



The complete chloroplast genome sequence of *Solanum hougasii*, one of the potato wild relative species

Kwang-Soo Cho^a , Ji-Hong Cho^a, Ju-Sung Im^a, Jang-Gyu Choi^a, Young-Eun Park^a, Dong-Chil Jang^a, Su-Young Hong^a and Tae-Ho Park^b 

^aHighland Agriculture Research Institute, National Institute of Crop Science, Rural Development Administration, Pyeongchang, South Korea;

^bDepartment of Horticulture, Daegu University, Gyeongsan, South Korea

ABSTRACT

Solanum hougasii is a wild tuber-bearing species belonging to the family Solanaceae. The complete chloroplast genome of *S. hougasii* was constituted by *de novo* assembly, using a small amount of whole genome sequencing data. The chloroplast genome of *S. hougasii* was a circular DNA molecule with a length of 155,549 bp and consisted of 85,990 bp of large single copy, 18,373 bp of small single copy, and 25,593 bp of a pair of inverted repeat regions. A total of 158 genes were annotated, including 105 protein-coding genes, 45 tRNA genes, and eight rRNA genes. Maximum likelihood phylogenetic analysis with 25 Solanaceae species revealed that *S. hougasii* is most closely grouped with *S. tuberosum*.

ARTICLE HISTORY

Received 26 May 2018

Accepted 18 June 2018

KEYWORDS

Chloroplast; genome; genome sequence; *Solanum hougasii*

Solanum hougasii, a wild tuber-bearing hexaploid species, is a relative to the cultivated potato, *S. tuberosum*. It was identified to be a source of resistance to late and early blight, root-knot nematode, and potato virus Y for potato breeding (Cockerham 1970; Brown et al. 1999; Inglis et al. 2007; Haynes and Qu 2016). Its EBN (Endosperm Balanced Number) value of four, theoretically makes it directly crossable for breeding purposes with cultivated tetraploid potatoes (Hawkes 1990; Ortiz and Ehlenfeldt 1992; Cho et al. 1997; Spooner et al. 2014; Haynes and Qu 2016). Moreover, its nuclear genome composition has evolutionally been identified by GISH analysis (Pendinen et al. 2012). *S. hougasii* has an allotropic behavior, that is, one genome belonged to AA and the other to BB. In addition, *S. hougasii* third genome is more intimately related to P genome or to the species related to P genome (Pendinen et al. 2012). The information of plastid genome of the wild species obtained in this study will provide an opportunity to investigate more detailed evolutionary and breeding aspects.

The *S. hougasii* (PI161174) was originally collected in Michoacan, Mexico by International Potato Centre (CIP), provided via Highland Agriculture Research Institute and stored at Daegu University, South Korea. A paired-end (PE) genomic library was constructed with total genomic DNA, according to the standard protocol (Illumina, San Diego, USA) and sequenced using an HiSeq2000 at Macrogen (<http://www.macrogen.com/kor/>). Low-quality bases with raw scores of 20 or less were removed and approximately 5.1 Gbp of high-quality PE reads were assembled by a CLC genome assembler

(CLC Inc, Aarhus, Denmark) (Kim et al. 2015). The reference chloroplast genome sequence of *S. commersonii* (KM489054, Cho et al. 2016) was used to retrieve principal contigs representing the chloroplast genome from the total contigs using Nucmer (Kurtz et al. 2004). The representative chloroplast contigs were arranged in an order based on BLASTZ analysis (Schwartz et al. 2003) with the reference sequence and were connected to a single draft sequence by joining overlapping terminal sequences. DOGMA (Wyman et al. 2004) and BLAST searches were used to predict the chloroplast genes.

The complete chloroplast genome of *S. hougasii* (GenBank accession no. MF471372) was 155,549 bp in length and included 25,593 bp inverted repeat (IRa and IRb) regions separated by small single copy (SSC) region of 18,373 bp and large single copy (LSC) region of 85,990 bp with the typical quadripartite structure of most plastids, and the structure and gene features were typically identical to those of higher plants. A total of 158 genes with an average size of 584.5 bp were annotated including 105 protein-coding genes with an average size of 766.6 bp, 45 tRNA genes, and 8 rRNA genes. An overall GC content was found to be 37.87%.

