angustifolia* is 152,836 bp in length and has four subregions: 93,726 bp in

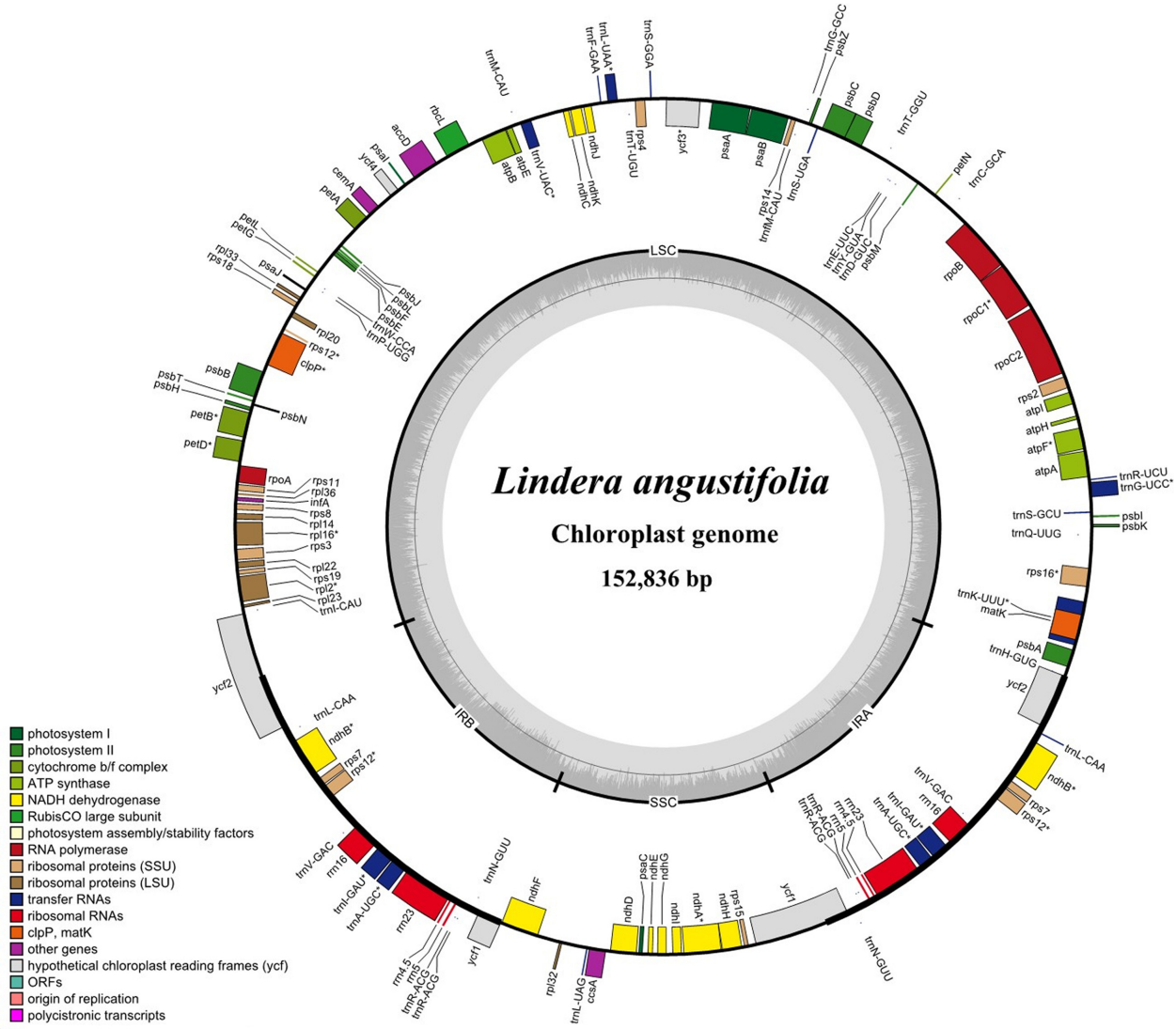

Fig. 1. Structural map of the chloroplast genome of *Lindera angustifolia*. Genes shown outside are transcribed clockwise, and inside the circle are transcribed counterclockwise. Genes are color-coded to distinguish different functional groups.

