

The complete chloroplast genome of *Aruncus dioicus* var. *kamtschaticus* (Rosaceae)

Hwa-Jung Suh^a , Juhyeon Min^{b,c} , Jongsun Park^{b,c}  and Sang-Hun Oh^a 

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

ABSTRACT

Aruncus dioicus var. *kamtschaticus* is an economically important herb in the cold temperate regions of East Asia, and displays highly variable morphological features. Completed chloroplast genome of *A. dioicus* var. *kamtschaticus* isolated in Korea is 157,859 bp long with four subregions: 85,972 bp of large single copy and 19,185 bp of small single-copy regions separated by 26,351 bp of inverted repeat regions. The genome includes 131 genes (86 protein-coding genes, eight *rRNAs*, and 37 *tRNAs*). Phylogenetic analyses show that our chloroplast genome was clustered with two partial chloroplast genomes of *A. dioicus*.

ARTICLE HISTORY

Received 27 November 2020
Accepted 16 March 2021

KEYWORDS

Aruncus; chloroplast genome; Korea; Rosaceae





Aruncus dioicus (Walter) Fernald is one of the three species of *Aruncus* L., broadly distributed in the temperate regions of the Northern Hemisphere from central Europe and the Caucasus to the Himalayas, China, Korea, Japan, Russia (Ussuri, Sakhalin, Kamchatka, and Kurile Islands), and North America (Ohwi 1965; Czerepanov 1995; Kim H and Ju 1997; Cuizhi and Alexander 2003; Lee 2006; Shetekauri and Jacoby 2009; Mellichamp 2015). It is a large perennial herb up to 3 m tall with stipulate, pinnately compound leaves, loose panicles with many small, unisexual flowers, and pendulous, follicular fruits. *Aruncus dioicus* var. *kamtschaticus* (Maxim.) H. Hara 1955 is a common East Asian variety, found in mixed forests at high elevation (Ohwi 1965; Cuizhi and Alexander 2003). In Korea, it has been traditionally used for foods (Chung et al. 2010) and for medicinal and cosmetic usages (Kim M-S et al. 2011; Youn et al. 2012; Zhang Q and Kim 2014). It shows a wide range of morphological variations, resulting in taxonomic confusion (Hara 1957; Cuizhi and Alexander 2003). The complete chloroplast genome will be helpful to understand the origin of the economically important plants.

Total DNA was extracted from fresh leaves collected on Mt. Hwangbyeong, Gangwon-do, Republic of Korea (37°42'49.03"N, 128°42'28.39"E; voucher number: Suh 7282 in Daejeon University Herbarium (TUT)) by using DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). Sequencing library was constructed using Illumina TruSeq Nano DNA Library Preparation Kit (Illumina, San Diego, CA) following manufacturer's recommendations with around 350-bp DNA fragments. 1.99-Gbp raw sequences were obtained using HiSeqX at Macrogen Inc., Korea, and filtered by Trimmomatic version

0.33 (Bolger et al. 2014). Chloroplast genome was *de novo* assembled with Velvet version 1.2.10 (Zerbino and Birney 2008) and base-pair confirmation was performed by SOAPGapCloser version 1.12, BWA version 0.7.17, and SAMtools version 1.9 (Li et al. 2009; Li 2013). All processes were conducted under the environment of Genome Information System (GeIS; <http://geis.infoboss.co.kr>). Geneious R11 version 11.0.5 (Biomatters Ltd., Auckland, New Zealand) was used for annotation based on *A. dioicus* chloroplast (KY419942).

Chloroplast genome of *A. dioicus* var. *kamtschaticus* (GenBank accession is MW115132) is 157,859 bp long (GC ratio is 36.4%) and has four subregions: 85,972 bp of large single copy (LSC; 34.2%) and 19,185 bp of small single copy (SSC; 30.0%) regions are separated by 26,351 bp of inverted repeat (IR; 42.4%). It contains 131 genes (86 protein-coding genes, eight *rRNAs*, and 37 *tRNAs*); 19 genes (eight protein-coding gene, four *rRNAs*, and seven *tRNAs*) are duplicated in IR regions. No structural variation was found among the genomes of the species of tribe Spiraeae examined.