Phylogenetic analysis was performed using chloroplast coding sequences of *S. hougasii* and 25 published species in Solanaceae family by a maximum likelihood method in MEGA 6.0 (Tamura et al. 2013). According to the phylogenetic tree, *S. hougasii* belonged to the same clade in *Solanum* species as expected, and was interestingly closely grouped with *S. tuberosum* (Figure 1).

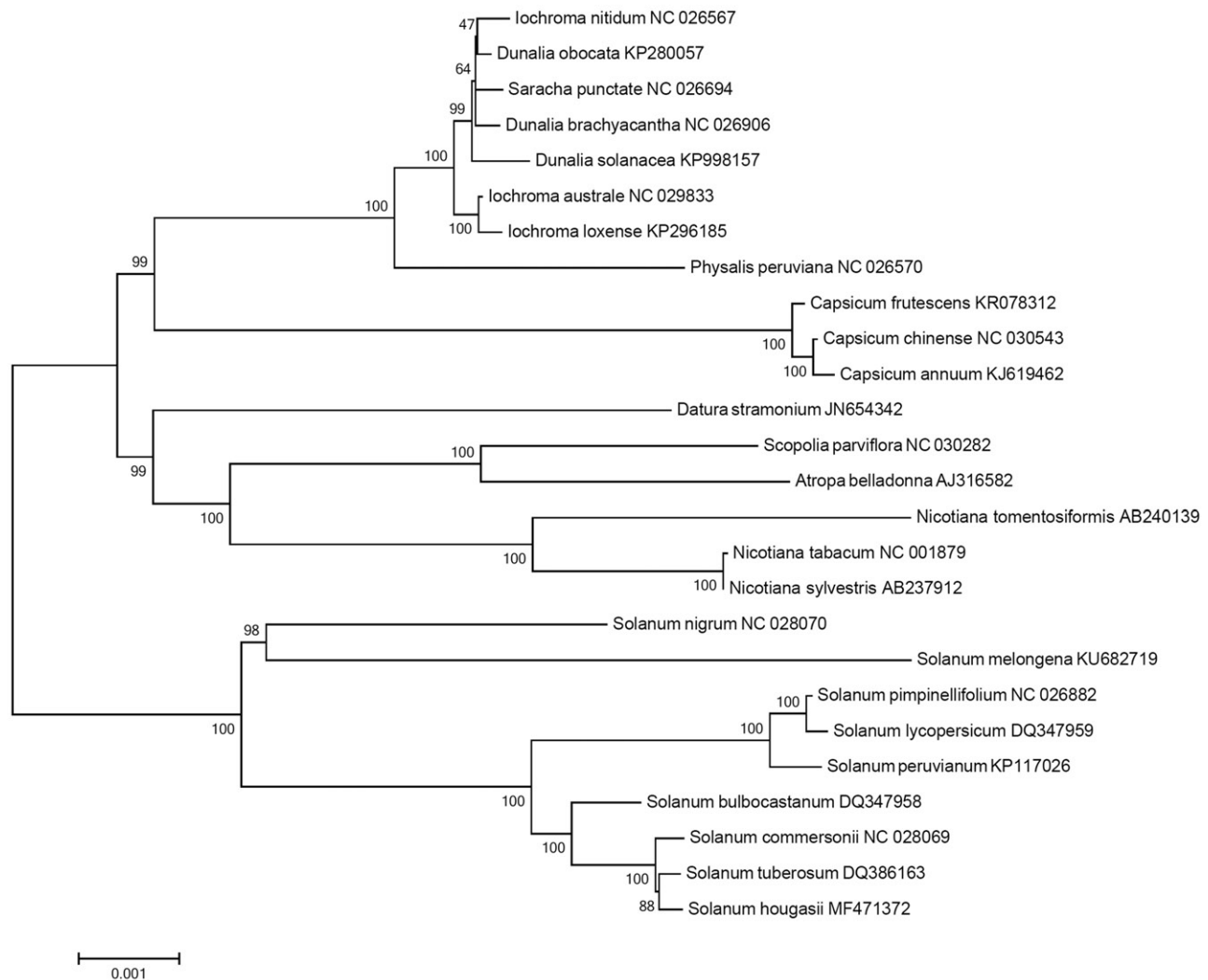


Figure 1. Maximum likelihood phylogenetic tree of *S. hougasii* with 25 species belonging to the Solanaceae based on chloroplast protein coding sequences. Numbers in the nodes are the bootstrap values from 1,000 replicates.

Disclosure statement

The author reports no conflicts of interest. The author alone is responsible for the content and writing of this article.

Funding

This study was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (IPET) through Golden Seed Project [Project No. 213009-05-2-WT411] funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA) and Korea Forest Service (KFS), South Korea.

ORCID

Kwang-Soo Cho  <http://orcid.org/0000-0002-0590-3650>
Tae-Ho Park  <https://orcid.org/0000-0002-4316-5732>

References

Brown CR, Hojtahedi H, Santo GS. 1999. Genetic analysis of resistance to *Meloidogyne chitwoodi* introgressed from *Solanum hougasii* into cultivated potato. *J Nematol.* 31:264–271.

Cho HM, Kim-Lee HY, Om YH, Kim JK. 1997. Influence of endosperm balance number (EBN) in interploidal and interspecific crosses between *Solanum tuberosum* dihaploids and wild species. *Korean J Breed.* 29:154–161.

Cho K-S, Cheon K-S, Hong S-Y, Cho J-H, Im J-S, Mekapogu M, Yu Y-S, Park T-H. 2016. Complete chloroplast genome sequences of *Solanum commersonii* and its application to chloroplast genotype in somatic hybrids with *Solanum tuberosum*. *Plant Cell Rep.* 35:2113–2123.