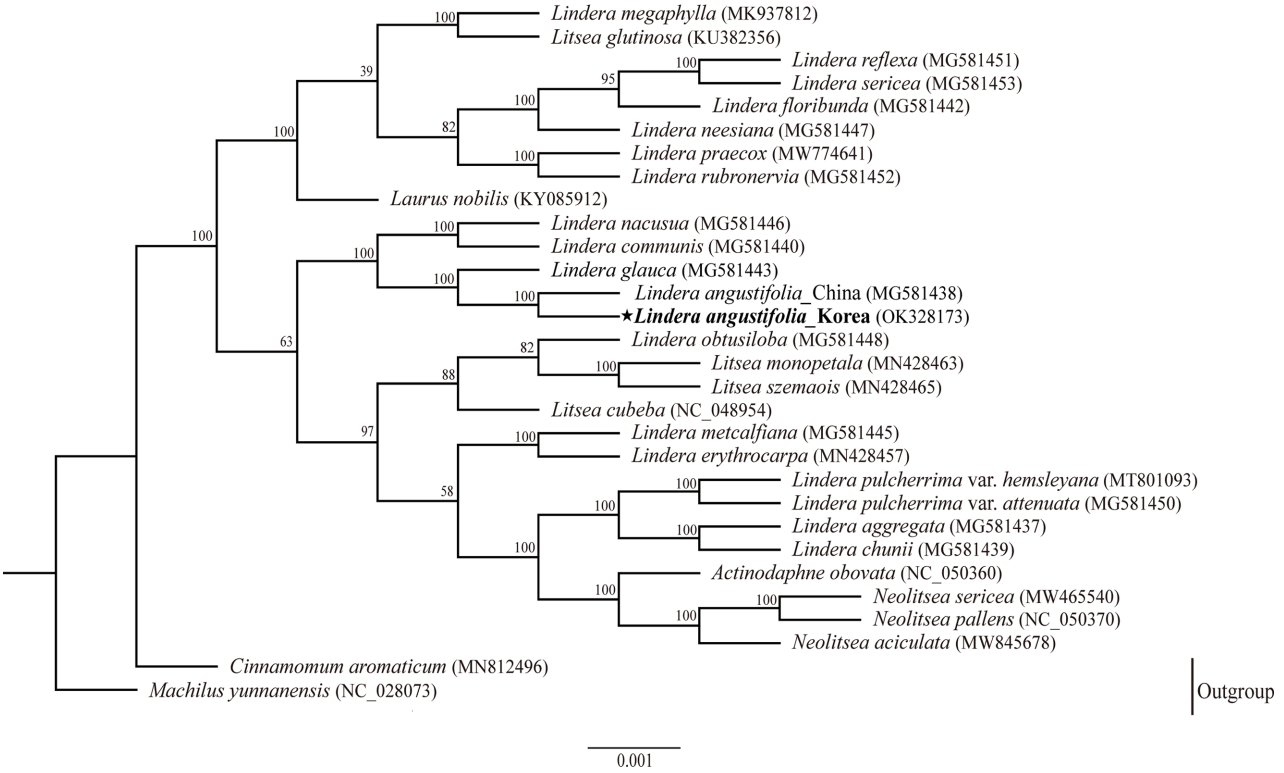

Fig. 2. A phylogenetic tree (RAxML) based on the plastome sequences of 30 taxa of Lauraceae and 75 common genes from 30 complete chloroplast genomes. The tree is rooted with the plastome sequence of *Machilus yunnanensis* and *Cinnamomum aromaticum*. The numbers above the nodes indicate bootstrap values with 1,000 replicates.

the large single-copy (LSC) region and 18,946 bp in small single-copy (SSC) region are separated by two inverted repeat (IR) regions of 20,082 bp (Fig. 1). The overall GC content of *L. angustifolia* is 39.2% and those in the LSC, SSC, and IR regions are 38.0%, 33.9%, and 44.4%, respectively. A total 125 genes were detected including 81 protein-coding genes, 36 tRNA genes, and eight rRNA genes.