Two partial chloroplasts (KY419932 and KY419926) were used for identifying intraspecific variations: 180 SNPs and 69 INDELS against KY419932 and 299 SNPs and 84 INDELS against KY419926 in LSC, SSC, and IRb regions. Numbers of intraspecific variations are relatively large based on the intraspecific variation analysis of chloroplast genomes (Park, Xi, et al. 2020). High levels of intraspecific variations were found in many plant species, such as some species in Orchidaceae (Oh et al. 2019a, 2019b; Park, Suh, et al. 2020; Kang et al. 2020) and Rosaceae (Cho et al. 2019; Heo et al. 2020), *Camellia japonica* (Park et al. 2019), *Euscaphis japonica*

CONTACT Sang-Hun Oh  soh42@du.kr  Department of Biology, Daejeon University, Daejeon, Korea; Jongsun Park  starflr@infoboss.co.kr  InfoBoss Inc, Seoul, Republic of Korea

© 2021 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.

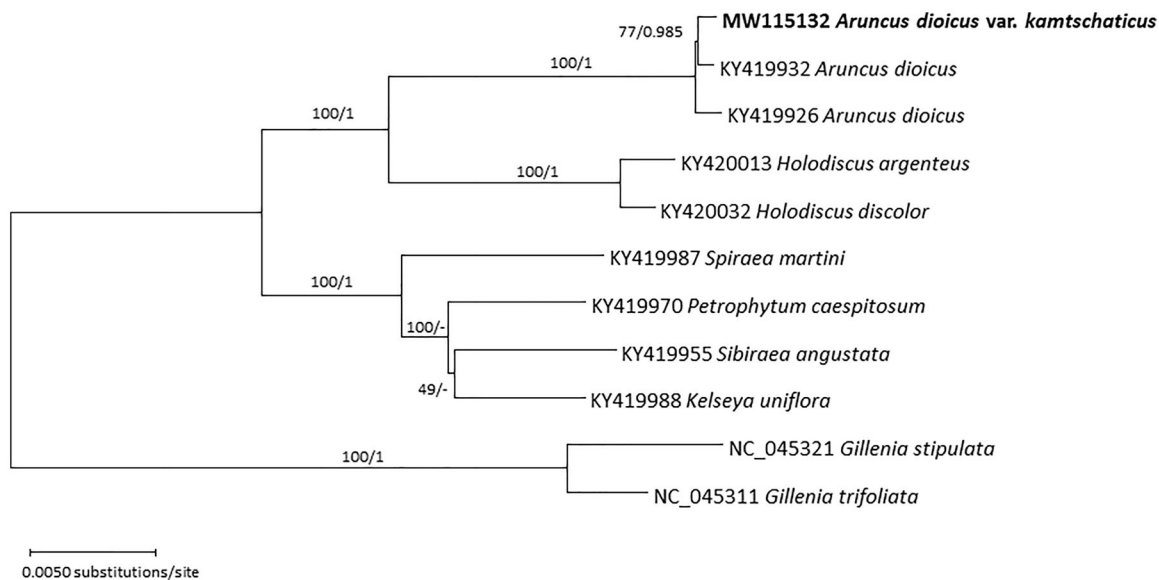


Figure 1. Maximum likelihood phylogenetic trees of ten chloroplast genomes of seven species of tribe Spiraeeae with two *Gillenia* species as outgroups. Phylogenetic tree was drawn based on maximum likelihood tree. The numbers above branches indicate bootstrap support values of maximum likelihood and posterior probability from the BI inference, respectively.

(Oh and Park 2020), *Selaginella tamariscina* (Park, Kim, et al. 2020), and *Marchantia polymorpha* subsp. *ruderalis* (Kwon et al. 2019).

Thirteen chloroplast genomes to represent the major lineages of tribe Spiraeeae (Potter et al. 2007) were used in phylogenetic analysis of the maximum likelihood (ML) and Bayesian inference (BI). Seventy-eight genes of LSC, SSC, and IRb regions were included in the analyses. A heuristic search was used with nearest-neighbor interchange branch swapping, the Tamura-Nei model, and uniform rates among sites to construct ML phylogenetic tree with default values for other options using MEGA X (Kumar et al. 2018). Bootstrap analysis with 1000 pseudoreplicates was also conducted. BI tree was constructed by MrBayes version 3.2.7a (Ronquist et al. 2012). The GTR model with gamma rates was used as a molecular model. A Markov-chain Monte Carlo (MCMC) algorithm was employed for 10,000,000 generations, sampling trees every 200 generations, with four chains running simultaneously.

The phylogenetic tree shows that *A. dioicus* var. *kamtschaticus* clustered with two previously sequenced chloroplasts of *A. dioicus* forming a strongly supported clade (Figure 1). Our new chloroplast genome sequences are determined from the Korean sample. Geographic origin of the two *A. dioicus* published chloroplast genomes (Zhang SD et al. 2017) is unknown, but it is unlikely to be from Korea. The high level of intraspecific variation suggests that the chloroplast genome can be utilized to understand phylogeographic pattern of *A. dioicus* across the whole distribution range as well as relationships among the several varieties of the species.

