

A complete chloroplast genome sequence of *Gastrodia elata* (Orchidaceae) represents high sequence variation in the species

Jongsun Park^{a,b} , Youngbae Suh^c and Sangtae Kim^d 

^aInfoboss Co., Ltd, Gangnam-gu, Seoul, Korea; ^bInfoBoss Research Center, Gangnam-gu, Seoul, Korea; ^cNatural Products Research Institute, Seoul National University, Seoul, Korea; ^dDepartment of Biology, Sungshin University, Seoul, Korea

ABSTRACT

Gastrodia elata is a non-photosynthetic saprophytic plant of medicinal use in the oriental countries. We report the second complete chloroplast (cp) genome of *G. elata* from a sample collected in Korea. The length of cp genome is only 35,180 bp: there is no inverted repeated region and many photosynthesis genes are missing compared to typical angiosperm cp genomes. It includes 20 protein-coding genes, 3 rRNAs, and 5 tRNAs. The overall GC content of the genome was 26.7%. Relatively, high intra-specific variation (457 SNPs and 670 indels) is detected in the species comparing it with other seed plants.

ARTICLE HISTORY

Received 21 November 2019
Accepted 23 December 2019

KEYWORDS

Gastrodia; chloroplast genome; *Gastrodia elata*; intra-specific variation; Orchidaceae

Gastrodia elata Blume is a perennial saprophytic species in Orchidaceae. It is distributed in South and East Asia including Nepal, Bhutan, India, China, Taiwan, Japan, and Korea. Stomach-shaped underground rhizome of this plant has been used in traditional oriental medicine, especially for treating headaches, dizziness, tetanus, and epilepsy (Tsaia et al. 2011).

Total genomic DNA was extracted from fresh leaves of *G. elata* (voucher: *S. Kim* 2019-115 deposited in the Herbarium of the Sungshin University; N33°18'02.94", E126°35'31.65"; Is. Jeju, Korea) by using a DNeasy Plant Mini kit (QIAGEN, Hilden, Germany). Genome sequencing was performed using HiSeq4000 at Macrogen Inc. and *de novo* assemblies have performed using both Velvet 1.2.10 (Zerbino and Birney 2008) and SOAPGapCloser (Zhao et al. 2011) under the in-house bioinformatics pipeline. Geneious R11.0.5 (Biomatters Ltd., Auckland, New Zealand) was used for genome annotation based on previously reported *G. elata* chloroplast (cp) genome (NC_037409; Yuan et al. 2018).

The length of the assembled cp genome of *G. elata* (Genbank accession number: MN296709) is 35,810 bp. In comparison to the previous report on cp genome from the Chinese plant of *G. elata* (NC_037409; Yuan et al. 2018), there are no major structural changes in this genome. As previously reported in *G. elata*, many genes related to photosynthesis are also missing and there are no inverted repeat regions. It contains 28 genes containing 20 protein-coding genes, 3 rRNAs, and five tRNAs. The overall GC content of *G. elata* was 26.7%.

Intra-specific cp genome variation in *G. elata* is considerably high compared to those in other Orchidaceae species.

Four-hundred and fifty-seven single nucleotide polymorphisms (SNPs; 1.27%) and 670 insertions and deletions (INDELs; 1.87%) are identified. In another case of the Orchidaceae, *Goodyera schlechtendaliana* (LC085346 and NC_029364; Oh et al., 2019) showed 700 SNPs (0.45%) and 2133 INDELs (1.38%). Intra-specific cp genome variation in Orchidaceae is significantly higher than those in other seed plants: 121 SNPs (0.076%) and 781 INDELs (0.49%) in *Pyrus ussuriensis* (Rosaceae; Cho et al. 2019); 48 SNPs (0.031%) and 58 INDELs (0.037%) in *Duchesnea chrysantha* (Rosaceae; Park, Kim, Lee 2019); 93 SNPs (0.060%) and 56 INDELs 0.036% in *Abeliophyllum distichum* (Oleaceae; Park, Kim, Xi, Jang, et al. 2019); 84 SNPs (0.056%) and 125 INDELs (0.084%) in *Pseudostellaria palibiniana* (Caryophyllaceae; Kim et al. 2019); 78 SNPs (0.050%), and 643 INDELs (0.41%) in *Camellia japonica* (Theaceae; Park, Kim, Xi, Oh, et al. 2019); 21 SNPs (0.015%) and 114 INDELs (0.080%) in *Illicium anisatum* (Illiciaceae; Park, Kim, Xi 2019); 84 INDELs (0.054%) in *Coffea Arabica* (Rubiaceae; Park, Kim, Xi, Heo 2019).