Cockerham G. 1970. Genetical studies on resistance to potato viruses X and Y. *Heredity.* 25:309–348.

Hawkes JG. 1990. The potato: Evolution, biodiversity and genetic resources. Belhaven Press, London, UK.

Haynes KG, Qu X. 2016. Late blight and early blight resistance from *Solanum hougasii* introgressed into *Solanum tuberosum*. *Amer J of Potato Res.* 93:86–95.

Inglis DA, Brown CR, Gundersen BG, Porter LD, Miller JS, Johnson DA, Lozoya-Saldaña H, Haynes KG. 2007. Assessment of *Solanum Hougasii* in Washington and Mexico as a source of resistance to late blight. *Amer J of Potato Res.* 84:217–228.

Kim K, Lee SC, Lee J, Lee H, Joh HJ, Kim NH, Park HS, Yang TJ. 2015. Comprehensive survey of genetic diversity in chloroplast genomes and 45S rDNAs within *Panax ginseng* species. *PLoS One.* 10:e0117159

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12

- Ortiz R, Ehlenfeldt MK. 1992. The importance of endosperm balance number in potato breeding and the evolution of tuber-bearing *Solanum* species. *Euphytica*. 60:105–113.
- Pendinen G, Spooner DM, Jiang J, Gavrilenko T. 2012. Genomic *in situ* hybridization reveals both auto- and allopolyploid origins of different North and Central American hexaploid potato (*Solanum* sect. *Petota*) species. *Genome*. 55:407–415.
- Schwartz S, Kent WJ, Smit A, Zhang Z, Baertsch R, Hardison RC, Haussler D, Miller W. 2003. Human-mouse alignments with BLASTZ. *Genome Res*. 13:103–107.
- Spooner DM, Ghislain M, Simon R, Jansky SH, Gavrilenko T. 2014. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *Bot Rev*. 80:283–383.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 30:2725–2729.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 20:3252–3255.