MITOGENOME ANNOUNCEMENT

The complete chloroplast genome of Abeliophyllum distichum f. lilacinum Nakai (Oleaceae) from the Chungbuk Province, Korea

Jongsun Park*a,b c, Suhyeon Park*a,b c, Taewon Jang*c, Gwanho Kim*d and Jae-Ho Park*c

*aInfoBoss Inc, Seoul, Republic of Korea; bInfoBoss Research Center, Seoul, Republic of Korea; cDepartment of Pharmaceutical Science, JungWon University, Goesan, Republic of Korea; dGoesan Bunjae Nongwon, Goesan, Republic of Korea

ABSTRACT

The chloroplast genome of Abeliophyllum distichum f. lilacinum Nakai, classified to a monotypic in this genus, and an endemic species in Korea, was sequenced to understand the genetic differences among intraspecies and cultivars of A. distichum. The chloroplast genome length is 156,015 bp (GC ratio is 37.8%) and has a typical quadripartite structure: 86,779 bp large single copy (35.8%) and 17,828 bp small single copy (31.9%) regions separated by two 25,704 bp inverted repeat (43.2%) regions. The genome encodes for 133 genes (88 protein-coding genes, eight rRNAs, and 37 tRNAs). Six to 99 SNPs and seven to 18 INDEL regions (19 bp to 72 bp) were identified against available chloroplast genomes of A. distichum. Phylogenetic trees show that A. distichum f. lilacinum is clustered with the Dae Ryun cultivar which has a larger fruit body. Our analyses suggest additional research, such as Genotyping-By-Sequencing, for understanding relationship between morphology and genotype of A. distichum.

CONTACT Jongsun Park; starflr@infoboss.co.kr. (Received 17 March 2021 Accepted 12 May 2021)

ARTICLE HISTORY

Accepted 12 May 2021

KEYWORDS

Abeliophyllum distichum f. lilacinum; chloroplast genome; intraspecific variation; Korea endemic

Abeliophyllum distichum Nakai, a monotypic species in Abeliophyllum (Nakai 1919; Melchior 1964), is an endemic species in Korea (Park et al. 2020). It presents diverse morphological features such as leaves and flowers, resulting in the four forma: i) A. distichum f. lilacinum, pink-colored flower, ii) A. distichum f. edburneum, ivory-colored flower, iii) A. distichum f. viridicalycinum, light-green-colored flower, and iv) A. distichum f. rotundicarpum, round-shaped fruit (Lee 1976). Due to these morphological characteristics, one cultivar was developed and at least two more cultivars are being developed for horticultural purposes (Min et al. 2019; Park et al. 2019; Park et al. 2019). Moreover, A. distichum contains useful phytocompounds including acteoside (Jang et al. 2020; Park et al. 2019) and hirsutrin (Oh et al. 2003). Here, we sequenced and reported the chloroplast genome of A. distichum f. lilacinum to better understand its intraspecific genetic diversity.

Fresh leaves material was collected in GoesanBunjae-Nongwon in Goesan-gun, Chungbuk Province, Republic of Korea (36.797369 N, 127.782997E). Total genomic DNA was extracted from the collected leaves material using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). The voucher was deposited in the InfoBoss Cyber Herbarium (IN; http://herbarium.infoboss.co.kr/; Voucher number: IB-01092; Contact: Suhyeon Park; shpark817@infoboss.co.kr). Genome sequencing was conducted using NovaSeq6000 at Macrogen Inc., Korea, and de novo assembly and sequence confirmation were conducted using Velvet v1.2.10 (Zerbino and Birney 2008), GapCloser v1.12 (Zhao et al. 2011), BWA v0.7.17 (Li 2013), and SAMtools v1.9 (Li et al. 2009) under the environment of Genome Information System (GelS; http://geis.infoboss.co.kr/). Geneious Prime® v2020.2.4 (Biomatters Ltd., Auckland, New Zealand) was used for chloroplast genome annotation based on the A. distichum chloroplast genome (MN116559; Park, Min, et al. 2019).

