

## A chloroplast genome of *Forsythia saxatilis* (Nakai) Nakai, an endemic species in Korea

Sang Chul Choi  and Sangtae Kim 

Department of Biotechnology, Sungshin University, Seoul, Korea

### ABSTRACT

*Forsythia saxatilis* (Nakai) Nakai is an endemic species in Korea. We present the second complete chloroplast genome sequence of *F. saxatilis* showing that the chloroplast genome is 156,376 bp in length containing four subregions: a large single-copy (LSC) region of 87,097 bp, a small single-copy (SSC) region of 17,859 bp, and a pair of inverted repeat regions (IRs) of 25,710 bp. The genome contains 131 genes including 88 protein-coding genes, 35 tRNAs, and eight rRNAs. We found 40 base differences in 17 sites between our chloroplast genome and the previously reported chloroplast genome sequences of *F. saxatilis*. All of the differences were found in a region of 832 bp in the LSC region except for one site, which indicated potential sequencing errors in building the previously reported genome because of little substitutions in the homologous region of the other *Forsythia* species and even of *Abeliophyllum distichum*, a sister to *Forsythia*. The new chloroplast genomes of *F. saxatilis* and those of *F. x intermedia* and *F. viridissima* are identical, which suggests that *F. saxatilis* is maternally closely related to the other two species.

### ARTICLE HISTORY

Received 15 September 2020  
Accepted 6 October 2020

### KEYWORDS

Chloroplast genome;  
endangered species;  
*Forsythia saxatilis*; Oleaceae

*Forsythia* Vahl (Oleaceae) contains 13 species distributed mostly in eastern Asia except for one species in southeastern Europe (Ha et al. 2018). Species of *F. koreana* (Rehder) Nakai, *F. nakaii* (Uyeki) T. B. Lee, *F. ovata* Nakai, and *F. saxatilis* (Nakai) Nakai have been reported as endemic species in Korea (National Institute of Biological Sciences 2013). In particular, most distribution of *F. saxatilis* is in the Bukhansan National Park in Korea as an endangered and flagship species of the park (Kim and Chang 2007). Additional fragmented distributions of *F. saxatilis* are found in limestone area of Gangwon province and Eusung-gun in Gyungbuk province (Han et al. 2011).

The first chloroplast (cp) genome from *F. saxatilis* has been reported in the phylogenetic study of tribe Forsythieae in Oleaceae (Ha et al. 2018). In this study, we found a new population of *F. saxatilis* in Opaesan (N20°27'55.14", W105°17'31.21") where is near from the Bukhansan National Park and determined the second cp genome with a sample from the new population. It provides comparative genomic evidence that the first cp genome could be affected by potential sequencing errors.

We collected a plant and leaves were preserved in silica gel. A voucher specimen was deposited in the herbarium of the Sungshin University, Korea (S. Kim 2020-001, SWU). DNA extraction was performed using a commercial plant DNA extraction kit (Exgene™; GeneAll, Seoul, Korea), and the extracted DNA was sequenced using the MGISEQ-2000 platform (MGI Tech Co. Ltd., Shenzhen, China) with paired-end

reads (150 bp in length). The sequencing library was constructed using the MGIEasy DNA library Prep Kit (MGI Tech Co. Ltd., Shenzhen, China) with an insert size of 500 bp.

