


The complete chloroplast genome of *Euscaphis japonica* (Thunb.) Kanitz (Staphyleaceae) isolated in Korea

Sang-Hun Oh^a  and Jongsun Park^{b,c} 

^aDepartment of Biology, Daejeon University, Daejeon, Republic of Korea; ^bInfoBoss Inc., Seoul, Republic of Korea; ^cInfoBoss Research Center, Seoul, Republic of Korea

ABSTRACT

The complete chloroplast genome of *Euscaphis japonica* (Thunb.) Kanitz isolated in Korea is 160,606 bp long and has four subregions: 89,232 bp of large single-copy and 18,734 bp of small single-copy regions are separated by 26,320 bp of inverted repeat regions including 129 genes (84 CDS, 8 rRNAs, and 37 tRNAs) and three pseudogenes. There were 424 SNPs and 809 INDELS compared with the Chinese *E. japonica*, useful to develop markers for phylogeographic study of the species. Phylogenetic trees show that *E. japonica*, representing Crossosomatales, is nested within the Malvids clade, confirming previous studies.

ARTICLE HISTORY

Received 14 July 2020
Accepted 8 October 2020

KEYWORDS

Crossosomatales; *Euscaphis japonica*; intraspecific variations; Jeju Island; Korea

Euscaphis japonica (Thunb.) Kanitz, commonly called the Korean sweetheart tree, is a deciduous tree distributed in warm temperate regions of Korea, China, Japan, and Taiwan, and widely cultivated in garden for its showy fruits. It is characterized by having follicular fruits and arillate seeds in Staphyleaceae (Li et al. 2008), a member of the recently re-circumscribed angiosperm order, Crossosomatales (Oh and Potter 2006). A recent phylogenetic analysis of angiosperms placed Crossosomatales as sister to a large clade of Picramniales, Malvales, Brassicales, Huerteales, and Sapindales (Soltis et al. 2011). We determined complete chloroplast genome of *E. japonica* to confirm the phylogenetic position of Crossosomatales and to investigate intraspecific variations (Xiang et al. 2019).





Total DNA of *E. japonica* collected on Jeju Island in Korea was extracted from fresh leaves by using a DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). The voucher specimen was deposited in the herbarium of Daejeon University (TUT) (*Oh* 7720; 33°17'24.35"N, 126°26'34.41"E). Genome sequencing and *de novo* assembly were done by the method described in Park et al. (2019b). Genome annotation was based on *E. japonica* chloroplast genome (MN159078) using Geneious R11 v11.0.5 (Biomatters Ltd, Auckland, New Zealand).

Chloroplast genome of *E. japonica* (GenBank accession is MT702885) is 160,606 bp (GC ratio is 37.3%) and has four subregions: 89,232 bp of large single-copy (LSC; 35.4%) and 18,734 bp of small single-copy (SSC; 31.6%) regions are separated by 26,320 bp of inverted repeat (IR; 42.7%). It contains 129 genes (84 protein-coding genes, 8 rRNAs, and 36 tRNAs) and three pseudogenes; 17 genes (six protein-coding genes,

four rRNAs, and seven tRNAs) and one pseudogene are duplicated in IR regions.

Based on pairwise alignment with the chloroplast of *E. japonica* from China (MN159078), 424 SNPs (0.26%) and 809 INDELS (0.50%) were identified, displaying relatively large amount of intraspecific variations: larger than those of *Pyrus ussuriensis* (Cho et al. 2019), *Agrimonia pilosa* (Heo et al. 2020), *Camellia japonica* (Park et al. 2019b), and *Marchantia polymorpha* (Kwon et al. 2019). However, it is less than those of *Goodyera schlechtendaliana* (844 SNPs and 2045 INDELS; Oh et al. 2019b, 2019a) and *Gastrodia elata* (493 SNPs and 650 INDELS; Kang et al. 2020; Park et al. 2020) from Orchidaceae. Our chloroplast genome displays that direction of both LSC and SSC is reversely oriented compared with the chloroplast from China (MN159078), showing a structural variation within a species (Park et al. 2019a; Xi et al. 2019).

Seventy-two conserved genes from 11 chloroplast genomes including two *E. japonica* chloroplast genomes were included in phylogenetic analysis using the Bayesian inference using MrBayes v3.2.7a (Huelsenbeck and Ronquist 2001) and the maximum likelihood method using IQ-TREE v1.6.12 (Nguyen et al. 2015). The phylogenetic trees show that the Korean plant of *E. japonica* is sister to the Chinese accession of the species (Figure 1). Our phylogenetic analyses of chloroplast genomes confirm that *E. japonica* representing Crossosomatales is sister to the clade of Brassicales, Malvales, Huerteales, and Sapindales, consistent with previous studies (Soltis et al. 2011; Chase et al. 2016). Our chloroplast genome can be utilized to understand phylogeographic pattern of the warm temperate tree.