Phylogenetic analysis was performed for 41 previously reported cp genomes in Orchioieae and Epidendroideae including the cp genome of the current study and seven out-group taxa in Cyprapedioideae using nine conserved chloroplast genes (*rpl20*, *ycf2*, *rps18*, *rps8*, *rpl16*, *rps7*, *clpP*, *rps19*, and *rpl14*). The tree showed that that two *G. electa* cp genomes were placed together to form a clade, which is nested in the clade of Orchioieae and Epidendroideae with a long autapomorphic branch (Figure 1). *Calanthe triplicate* and *C. davidii* were sister groups to the clade of *G. eleata*.

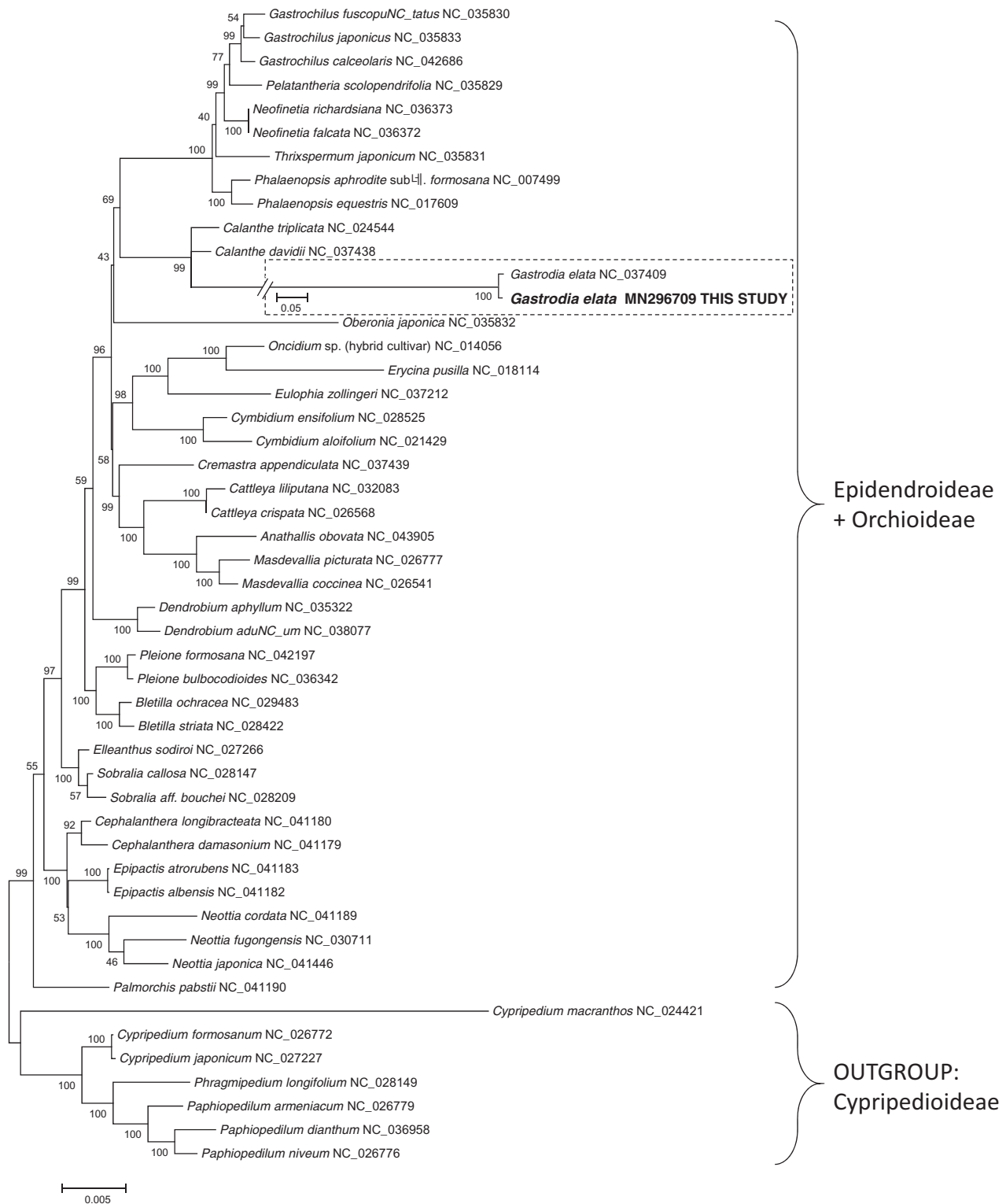


Figure 1. A maximum-likelihood tree based on "GTR + gamma + I" model using MEGA X (Kumar et al. 2018). Outgroup taxa were selected based on Górnaiak et al. (2010). The dashed box indicates a different-scaled region from the main tree. Numbers above the nodes indicate supporting values in 500 replicates of bootstrap analysis.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Jongsun Park  <http://orcid.org/0000-0003-0786-4701>
Sangtae Kim  <http://orcid.org/0000-0002-1821-4707>

References

- 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*. Mitochondr DNA B. 4(2): 2424–2425.
- Górnaiak M, Paun O, Chase MW. 2010. Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene,

- Xdh*: congruence with organellar and nuclear ribosomal DNA results. *Mol Phylogenet Evol.* 56(2):784–795.
- Kim Y, Heo K-I, Park J. 2019. The second complete chloroplast genome sequence of *Pseudostellaria palibiniana* (Takeda) Ohwi (Caryophyllaceae): intraspecies variations based on geographical distribution. *Mitochondr DNA B.* 4(1):1310–1311.
- 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.