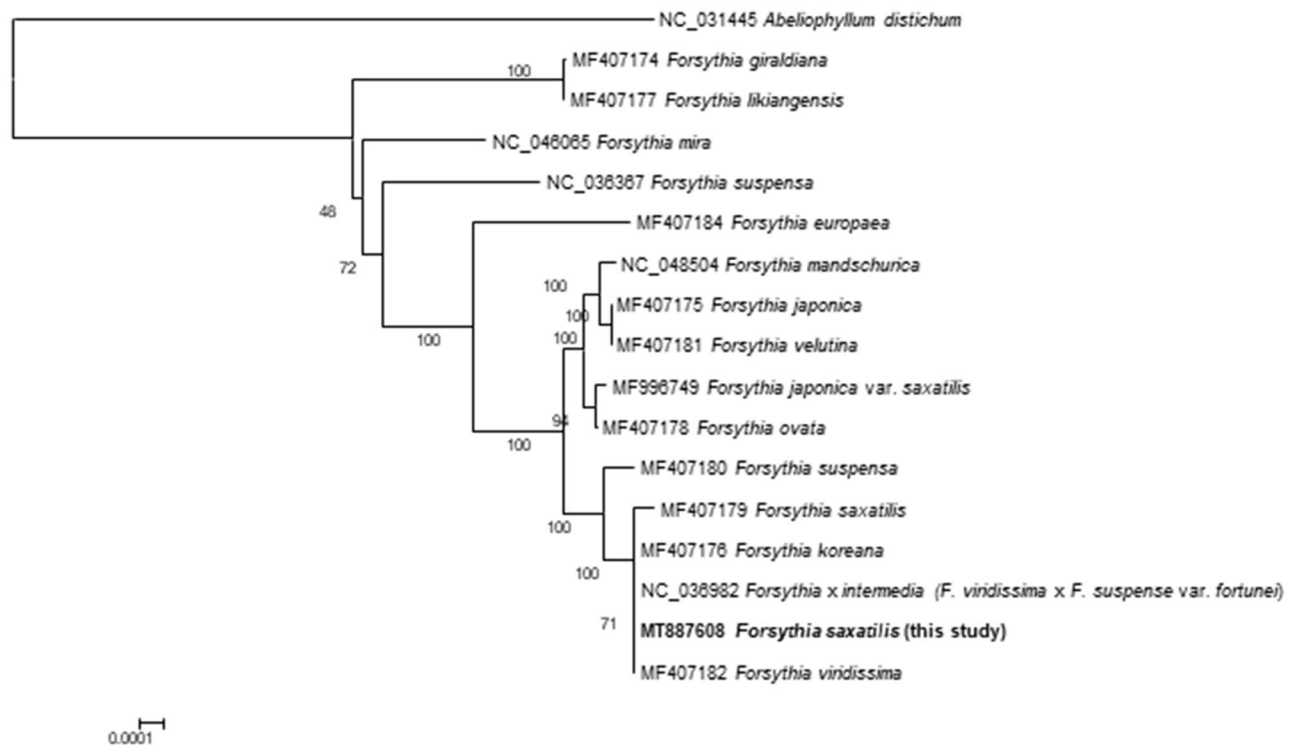
The sequencing generated a total of 20,477,062 raw reads. After short-read quality filtering, the reads were mapped against the previously reported cp genome from *F. saxatilis* (MF407179) using Geneious version 9.1.8 (Kearse et al. 2012). We also performed *de novo* assembly of the mapped reads using Geneious, which produced the same sequence as the mapped result. Comparing our cp genome with the previously reported one from *F. saxatilis*, we found 40 base differences in 17 sites. The sites included two short repeating unit differences (TTATAAA and ATGCTA), two three-concatenated base changes, two homopolymer indels, three homopolymer size differences, and eight single-nucleotide substitutions between the two cp genomes. Remarkably, all of the differences except for one site were found in the intergenic region between *aptF* and *atpH* and parts of the two genes, which was a relatively short region (832 bp) of the large single-copy (LSC). Furthermore, there were not so many substitutions in the region of the other *Forsythia* species and even of *Abeliophyllum distichum*, a sister to *Forsythia* in Oleaceae (Ha et al. 2018). Therefore, we believe that this region in the previously reported cp genome from *F. saxatilis* is likely to be affected by potential sequencing errors.

The assembled cp genome (MT887608) was annotated using Geneious version 9.1.8 (Kearse et al. 2012) based upon the annotation of *F. mandshurica* (NC\_048504.1), one of the

**CONTACT** Sangtae Kim  [amborella@sungshin.ac.kr](mailto:amborella@sungshin.ac.kr)  Department of Biology, Sungshin University, Seoul 01133, Korea

© 2020 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.



