Korean J. Pl. Taxon.
46(1): 60-64 (2016)
http://dx.doi.org/10.11110/kjpt.2016.46.1.60

pISSN 1225-8318
eISSN 2466-1546
Korean Journal of
Plant Taxonomy

The complete plastid genome of *Scopolia parviflora* (Dunn.) Nakai (Solanaceae)

Jin Hee Park[1] and Jungho Lee[2]*

[1]*Freshwater Bioresources Research Division, Nakdonggang National Institute of Biological Resources, Sangju 742-350, Korea*
[2]*Green Plant Institute, B-301, Heungdeok IT Valley, Giheung-gu, Yongin 446-908, Korea*
(Received 7 March 2016; Revised 17 March 2016; Accepted 18 March 2016)

ABSTRACT: *Scopolia parviflora* of the family Solanaceae is an endemic species of Korea and a traditional Korean medicinal plant. The plastid genome was sequenced by next-generation sequencing (NGS) method. The characterized cp genome is 156,193 bp in size; the large single-copy (LSC) region is 86,364 bp, the inverted repeat (IR) is 25,905 bp, and the small single copy (SSC) region is 18,019 bp. The overall GC content of the plastid genome amounts to 37.61%. The cp genome contains 113 genes and 21 introns, including 80 protein-coding genes, four RNA genes, 30 tRNA genes, 20 group II introns, and one group I intron. A phylogenetic analysis showed that *Scopolia parviflora* was closely related to *Hyoscyamus niger*.

Keywords: Chloroplast, *Scopolia parviflora*, genome sequence, medicinal plant

Michikwangipul, *Scopolia parviflora* (Dunn.) Nakai (Dunn, 1912; Nakai, 1933) of the family Solanaceae, limitedly occurs in Korean peninsula, and is similar to *S. japonica* (Maximowicz, 1873) occurring in Japan. *S. parviflora* is distinguished from *S. japonica* by ITS and some phenetic characters (Hong and Paik, 2001; Kim et al., 2003). The members of the genus *Scopolia* are known as medicinal plants (Mino, 2002; Jung et al., 2003; Min et al., 2007). The genetic makeup of the plastid genome of *Scopolia* is poorly known. Here we sequenced the chloroplast genome of *Scopolia parviflora* as a representative of the genus *Scopolia*.

Materials and Methods

The plant material of Michikwangipul used in this study was collected from the wild population of Mt. Cheonma, Korea (N37° 40' 34.85" E127° 15' 33.00"). The voucher of the plant specimen (Parkjh 20150505-141) was deposited in NNIBR, Herbarium in Nakdonggang National Institute of Biological Resources. The total DNA was prepared as described by Lee et al. (2015). The Illumina paired-end (PE) genomic library of 200 bp was constructed and sequenced using an Illumina HiSeq 2000 platform. The plastid sequence was obtained using CLC Genomics Workbench version 7.05 as described by Jeong et al. (2014). Circular structures of each replicon were confirmed by polymerase chain reaction (PCR) amplification at their ends and by joining of Sanger sequence reads derived from the amplicons. The assemblies were further verified by examining paired-end distance and depth after re-mapping reads on the contig sequences. The BLAST searches of a large contig were verified to be plastid genomes. For gene annotation of organelle genomes, protein-coding and ribosomal RNA genes were annotated using DOGMA (http://dogma.ccbb.utexas.edu/; Wyman et al., 2004). The boundaries of each annotated gene were manually determined by comparison with orthologous genes from other known cp genomes. Genes encoding tRNAs were first predicted using tRNAscan (http://lowelab.ucsc.edu/tRNA scan-SE; Lowe and Eddy, 1997) and ARAGORN version 1.2 (http://130.235.46.10/ARAGORN/; Laslett and Canback, 2004), and were manually verified by predicting the tRNA secondary structure. Circular genome maps were drawn using GenomeVx (Conant and Wolfe, 2008) followed by manual modification. The sequencing data and gene annotations were submitted to GenBank with accession number

*Author for correspondence: jlee@greenplant.re.kr

KU900232. DNA sequences of seven cp protein genes, including *psaA*, *psaB*, *psbA*, *psbB*, *psbC*, *psbD*, and *rbcL* were used to construct cp phylogenetic tree by Maximum Parsimony criterion using Paup ver. 6.0. Bootstrap and jackknife analyses of the MP tree were also performed with 1,000 replicates.

