

MITOGENOME ANNOUNCEMENT

 OPEN ACCESS

The complete plastome sequence of *Diospyros blancoi* A. DC. (Ebenaceae)

Sangjin Jo, Hoe-Won Kim, Young-Kee Kim, Se-Hwan Cheon and Ki-Joong Kim

School of Life Sciences, Korea University, Seoul, Korea

ABSTRACT

The plastome sequences of *Diospyros blancoi* A. DC. (Ebenaceae) were completed in this study (NCBI acc. no. KX426216). The gene order and structure of the *D. blancoi* plastome are collinear with the typical plastome of land plants. The complete plastome size is 157,745 bp in length and consists of a large single-copy region of 87,246 bp and a small single-copy region of 18,323 bp, which are separated by a pair of 26,088 bp-long inverted repeat regions. The overall A-T content of the plastome sequence is 62.6%. The plastome contains 113 genes, of which 79 are protein-coding genes, 30 are tRNA genes, and 4 are rRNA genes. Sixteen genes contain one intron and two genes have two introns. A total of 45 simple sequence loci were identified from the genome. Phylogenetic analysis revealed that *D. blancoi* is a sister group of Primulaceae with 100% bootstrap support.

ARTICLE HISTORY

Received 28 July 2016
Accepted 29 July 2016

KEYWORDS

Plastome; tropical fruit; Ebenaceae; *Diospyros blancoi*

Diospyros blancoi A. DC. is commonly referred to as velvet apple. It is a widely cultivated tropical Asian fruit that originated in Philippine (Kim 2011). But, this is a rare plant in wild population. The fruit and stem of this species is used for both food and timber. *Diospyros blancoi* belongs to the family Ebenaceae of Ericales (Byng et al. 2016). Ericales consist of twenty-two families (Byng et al. 2016). Ericaceae consist of four genera and approximately 800 species (Christenhusz & Byng 2016). Most of them belong to the genus *Diospyros*. The recent studies reported a few plastome from *Diospyrus*, but no sequence data was released from the NCBI so far (Turner et al. 2016; Fu et al. 2016). Furthermore, the two studies did not include the *D. blancoi*. The complete plastome sequence of *D. blancoi* will aid us in developing molecular markers for the identification and improvement of cultivars of this economically important tropical fruit species.

The leaves of *D. blancoi* used in this study were collected from the Korea University greenhouse, where we grew the plants from the seeds that were originally collected from the Philippines. The plants flowered and fruited in the greenhouse. A voucher specimen was deposited in the Korea University Herbarium (KUS acc. no. 2014-0240). Fresh leaves were ground into powder in liquid nitrogen and total DNAs were extracted using the CTAB method (Doyle & Doyle 1987). The DNAs were further purified by ultracentrifugation and dialysis (Palmer 1986). The genomic DNAs are deposited in the Plant DNA Bank in Korea (PDBK acc. no. 2014-0240). The complete plastome sequence was generated using an Illumina HiSeq 2000 system (Illumina, Inc., San Diego, CA). Annotations were performed using the National Center for Biotechnology Information (NCBI) BLAST, DOGMA (Wyman et al. 2004), and tRNAscan-SE programs (Lowe & Eddy 1997).

For the phylogenetic analysis, we selected and downloaded 24 complete plastome sequences based on the APG IV system (Byng et al. 2016) from the NCBI database. All the 25 plastome sequences were from super-asterid plants.

The gene order and structure of the *D. blancoi* plastome are similar to those of a typical angiosperm (Shinozaki et al. 1986; Kim & Lee, 2004; Yi & Kim 2012). The complete plastome is 157,745 bp in length, and consists of a large single-copy (LSC) region of 87,246 bp and a small single-copy (SSC) region of 18,323 bp separated by two inverted repeats (IR) of 26,088 bp. The published *D.* plastome sizes are ranged from 157,300 bp to 157,784 bp (Turner et al. 2016; Fu et al. 2016). Therefore, the plastome size of *D. blancoi* also belongs to the suggested range. The plastome of *D. blancoi* comprises 113 unique genes (79 protein-coding genes, 30 tRNA genes, and 4 rRNA genes). Seven protein-coding, seven tRNA, and four rRNA genes are duplicated in the IR regions. The major portion of the *D. blancoi* plastome consists of protein-coding genes (54.8%), tRNA genes (1.8%), and rRNA genes (5.8%). The average A-T content of the plastome is 62.6%, whereas that in the LSC, SSC, and IR regions is 64.7%, 69.1%, and 56.9%, respectively. Sixteen genes have one intron and two genes, *ycf3* and *clpP*, comprising two introns. A total of 45 simple sequence repeat (SSR) loci, which can be defined as having more than 10 duplications of simple nucleotide(s), are scattered among the plastome. Among these, 40, 3, and 2 are mono-SSR, di-SSR, and tri-SSR loci, respectively. Some of these loci will be useful in identifying cultivars of *D. blancoi*.