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 SHO.

ORCID

Hwa-Jung Suh  <http://orcid.org/0000-0003-2528-425X>
 Juhyeon Min  <http://orcid.org/0000-0002-9466-5681>
 Jongsun Park  <http://orcid.org/0000-0003-0786-4701>
 Sang-Hun Oh  <http://orcid.org/0000-0001-8778-4458>

Data availability statement

Chloroplast genome sequence can be accessed via accession number of MW115132 in GenBank of NCBI at <https://www.ncbi.nlm.nih.gov>. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA668550, SAMN16414964, and SRR12807287, respectively.

References

- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15):2114–2120.
- 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.
- Chung H-S, Choi M-G, Moon K-D. 2010. Maintaining quality of wild vegetables (*Aster glehni* and *Aruncus dioicus* var. *kamtschaticus*) from Ulleungdo (Island), Korea by modified atmosphere packaging. *Food Sci Biotechnol*. 19(1):193–199.
- Cuizhi G, Alexander C. 2003. *Aruncus* Linnaeus. *Flora of China*. Vol. 9. Beijing (China): Science Press, St. Louis (MO): Missouri Botanical Garden Press.
- Czerepanov SK. 1995. *Vascular plants of Russia and adjacent states (the former USSR)*. Cambridge: Cambridge University Press.
- Hara H. 1957. Critical notes on some type specimens of East-Asiatic plants in foreign herbaria. *J Jap Bot*. 30(3):65–72.
- 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.
- 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.
- Kim H, Ju I. 1997. *Aruncus*. *Flora Coreana*. Vol. 3. Pyoungyang (North Korea): The Science and Technology Publishing House.
- Kim M-S, Kim K-H, Jo J-E, Choi J-J, Kim Y-J, Kim J-H, Jang S, Yook H-S. 2011. Antioxidative and antimicrobial activities of *Aruncus dioicus* var. *kamtschaticus* Hara extracts. *J Korean Soc Food Sci Nutr*. 40(1):47–55.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 35(6):1547–1549.
- 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.
- Lee Y. 2006. *New Flora of Korea*. Seoul (Korea): Kyo-Hak Publishing Co. Ltd.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. 25(16):2078–2079.
- Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:13033997.
- Mellichamp T. 2015. *Aruncus* Linnaeus. *Flora of North America*. Vol. 9. New York: Oxford University Press.
- Oh S-H, Park J. 2020. The complete chloroplast genome of *Euscaphis japonica* (Thunb.) Kanitz (Staphyleaceae) isolated in Korea. *Mitochondrial DNA Part B*. 5(3):3769–3771.
- 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.
- Ohwi J. 1965. *Flora of Japan*. Washington, DC: Smithsonian Institution; p. 1067.
- Park J, Kim Y, Lee G-H, Park C-H. 2020. The complete chloroplast genome of *Selaginella tamariscina* (Beauv.) Spring (Selaginellaceae) isolated in Korea. *Mitochondrial DNA Part B*. 5(2):1654–1656.
- Park J, Kim Y, Xi H, Oh YJ, Hahm KM, Ko J. 2019. 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.
- Park J, Xi H, Kim Y. 2020. The complete chloroplast genome of *Arabidopsis thaliana* isolated in Korea (Brassicaceae): an investigation of intraspecific variations of the chloroplast genome of Korean *A. thaliana*. *Int J Genomics*. 2020:3236461.
- Potter D, Eriksson T, Evans RC, Oh S-H, Smedmark J, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, et al. 2007. Phylogeny and classification of Rosaceae. *Plant Syst Evol*. 266(1–2):5–43.
- Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 61(3):539–542.
- Shetekauri S, Jacoby M. 2009. *Gebirgsflora & Bäume des Kaukasus*. Koenigstein (Germany): Koeltz Scientific Books.
- Youn J-S, Shin S-Y, Wu Y, Hwang J-Y, Cho J-H, Ha Y-G, Kim J-K, Park M-J, Lee S-H, Kim T-H, et al. 2012. Antioxidant and anti-wrinkling effects of *Aruncus dioicus* var. *kamtschaticus* extract. *Korean J Food Preserv*. 19(3):393–399.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res*. 18(5):821–829.
- Zhang Q, Kim H-Y. 2014. DNA damage protection and anti-inflammatory activity of different solvent fractions from *Aruncus dioicus* var. *kamtschaticus*. *Korean J Plant Res*. 27(6):714–719.
- Zhang SD, Jin JJ, Chen SY, Chase MW, Soltis DE, Li HT, Yang JB, Li DZ, Yi TS. 2017. Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytol*. 214(3):1355–1367.