The chloroplast genome of A. distichum f. lilacinum (GenBank accession is MW426545) is 156,015 bp (GC ratio is 37.8%) and has four subregions: 86,779 bp of large single copy (GC ratio is 35.8%) and 17,828 bp of small single copy (GC ratio is 31.9%) regions are separated by two 25,704 bp inverted repeat (IR; GC ratio is 43.2%). The chloroplast genome contains 133 genes consisting of 88 protein-coding genes, eight rRNAs, and 37 tRNAs. Nineteen genes comprising seven protein-coding genes, four rRNAs, and seven tRNAs are duplicated in the IR regions.

Seven single nucleotide polymorphisms (SNPs) and nine insertion and deletion (INDEL) regions (22 bp in total) were found between A. distichum f. lilacinum and A. distichum Dae Ryun (MN116559) using ‘Find Variations/SNPs’ function in Geneious Prime® v2020.2.4 (Biomatters Ltd., Auckland, New Zealand) based on the pair-wise alignment of the two chloroplast genomes. In addition, 6 SNPs and 7 INDEL regions (19 bp) were found against Sang Jae and Ok Hwang 1ho sequences (MN127986 and MK616470) and 99 SNPs and 18 INDEL regions (72 bp) and 6 SNPs and 8 INDEL regions (21 bp) were identified according to NC_031445 and MF407183.
respectively. These numbers of intraspecific variations except those against NC_031445 are smaller than those identified between Korean isolates of other plant species (Cho et al. 2019; Kim et al. 2019; Oh et al. 2019a, 2019b; Park and Kim 2019; Park et al. 2019a, 2019b; Choi et al. 2020; Park et al. 2021), suggesting that genetic diversity in *A. distichum* may not be sufficient enough so that morphological features are not tightly linked to those intraspecific variations on chloroplast genomes.

Twenty-three chloroplast genomes representing *Abeliophyllum*, *Forsythia*, and *Fontanesia* genera were used for constructing the bootstrapped Maximum-Likelihood (ML; bootstrap repeat is 1,000) and Bayesian Inference (BI; number of generations is 1,100,000) phylogenetic trees using MEGA X (Kumar et al. 2018) and MrBayes v3.2.7a (Ronquist et al. 2012), respectively, after aligning the whole chloroplast genomes with MAFFT v7.450 (Katoh and Standley 2013). ML and BI phylogenetic trees show that *A. distichum* f. *lilacinum* is clustered with the Dae Ryun cultivar (MN116559; Figure 1) which displays a flower that is bigger than that of the wild type of *A. distichum* (Park et al. 2019). The phylogenetic reconstructions also show that the other *A. distichum* forma chloroplast genomes are clustered with relatively low supportive values of ML and BI trees (54 and 0.8717, respectively, in Figure 1) or are incongruent in both trees (the clade covering NC_031445, MN127986, and MK616470 in Figure 1). These analyses indicate that phenotypic variations of *A. distichum* are not clearly resolved using whole chloroplast genome sequences like the case of *Sueda japonica* (Kim et al. 2020), requiring other genome-wide analysis methods such as Genotyping-By-Sequencing (GBS) which has been applied to various plant species to find relationships between morphological features (Wong et al. 2015; Guajardo et al. 2020).

**Disclosure statement**

The authors declare that they have no competing interests.

**Funding**

This work was supported by the Jungwon University Research Grant [2020-025].

**ORCID**

Jongsun Park [http://orcid.org/0000-0003-0786-4701](http://orcid.org/0000-0003-0786-4701)

Jae-Ho Park [http://orcid.org/0000-0003-0942-7871](http://orcid.org/0000-0003-0942-7871)

**Data availability statement**

Chloroplast genome sequence can be accessed via accession number MW426545 in GenBank of NCBI at [https://www.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov). The associated BioProject, SRA, and Bio-Sample numbers are PRJNA689053, SAMN17198782, and SRR13340840, respectively.
