


## The complete mitochondrial genome sequences of potato (*Solanum tuberosum* L., Solanaceae)

Kwang-Soo Cho<sup>a</sup> , Ji-Hong Cho<sup>a</sup>, Ju-Sung Im<sup>a</sup>, Jang-Gyu Choi<sup>a</sup>, Young-Eun Park<sup>a</sup>, Su-Young Hong<sup>a</sup>, Min Kwon<sup>a</sup>, Jin-Ho Kang<sup>b</sup> and Tae-Ho Park<sup>c</sup>

<sup>a</sup>Highland Agriculture Research Institute, National Institute of Crop Science, Rural Development Administration, Pyeongchang, South Korea;

<sup>b</sup>Graduate School of International Agricultural Technology and Crop Biotechnology Institute, Seoul National University, Pyeongchang, South Korea; <sup>c</sup>Department of Horticulture, Daegu University, Gyeongsan, South Korea

### ABSTRACT

Potato (*Solanum tuberosum*) from the Solanaceae is the fourth most important food crop worldwide. In this study, five complete mitochondrial genome sequences of *S. tuberosum* were characterized through *de novo* assembly of whole genome sequencing data. The resulting circular mitochondrial DNA molecules ranged from 49,171 bp to 297,014 bp in size and contained a total of 80 non-redundant genes, comprising 34 protein-coding genes, 24 hypothetical open reading frames, 19 tRNA genes, and 3 rRNA genes. Phylogenetic analysis using common protein-coding sequences confirmed that *S. tuberosum* belongs to the Solanoideae subfamily in the Solanaceae family.

### ARTICLE HISTORY

Received 10 October 2017  
Accepted 26 October 2017

### KEYWORDS

Potato; *Solanum tuberosum*;  
Mitochondrial genome;  
Solanaceae

Mitochondrial (mt) genomes of higher plants are more complex and encode a higher number of genes compared to those of other organisms. Their distinct features are the frequent uptake of foreign DNA by gene transfer (Richardson and Palmer 2007; Goremykin et al. 2009), very low mutation rate (Wolfe et al. 1987), and dynamic structure (Palmer and Herbon 1988) with an unusual size variation (Alverson et al. 2010). Furthermore, they play important roles in plant development and productivity (Ogihara et al. 2005; Li et al. 2009). Nevertheless, complete sequences of mt genomes from higher plants are rarely acquired, and none has been reported for *Solanum* species since the mt genome of *Nicotiana tabacum*, another member of the Solanaceae family, was reported (Sugiyama et al. 2005). In this study, we report the complete mt genome sequences of *Solanum tuberosum* L. (potato), one of the most important crops worldwide.

Total genomic DNA was extracted from fresh leaves of *S. tuberosum* (breeding line no. PT56 maintained in vitro plant in HARI) grown in Highland Agriculture Research Institute (HARI), National Institute of Crop Science, Rural Development Administration, Pyeongchang, South Korea, and used to construct a paired-end (PE) library with ~670 bp insert size according to standard Illumina PE library protocol. The library was sequenced using an Illumina MiSeq platform at LabGenomics (Seongnam, Korea). High quality PE reads of about 4.1 Gb were employed for *de novo* assembly using CLC genome assembler (v. beta 4.6; Qiagen, Aarhus, Denmark) as described previously (Cho et al. 2015). Contigs with

mitochondrial genome sequences were selected from initial assembled contigs, of which the longest mitochondrial contig was set as the initiation seed contig. The contigs were extended and joined by a series of PE read mappings and gap-fillings to generate final complete mitochondrial genome sequences. The assembled sequences were manually inspected based on read mapping status and BLASTN searches. The genome sequence was annotated using GeSeq (Tillich et al. 2017) and manual curation based on BLAST searches.