To validate the phylogenetic relationships of *D. blancoi* in asterids, we constructed a maximum-likelihood (ML) tree by using 25 super-asterid taxa. Phylogenetic analysis was performed on a data set that included 79 protein-coding genes

CONTACT Ki-Joong Kim  kimkj@korea.ac.kr  School of Life Sciences, Korea University, Seoul 02841, Korea

© 2016 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.

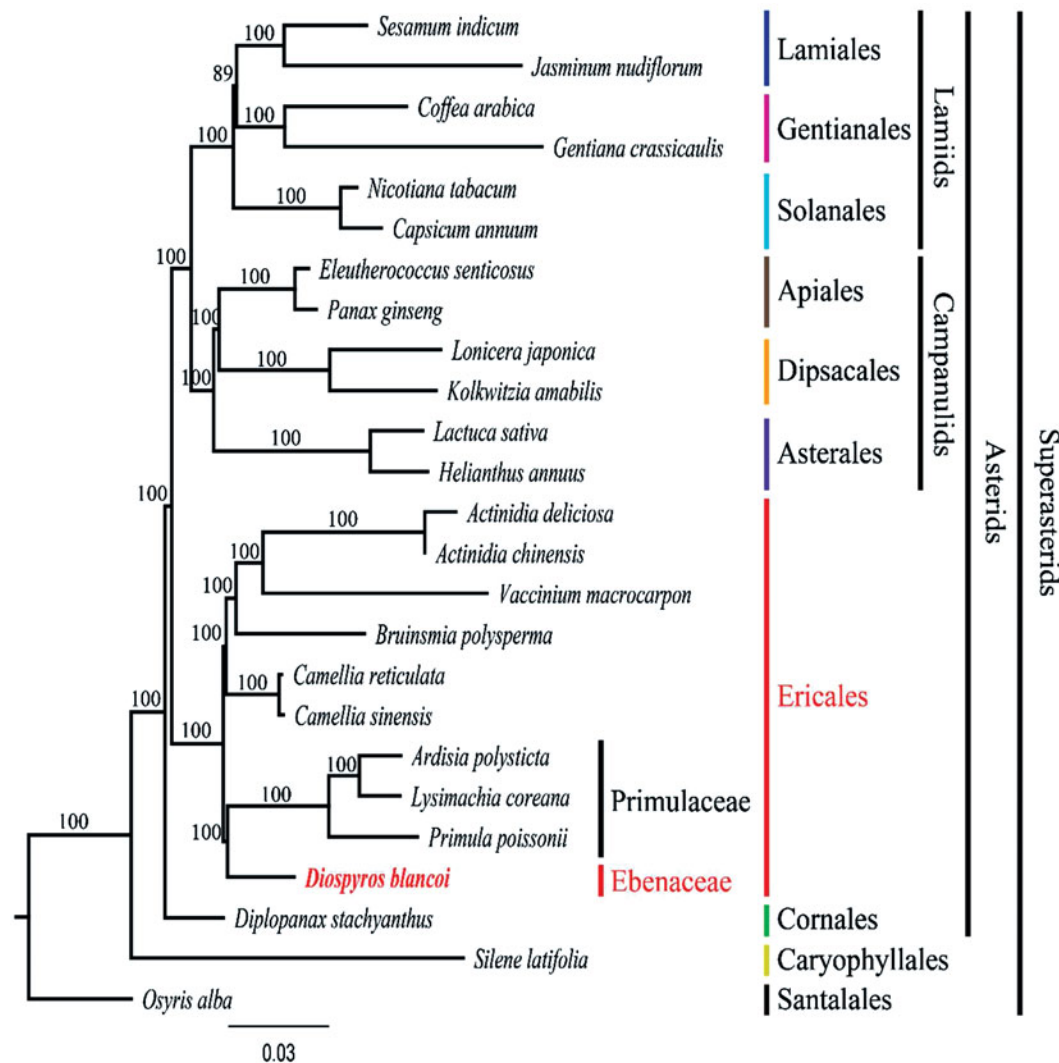


Figure 1. Maximum-likelihood (ML) tree based on 79 protein-coding and 4r rRNA genes from 25 plstomes as determined by RAxML (–ln L = –413544.670182). The numbers at each node indicate the ML bootstrap values. Genbank accession numbers of taxa are shown below, *Actinidia chinensis* (NC_026690), *A. deliciosa* (NC_026691), *Ardisia polysticta* (NC_021121), *Bruinsmia polysperma* (NC_030180), *Camellia reticulata* (NC_024663), *C. sinensis* (NC_020019), *Capsicum annuum* (NC_018552), *Coffea arabica* (NC_008535), *D. blancoi* (NC_022408), *Diplopanax stachyanthus* (NC_029750), *Eleutherococcus senticosus* (NC_016430), *Gentiana crassicaulis* (NC_027442), *Helianthus annuus* (NC_007977), *Jasminum nudiflorum* (NC_008407), *Kolkwitzia amabilis* (NC_029874), *Lactuca sativa* (NC_007578), *Lonicera japonica* (NC_026839), *Lysimachia coreana* (NC_026197), *Nicotiana tabacum* (NC_001879), *Osyris alba* (NC_027960), *Panax ginseng* (NC_006290), *Primula poissonii* (NC_024543), *Sesamum indicum* (NC_016433), *Silene latifolia* (NC_016730), and *Vaccinium macrocarpon* (NC_019616).

and 4 rRNA genes from the 25 taxa using RAxML v. 7.7.1 (Stamatakis et al. 2008). The 80 gene sequences (81,895 bp in length) were aligned with the MUSCLE program using Geneious v. 6.1.8 (Biomatters Ltd.). As a result, *D. blancoi* form a clade with Primulaceae with a 100% bootstrap value (Figure 1). This results are similar to the recent updated phylogeny of the group (www.mobot.org/MOBOT/research/APweb/). However, the results are somewhat different from other data (Fu et al. 