Results and Discussion

The cp-genome of *Scopolia parviflora*, was determined (Fig. 1) and found to be 156,193 bp in length. It includes small and large single copy (SSC, LSC) regions of 18,019 bp and 86,364 bp, respectively, separated by a pair of 25,905 bp inverted repeats (IRs). A total of 113 genes were detected, including 80 protein coding genes, 30 tRNA genes, and four rRNA genes (Table 1). This cp-genome was also found to contain 21 different introns, including 20 group II introns and a group I intron with a cyanobacterial origin (Besendahl et al., 2000) found within the *trnL_uaa* gene. Three protein coding genes, including *clpP*, *rps12*, and *ycf3*, contain two group II introns (*clpP.i1, clpP.i2, rps12.i1, rps12.i2, ycf3.i1* and *ycf3.i2*), and 14 genes contain a

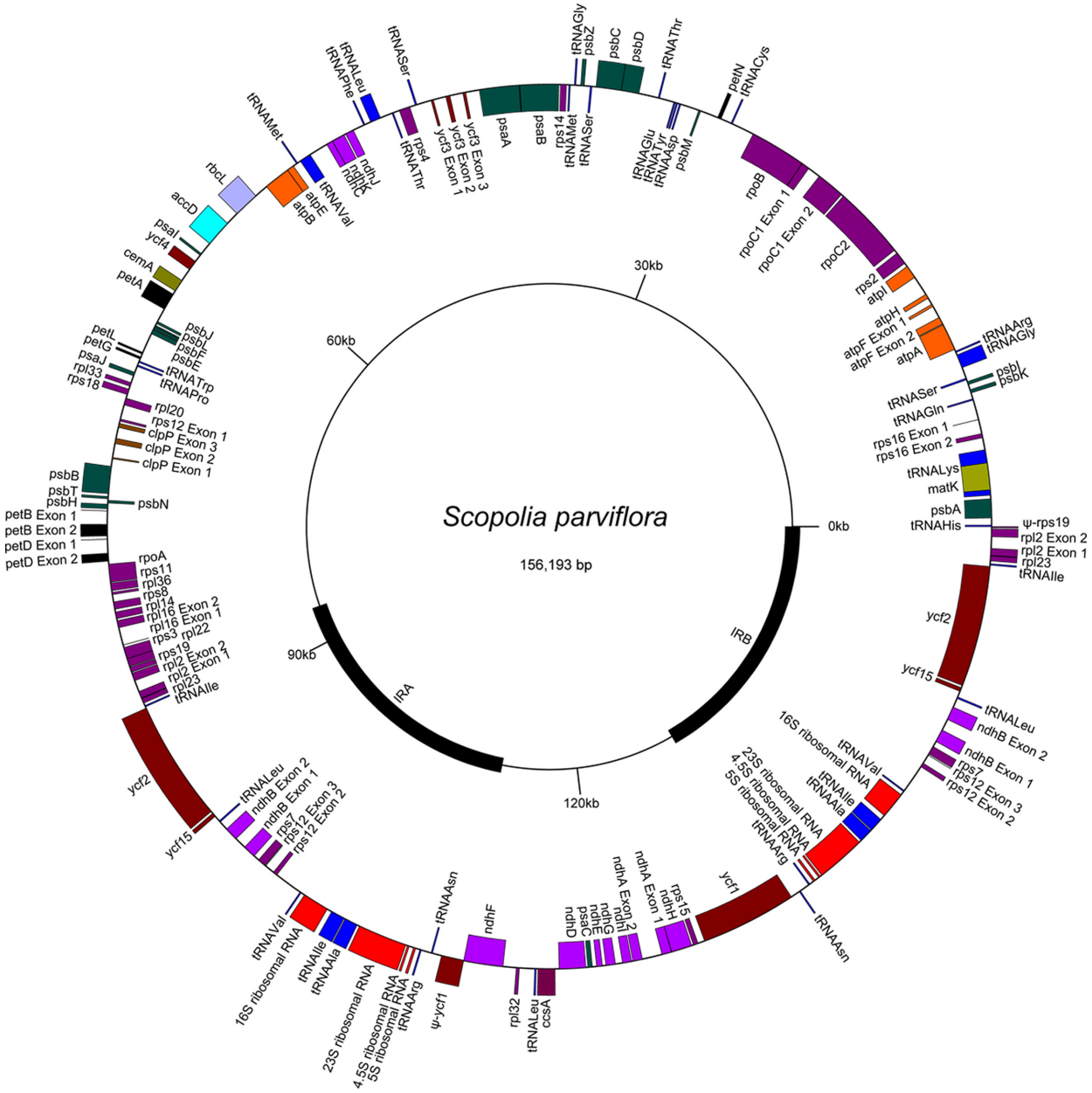

Fig. 1. Plastid genome map of *Scopolia parviflora*.