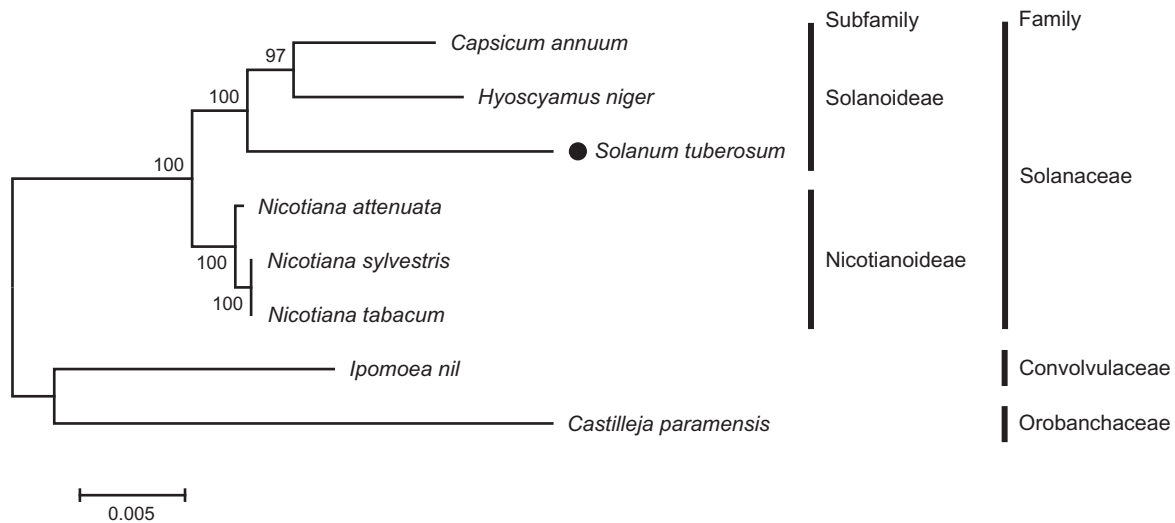
Five mitochondrial genome sequences of *S. tuberosum* were completed. The circular DNA molecules were 49,171 bp, 49,230 bp, 112,800 bp, 247,843 bp, and 297,014 bp long (GenBank accession nos. MF989953–MF989957). The largest mitochondrial genome sequence (297,014 bp) included sequences identical to the 247,843-bp and 49,171-bp mitochondrial genomes. A total of 80 non-redundant genes were annotated in the five genomes, comprising 34 protein-coding gene, 24 hypothetical open reading frames, 19 tRNA genes, and three rRNA genes. Among protein-coding genes, exons of trans-splicing *nad2* were dispersed on different genomes. Exons 1 and 2 were present only on the 49,230-bp genome, while Exons 3, 4, and 5 were found in two copies on 297,014-bp and 247,843-bp genomes.

Phylogenetic analysis based on multiple alignments of common 23 protein-coding sequences in the mt genomes confirmed that *S. tuberosum* belongs to the Solanoideae subfamily in the Solanaceae family (Figure 1).

**CONTACT** Tae-Ho Park  thzoo@daegu.ac.kr  Department of Horticulture, Daegu University, Gyeongsan, South Korea

© 2017 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 tree of *Solanum tuberosum* and related taxa based on mitochondrial genome sequences. Common protein-coding sequences in mitochondrial genome sequences were aligned using MAFFT (<http://mafft.cbrc.jp/alignment/server/index.html>) and used to generate maximum-likelihood phylogenetic tree by MEGA 6.0 (Tamura et al. 2013). Numbers on the nodes indicate bootstrap support values (>50%) from 1000 replicates. Scale bar represents the number of nucleotide substitution per site. Mitochondrial genome sequences used for this tree are *Capsicum annuum*, NC\_024624; *Castilleja paramensis*, NC\_031806; *Hyoscyamus niger*, NC\_026515; *Ipomoea nil*, NC\_031158; *Nicotiana attenuata*, MF579563; *N. sylvestris*, NC\_029805; *N. tabacum*, NC\_006581; *Solanum tuberosum*, MF989953–MF989957.

## Disclosure statement

The authors report no conflicts of interest.

## Funding

This study was supported by the Cooperative Research Program for Agriculture Science & Technology Development, Rural Development Administration, Republic of Korea; Project no. PJ01184101.

## ORCID

Kwang-Soo Cho  <http://orcid.org/0000-0002-0590-3650>

## References

- Alverson AJ, Wei X, Rice DW, Stem DB, Barry K, Palmer JD. 2010. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucubita pepo* (Cucurbitaceae). *Mol Biol Evol.* 27:1436–1448.
- Cho K-S, Yun B-K, Yoon Y-H, Hong S-Y, Mekapogu M, Kim K-H, Yang T-J. 2015. Complete chloroplast genome sequence of tartary buckwheat (*Fagopyrum tataricum*) and comparative analysis with common buckwheat (*F. esculentum*). *PLoS One.* 10:e0125332.
- Goremykin VV, Salamini F, Velasco R, Viola R. 2009. Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. *Mol Biol Evol.* 26:99–110.
- Li L, Wang B, Liu Y, Qiu YL. 2009. The complete mitochondrial genome sequence of the hornwort *Megaceros aenigmaticus* shows a mixed mode of conservative yet dynamic evolution in early land plant mitochondrial genomes. *J Mol Evol.* 68: 665–678.
- Ogihara Y, Yamazaki Y, Murai K, Kanno A, Terachi T, Shiina T, Miyashita N, Nasuda S, Nakamura C, Mori N, et al. 2005. Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acids Res.* 33:6235–6250.
- Palmer J, Herbon L. 1988. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *J Mol Evol.* 28:87–97.
- Richardson AO, Palmer JD. 2007. Horizontal gene transfer in plants. *J Exp Bot.* 58:1–9.
- Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, Sugiura M. 2005. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Mol Genet Genomics.* 272: 603–615.
- Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 30:2725–2729.
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research.* 45: W6–W11.
- Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA.* 84:9054–9058.