CONTACT Sang-Hun Oh  soh42@dju.kr  Department of Biology, Daejeon University, Daejeon 34520, Republic of Korea; Jongsun Park  starflr@infoboss.co.kr  InfoBoss Inc., 301 Room, Haeun Bldg., 670, Seolleung-ro, Gangnam-gu, Seoul, Republic of 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.

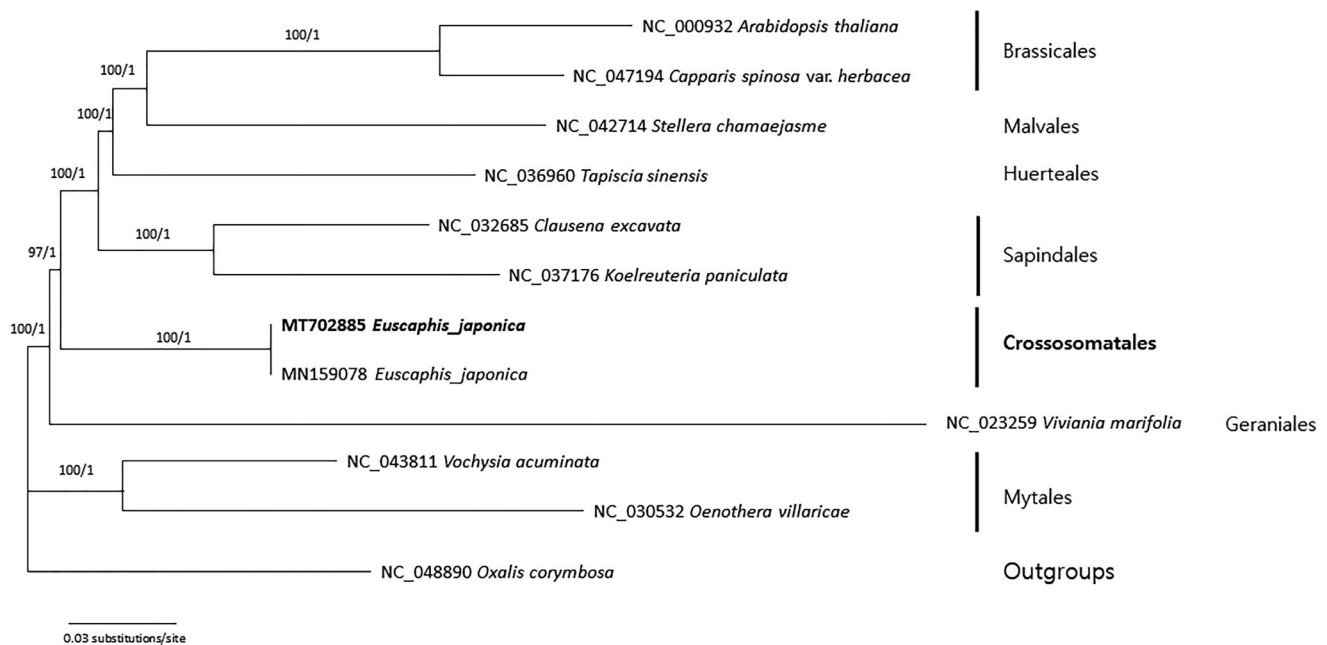


Figure 1. A maximum-likelihood tree from the analysis of 72 conserved genes, originated from two accessions of *Euscaphis japonica* (MT702885 in this study and MN159078), *Arabidopsis thaliana* (NC_000932), *Capparis spinosa* var. *herbacea* (NC_047194), *Clausena excavata* (NC_032685), *Koelreuteria paniculata* (NC_037176), *Oenothera villaricae* (NC_030532), *Stellera chamaejasme* (NC_042714), *Tapiscia sinensis* (NC_036960), *Vochysia acuminata* (NC_043811), and an outgroup species (*Oxalis corymbosa* (NC_048890)). The GTR model with gamma rates was used as a molecular model to construct the ML tree. A heuristic search was used with nearest-neighbor interchange (NNI) branch swapping, the Tamura-Nei model, and uniform rates among sites to construct maximum likelihood phylogenetic tree. All other options were set to their default values in IQ-TREE. Bootstrap analyses with 1000 pseudoreplicates were conducted with the same options. The same topology was produced from Bayesian inference. A Markov-chain Monte Carlo (MCMC) algorithm was employed for 1,000,000 generations, sampling trees every 200 generations, with four chains running simultaneously. Values above branches are bootstrap supports from the analysis of maximum likelihood and posterior probabilities. Names of orders followed APG IV (Chase et al. 2016).