Table 1. Gene list of *Scopolia parviflora* plastid.

	Gene								
Genetic system genes									
Conserved off	*ycf1*	*ycf2* ×2	*ycf3***	*ycf4*	*ycf15*				
Maturase K	*matK*								
RNA polymerase	*rpoA*	*rpoB*	*rpoC1**	*rpoC2*					
Ribosomal protein									
Large subunits	*rpl2**×2	*rpl14*	*rpl16**	*rpl20*	*rpl22*	*rpl23* ×2	*rpl32*	*rpl33*	*rpl36*
Small subunits	*rps2*	*rps3*	*rps4*	*rps7* ×2	*rps8*	*rps11*	*rps12***α ×2	*rps14*	*rps15*
	*rps16**	*rps18*	*rps19*						
Photosynthesis genes									
Acetyl-CoA carboxylase	*accD*								
ATP-dependent Clp protease	*clpP***								
ATP synthase	*atpA*	*atpB*	*atpE*	*atpF**	*atpH*	*atpI*			
Cytochrome b	*petB**								
Cytochrome b/f	*petD**	*petG*	*petL*	*petN*					
Cytochrome f	*petA*								
Cytochrome C biogenesis	*ccsA*								
Membrane protein	*cemA*								
NADH dehydrogenase	*ndhA**	*ndhB** ×2	*ndhC*	*ndhD*	*ndhE*	*ndhF*	*ndhG*	*ndhH*	*ndhI*
	ndhJ	*ndhK*							
Photosystem I	*psaA*	*psaB*	*psaJ*	*psaC*	*psaI*				
Photosystem II	*psbA*	*psbB*	*psbC*	*psbD*	*psbE*	*psbF*			
	psbH	*psbI*	*psbJ*	*psbK*	*psbL*	*psbM*	*psbN*	*psbT*	*psbZ*
Rubisco	*rbcL*								
Ribosomal RNA	*rrn16S* ×2	*rrn23S* ×2	*rrn4.5S* ×2	*rrn5S* ×2					
Transfer RNA	*trnA_UGC** ×2	*trnC_GCA*	*trnD_GUC*	*trnE_UUC*	*trnF_GAA*	*trnfM_CAU*	*trnG_GCC*	*trnG_UCC**	*trnH_GUG*
	trnI_CAU ×2	*trnI_GAU** ×2	*trnK_UUU**	*trnL_CAA* ×2	*trnL_UAA**	*trnL_UAG*	*trnM_CAU*	*trnN_GUU* ×2	*trnP_UGG*
	trnQ_UUG	*trnR_ACG* ×2	*trnR_UCU*	*trnS_GCU*	*trnS_GGA*	*trnS_UGA*	*trnT_GGU*	*trnT_UGU*	*trnV_GAC* ×2
	*trnV_UAC**	*trnW_CCA*	*trnY_GUA*						
Pseudo gene	*ψ-rps19*	*ψ-ycf1*							

*intron $^{\alpha}$cp genome contains a copy of *rps12* exon 1 in LSC and two copies of *rps12* exon 2 and 3 in IR

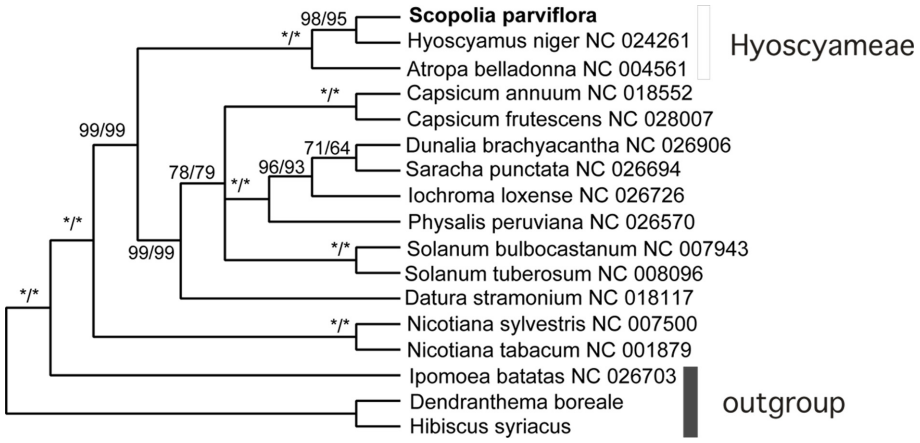

Fig. 2. Maximum parsimonious tree of 14 Solanaceae plastids, using seven protein coding gene (*psaA, psaB, psbA, psbB, psbC, psbD,* and *rbcL*) sequences. */*: Bootstrap value 100%/Jacknife value 100%.