2016). Fu et al. (2016) data suggested that Primulaceae is the basal family of the Ericales, and Ebenaceae is the sister group to other members of Ebenales. Our complete plastome sequence of *D. blancoi* will be helpful to resolve not only the interfamilial relationships of Ericales, but also the species relationship within a *Diospyros* (Ebenaceae).

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

This work was supported by the National Research Foundation of Korea (NRF) [NRF-2015M3A9B8030588 and NRF-2015M3A9B8047398]

References

- Byng JW, Chase MW, Christenhusz MJ, Fay MF, Judd WS, Mabberley DJ, Sennikov AN, Soltis DE, Soltis PS, Stevens PF. 2016. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc. 181:1–20.
- Christenhusz MJ, Byng JW. 2016. The number of known plant species in the world and its annual increase. Phytotaxa. 261:201–217.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19:11–15.
- Fu J, Liu H, Hu J, Liang Y, Liang J, Wuyun T, Tan X. 2016. Five complete chloroplast genome sequences from diospyros: genome organization and comparative analysis. PLoS One. 11:e0159566.
- Kim K-J, Lee HL. 2004. Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. DNA Res. 11:247–261.

- Kim K-J. 2011. Tropical fruit resources. Seoul: Geobook.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955–964.
- Palmer JD. 1986. Isolation and structural analysis of chloroplast DNA. *Method Enzymol.* 118:167–186.
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *Embo J.* 5:2043–2049.
- Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol.* 57:758–771.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.
- Yi D-K, Kim K-J. 2012. Complete chloroplast genome sequences of important oilseed crop *Sesamum indicum* L. *PLoS One.* 7:e35872.
- Turner B, Paun O, Munzinger J, Chase MW, Samuel R. 2016. Sequencing of whole plastid genomes and nuclear ribosomal DNA of *Diospyros* species (Ebenaceae) endemic to New Caledonia: many species, little divergence. *Ann Botany.* 117:1175–1185.