- Oh S-H, Suh HJ, Park J, Kim Y, Kim S. 2019. The complete chloroplast genome sequence of *Goodyera schlechtendaliana* in Korea (Orchidaceae). *Mitochondr DNA B.* 4(2):2692–2693.
- Park J, Kim Y, Lee K. 2019. The complete chloroplast genome of Korean mock strawberry, *Duchesnea chrysantha* (Zoll. & Moritz) Miq.(Rosoidae). *Mitochondr DNA B.* 4:864–865.
- Park J, Kim Y, Xi H, Heo K-I. 2019. The complete chloroplast genome of ornamental coffee tree, *Coffea arabica* L. (Rubiaceae). *Mitochondr DNA B.* 4(1):1059–1060.
- Park J, Kim Y, Xi H, Jang T, Park J-H. 2019. The complete chloroplast genome of *Abeliophyllum distichum* Nakai (Oleaceae), cultivar Ok Hwang 1ho: insights of cultivar specific variations of *A. distichum*. *Mitochondr DNA B.* 4(1):1640–1642.
- Park J, Kim Y, Xi H, Oh Y-j, Hahm KM, Ko J. 2019. The complete chloroplast genome of common camellia tree, *Camellia japonica* L. (Theaceae), adapted to cold environment in Korea. *Mitochondr DNA B.* 4(1):1038–1040.
- Park J, Kim Y, Xi H. 2019. The complete chloroplast genome of aniseed tree, *Illicium anisatum* L. (Schisandraceae). *Mitochondr DNA B.* 4(1):1023–1024.
- Tsaia C-F, Huangb C-L, Lind Y-L, Lee Y-C, Yang Y-C, Huang N-K. 2011. The neuroprotective effects of an extract of *Gastrodia elata*. *J Ethnopharmacol.* 8:119–125.
- Yuan Y, Jin X, Liu J, Zhao X, Zhou J, Wang X, Wang D, Lai C, Xu W, Huang J, et al. 2018. The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. *Nat Commun.* 9(1):1615.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18(5):821–829.
- Zhao Q-Y, Wang Y, Kong Y-M, Luo D, Li X, Hao P. 2011. Optimizing *de novo* transcriptome assembly from short-read RNA-Seq data: a comparative study. *BMC Bioinformatics.* 12(Suppl 14):S2.