single group II intron: *rpoC1.i, rpl2.i, rpl16.i, rps16.i, atpF.i, petB.i, petD.i, ndhA.i, ndhB.i, trnA_ugc.i, trnG_ucc.i, trnI_gau.i, trnK_uuu.i*, and *trnV_uac.i*. Among the 20 group II introns, the intron in *rps12*, between exons 1 and 2, is trans-splicing, while the other 19 group II introns are cis-splicing.

Seventeen genes, five introns, and parts of three genes and an intron are found within the IR, which has two copies. These 17 genes include six protein-coding genes (*ndhB, rpl2, rpl23, rps7, ycf2, ycf15*), all four rRNA genes (*16S, 23S, 4.5S, 5S*), and seven tRNA genes (*trnA_ugc, trnI_cau, trnI_gau, trnL_caa, trnN_guu, trnR_acg, trnV_gac*). The five introns are *ndhB.i, rpl2.i, trnA_ugc.i, trnI_gau.i*, and *rps12.i2*. The IR contains the 5' end of *ycf1* at the border with the SSC, resulting in one intact *ycf1* and a 1,473-bp *ψ-ycf1* in the cp-genome. The IR also contains the 5' end of *rps19* at the border with the LSC, resulting in one intact *rps19* and a 84-bp *ψ-rps19* in the cp-genome. In addition, the IR contains parts of the *rps12* gene. This *rps12* gene consists of three exons, *rps12.e1*, *rps12.e2*, and *rps12.e3*, *rps12.e1* is in the LSC, but *rps12.e2* and *rps12.e3* are in the IR. Thus, the genome contains a single copy of *rps12.e1* but has two copies of *rps12.e2* and *rps12.e3*. A cis-splicing group II intron, *rps12.i2*, intervenes between *rps12.e2* and *rps12.e3*, but a trans-splicing intron, *rps12.i1t*, occurs between *rps12.e1* and *rps12.e2*. The *rps12.i1t* is split into two pieces, *rps12.i1t1* and *rps12.i1t2*, because the *rps12* gene is transcribed in two separate operons, the *clpP* operon (*clpP- rps12.e1- rps12.i1t1-rpl20*) and the *3' rps12* operon *(rps12.i1t2-rps12.e2-rps12.i2-rps12.e3-rps7-ndhB*).

Currently, more than 20 plastid genomes have been deposited in Genbank from 10 genera of Solanaceae. Phylogenetic analysis showed that *Scopolia parviflora* formed a strong clade with *Hyoscyamus niger* (Sanchez-Puerta & Abbona, 2014) and *Atropha belladonna* (Schmitz-Linneweber et al., 2002), and that *Hyoscyamus niger* was the closest to *Scopolia parviflora*. The results support monophyly of the tribe Hyoscyameae of Solanaceae. The three plants are known to be highly toxic and are also used as medicine. In contrast, another toxic plant, *Datura* (Yang et al., 2014), claded with potato (Fig. 2).

In Korea, two species of the genus *Scopolia* have been documented. One is purple flowered *Scopolia parviflora* (Dunn.) Nakai (1933) and the other is yellow flowered *S. lutescens* Y. Lee (1993). In contrast, purple flowered *Scopolia japonica* Maxim. (1873) and yellow flowered *Scopolia japonica* Maxim. f. *lutescens* Sugim. (1977) also occur in Japan. Currently, ITS sequence analysis suggested that *Scopolia parviflora* and *S. japonica* were clearly distinguished, but that *Scopolia parviflora* and *S. lutescens* Y. Lee were indistinguishable (Kim et al., 2003). Plastid DNA sequence of *Scopolia japonica* from Japanese collection (voucher Tsugaru & Sawada, 17731) was only available in *ndhF* (Genbank EU580945). Without 162 ambiguous sequence of EU580945, *Scopolia japonica* had 5 SNPs in *ndhF* distinguished from *Scopolia parviflora*. Further study of comparative plastid genomics would help our understanding on the relationship among the *Scopolia* species.

Acknowledgments

This work was supported by a grant from the National Institute of Biological Resources under the Ministry of Environment, Republic of Korea.