Disclosure statement

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

Funding

This work was supported by a research grant from the National Research Foundation of Korea [NRF-2020R111A3068464] to the first author.

ORCID

Sang-Hun Oh  <http://orcid.org/0000-0001-8778-4458>

Jongsun Park  <http://orcid.org/0000-0003-0786-4701>

Data availability statement

The chloroplast genome in this study is available in the NCBI GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) with an accession number, MT754180.

References

- Chase MW, Christenhusz M, Fay M, Byng J, Judd WS, Soltis D, Mabberley D, Sennikov A, 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):1–20.
- Cho M-S, Kim Y, Kim S-C, Park J. 2019. The complete chloroplast genome of Korean *Pyrus ussuriensis* Maxim. (Rosaceae): providing genetic background of two types of *P. ussuriensis*. *Mitochondrial DNA Part B.* 4(2): 2424–2425.
- Heo K-I, Park J, Xi H, Min J. 2020. The complete chloroplast genome of *Agrimonia pilosa* Ledeb. isolated in Korea (Rosaceae): investigation of intraspecific variations on its chloroplast genomes. *Mitochondrial DNA Part B.* 5(3):2264–2266.
- Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics.* 17:754–755.
- Kang M-J, Kim S-C, Lee H-R, Lee S-A, Lee J-W, Kim T-D, Park E-J. 2020. The complete chloroplast genome of Korean *Gastrodia elata* Blume. *Mitochondrial DNA Part B.* 5(1):1015–1016.
- Kwon W, Kim Y, Park J. 2019. The complete chloroplast genome of Korean *Marchantia polymorpha* subsp. *ruderalis* Bischl. & Boisselier: low genetic diversity between Korea and Japan. *Mitochondrial DNA Part B.* 4(1):959–960.
- Li D, Cai J, Wen J. 2008. Staphyleaceae. In: Wu ZY, Raven PH, Hong DY, editors. *Flora of China*. Vol. 11: Oxalidaceae through Aceraceae. Beijing; St. Louis: Science Press; Missouri Botanical Garden Press; p. 498–504.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Oh S-H, Potter D. 2006. Description and phylogenetic position of a new angiosperm family, Guamatelaceae, inferred from chloroplast *rbcl*, *atpB*, and *matK* sequences. *Syst Bot.* 31(4):730–738.
- Oh S-H, Suh HJ, Park J, Kim Y, Kim S. 2019a. The complete chloroplast genome sequence of a morphotype of *Goodyera schlechtendaliana* (Orchidaceae) with the column appendages. *Mitochondrial DNA Part B.* 4(1):626–627.
- Oh S-H, Suh HJ, Park J, Kim Y, Kim S. 2019b. The complete chloroplast genome sequence of *Goodyera schlechtendaliana* in Korea (Orchidaceae). *Mitochondrial DNA Part B.* 4(2):2692–2693.
- Park J, Kim Y, Xi H. 2019a. The complete chloroplast genome sequence of male individual of Korean endemic willow, *Salix koriyanagi* Kimura ex Goerz (Salicaceae). *Mitochondrial DNA Part B.* 4(1): 1619–1621.
- Park J, Kim Y, Xi H, Oh YJ, Hahm KM, Ko J. 2019b. The complete chloroplast genome of common camellia tree, *Camellia japonica* L. (Theaceae), adapted to cold environment in Korea. *Mitochondrial DNA Part B.* 4(1):1038–1040.

- Park J, Suh Y, Kim S. 2020. A complete chloroplast genome sequence of *Gastrodia elata* (Orchidaceae) represents high sequence variation in the species. *Mitochondrial DNA Part B*. 5(1):517–519.
- Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlswald BS, et al. 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *Am J Bot*. 98(4):704–730.
- Xi H, Park J, Kim Y. 2019. The complete chloroplast genome sequence of rose-gold pussy willow, *Salix gracilistyla* Miq. (Salicaceae). *Mitochondrial DNA Part B*. 4(2):2118–2120.
- Xiang S, Liu X-D, Sun W-H, Lan S-R, Liu Z-J, Zou S-Q. 2019. The complete chloroplast genome sequence of *Euscaphis japonica* (Staphyleaceae). *Mitochondrial DNA Part B*. 4(2):3484–3485.