**Figure 1.** A maximum-likelihood tree based on an alignment of entire cp genome sequences that are from *F. saxatilis*, 15 species of *Forsythia*, and *Abeliophyllum distichum* as an outgroup.

reference cp genomes in *Forsythia*. The size of the cp genome of *F. saxatilis* was 156,376 bp, consisting of a LSC region of 87,097 bp and a small single-copy (SSC) region of 17,859 bp separated by each of the two identical inverted repeat regions (IRs) of 25,710 bp. The cp genome encoded 88 protein-coding genes, 35 tRNAs, and eight rRNAs.

A phylogenetic analysis was performed using (1) the new cp genome of *F. saxatilis*, (2) 15 cp genomes of *Forsythia* species, and (3) the cp genome of *Abeliophyllum distichum* Nakai (NC\_031445) (Figure 1). These 17 sequences were aligned by the MAFFT program (Kato and Standley 2013) and a maximum-likelihood analysis was performed to create a phylogenetic tree using IQ-TREE (Nguyen et al. 2015) with the best model (K3Pu + F + I) including 1000 replicates of bootstrapping.

The tree showed that our cp genome was clustered with those of *F. koreana* (MF407176), *F. × intermedia* (= *F. viridissima* × *F. suspense* var. *fortunei*) (NC\_036982), and *F. viridissima* (MF407182) (Figure 1). The previous cp genome of *F. saxatilis* (MF407179) was more distantly related to them than our cp genome. The new cp genome sequence of *F. saxatilis* was exactly the same as those of *F. × intermedia* and *F. viridissima*, and differed only in two bases from that of *F. koreana*. Autapomorphic characters found in the previous cp genome of *F. saxatilis* probably due to potential sequencing errors resulted in the rather longer external branch leading to the previous cp genome of *F. saxatilis*.

We report the second complete chloroplast genome in *F. saxatilis*. This allowed comparison with the previously reported cp genome in the species, suggesting the possibility

of sequencing errors. The base substitution rate and phylogenetic relationship among species in *Forsythia* provided by this study elucidate the evolutionary history such as a hybridization origin.

## Disclosure statement

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

## Funding

This work was supported by the National Research Foundation of Korea [NRF-2017R1D1A1B03034952] to S. Kim.

## ORCID

Sang Chul Choi  <http://orcid.org/0000-0002-2340-8389>  
Sangtae Kim  <http://orcid.org/0000-0002-1821-4707>

## Data availability statement

The data that support the findings of this study are openly available in NCBI at <https://www.ncbi.nlm.nih.gov/> (reference number: MT887608) or available from the corresponding author.

## References

- Ha YH, Kim C, Choi K, Kim JH. 2018. Molecular phylogeny and dating of Forsythieae (Oleaceae) provide insight into the Miocene history of Eurasian temperate shrubs. *Front Plant Sci.* 9:1–15.
- Han SH, Kim DH, Park SY. 2011. *Forsythia saxatilis* Nakai, a rare and endangered species number 166 designated by the Korean Forestry

- Service. Seoul, Korea: National Institute of Forest Science (Korea). ISBN 978-89-8176-775-4 (Korean).
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol E.* 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.
- Kim KJ, Chang CS. 2007. Oleaceae. In: Park CW, editor. The genera of vascular plants of Korea. Seoul, Korea: Academic Publishing Co.
- National Institute of Biological Sciences. 2013. Endemic species of Korea. Incheon, Korea: National Institute of Biological Sciences.
- Nguyen LT, Schmidt HA, Haeseler AV, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol E.* 32(1):268–274.
- Park CW. 2007. The genera of vascular plants of Korea. *Flora of Korea* editorial committee. Seoul, Korea: Academy Publishing Co.; p. 1482.