Among the genes, 18 genes (*rps*16, *atp*F, *rpo*C1, *pet*B, *pet*D, *rpl*16, *rpl*2, *ndh*B×2, *ndh*A, *trn*K-UUU, *trn*G-UCC, *trn*L-UAA, *trn*V-UAC, *trn*I-GAU×2, and *trn*A-UGC×2) contain a single intron and three genes (*ycf*3, *clp*P, and *rps*12) have two introns.

As a result of comparing the complete chloroplast genomes of Chinese and Korean *L. angustifolia,* the total genome size is 152,836 bp and 152,832 bp long respectively, with a difference of only 4 bp. In addition, the two genes (*ycf*1 and *ycf*2) in the IR regions and one gene (*rpl*22) in the LSC region are pseudogenized in both Chinese and Korean samples. The nucleotide sequences of the 1,372 bp of 3′–*ycf*1, 3,162 bp of 5′–*ycf*2, and 387 bp of 5′–*rpl*22 were truncated at the LSC and IR boundaries separately. Furthermore, while five RNA editing genes (*psb*L, *rpl*2, *ndh*B×2, and *ndh*D) were identified in *L. angustifolia* from China, the "*ndh*D" gene was not recognized as an RNA editing site in that Korean individual.

The phylogenetic tree (RAxML) constructed for 30 species of the Lauraceae family shows that Korean *L. angustifolia* is most closely related to Chinese *L. angustifolia* with strong (100%) bootstrap support and forms a sister group of *L. glauca* (Fig. 2). The additional plastome sequence will be a valuable, fundamental tool for future studies of phylogenetic relationships among species of *Lindera*, and will provide a useful resource for researching geographically genetic variations.

ORCID: Amarsanaa GANTSETSEG https://orcid.org/0000-0002-9251-9509; Jung-Hyun KIM https://orcid.org/0000-0003-2699-5635; Chang Woo HYUN https://orcid.org/0000-0001-9677-5270; Eun-Kyeong HAN https://orcid.org/0000-0002-8530-9925; Jung-Hyun LEE https://orcid.org/0000-0003-0066-6032

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institute of Biological Resources (NIBR), funded by the

Ministry of Environment (MOE) of the Republic of Korea (NIBR202102103, NIBR202203102).

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

LITERATURE CITED

Dierckxsens, N., P. Mardulyn and G. Smits. 2017. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. Nucleic Acids Research 45: e18.

Jo, S., Y.-K. Kim, S.-H. Cheon, Q. Fan and K.-J. Kim. 2019. Characterization of 20 complete plastomes from the tribe Laureae (Lauraceae) and distribution of small inversions. PLoS ONE 14: e0224622.

Katoh, K. and H. Toh. 2010. Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics 26: 1899–1900.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes and A. Drummond. 2012. Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647–1649.

Park, Y.-J. and K.-S. Cheon. 2021. The complete chloroplast genome sequence of Korean *Neolitsea sericea* (Lauraceae). Korean Journal of Plant Taxonomy 51: 332–336.

Shaw, J. M. H. 2013. The correct name for the Korean willow-leaved spicebush. Hanburyana 7: 32–36.

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312–1313.

Zhao, Q., Y. Zhao and K. Wang. 2006. Antinociceptive and free radical scavenging activities of alkaloids isolated from *Lindera angustifolia* Chen. Journal of Ethnopharmacology 106: 408–413.