Literature Cited

Besendahl, A., Y. L. Qiu, Lee, J. D. Palmer and D. Bhattacharya. 2000. The cyanobacterial origin and vertical transmission of

the plastid tRNA(Leu) group-I intron. Current Genetics 37: 12–23.

Conant, G. C. and K. H. Wolfe. 2008. GenomeVx: simple web-based creation of editable circular chromosome maps. Bioinformatics 24: 861–862.

Dunn, S. T. 1912. Some additions to Korean flora. Bulletin of Miscellaneous Information, Royal Gardens, Kew 1912: 108–109.

Hong, S.-P. and J.-H. Paik. 2001. Leaf epidermal microstructure of the genus *Scopolia* Jacq. s.l. (Solanaceae-Hyoscymeae) and its systematic significance. Korean Journal of Plant Taxonomy 31: 267–282. (in Korean)

Jeong, H., J. M. Lim, J. Park, Y. M. Sim, H. G. Choi, J. Lee and W. J. Jeong. 2014. Plastid and mitochondrion genomic sequences from Arctic *Chlorella* sp. ArM0029B. BMC Genomics 15: 286. doi: 10.1186/1471-2164-15-286

Jung, H. Y., S. M. Kang, Y. M. Kang, M. J. Kang, D. J. Yun, J. D. Bahk, J. K. Yang and M. S. Choi. 2003. Enhanced production of scopolamine by bacterial elicitors in adventitious hairy root cultures of *Scopolia parviflora*. Enzyme and Microbial Technology 33: 987–90.

Kim, Y.-D., J.-H. Paik, S.-H. Kim and S.-P. Hong. 2003. Phylogeny of *Scopolia* Jaq. s. str. based on ITS sequences. Korean Journal of Plant Taxonomy 33: 373–386. (in Korean)

Laslett, D. and B. Canback. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Research 32: 11–16. doi: 10.1093/nar/gkh152

Lee, M., J. Park, H. Lee, S.-H. Sohn and J. Lee. 2015. Complete chloroplast genomic sequence of *Citrus platymamma* determined by combined analysis of Sanger and NGS data. Horticulture and Environmental Biotechnology 56: 704–711.

Lee, Y. 1993. New taxa on Korean flora (5). Korean Journal of Plant Taxonomy 23: 263–268 (in Korean).

Lowe, T. M. and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Research 25: 955–964.

Maximowicz, C. J. 1873. Diagnoses plantarum novarum Japoniae et Mandshuriae. Bulletin de l'Acadeìmie Impeìriale des Sciences de St. Peitersbourg 18: 35–72.

Min, J. Y., H. Y. Jung, S. M. Kang, Y. D. Kim, Y. M. Kang, D. J. Park, D. T. Prasad and M. S. Choi. 2007. Production of tropane alkaloids by small-scale bubble column bioreactor cultures of *Scopolia parviflora* adventitious roots. Bioresource Technology 98: 1748–1753.

Mino, Y. 2002. Amino acid sequences of ferredoxins from *Scopolia japonica* and *Lycium chinense*: their similarities to that of *Datura arborea*. Biological Pharmceutical Bulletin 25: 1367–1369.

Nakai, T. 1933. Notulae ad Plantas Japoniae & Koreae XLIII. Botanical Magazine 47: 235–267.

Sanchez-Puerta, M. V. and C. C. Abbona. 2014. The chloroplast genome of *Hyoscyamus niger* and a phylogenetic study of the tribe Hyoscyameae (Solanaceae). PLoS One 9: e98353. doi:10.1371/journal.pone.0098353.

Schmitz-Linneweber, C., R. Regel, T. G. Du, H. Hupfer, R. G. Herrmann and R. M. Maier. 2002. The plastid chromosome of *Atropa belladonna* and its comparison with that of *Nicotiana tabacum*: the role of RNA editing in generating divergence in the process of plant speciation. Molecular Biology and Evolution 19: 1602–1612.

Sugimoto, J. 1977. Notes on flora of Japan (3). Journal of Geobotany 26: 59–64. (in Japanese)

Wyman, S. K., R. K. Jansen and J. L. Boore. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20: 3252–3255. doi: 10.1093/bioinformatics/bth352

Yang, Y., Y. Dang, Q. Li, J. Lu, X. Li and Y. Wang. 2014. Complete chloroplast genome sequence of poisonous and medicinal plant *Datura stramonium*: organizations and implications for genetic engineering. PLoS One 9: e110656. doi: 10.1371/journal.pone.0110656