

## The complete chloroplast genome sequence of *Capsicum chinense* Jacq. (Solanaceae)

Hyun-Seung Park<sup>a\*</sup>, Jundae Lee<sup>b\*</sup>, Sang-Choon Lee<sup>a</sup>, Tae-Jin Yang<sup>a</sup> and Jae Bok Yoon<sup>c</sup>

<sup>a</sup>Department of Plant Science, Plant Genomics and Breeding Institute, and Research Institute for Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea; <sup>b</sup>Department of Horticulture, Chonbuk National University, Jeonju, Republic of Korea; <sup>c</sup>Research and Development Unit, Pepper and Breeding Institute, Business Incubator, College of Agricultural and Life Sciences, Seoul National University, Suwon, Republic of Korea

### ABSTRACT

*Capsicum chinense* is one of the domesticated pepper species and well known for its distinctive pungency. The complete chloroplast genome sequence of *C. chinense* was generated by *de novo* assembly using next generation sequencing data. The chloroplast genome is 156 807 bp long, containing large single-copy region of 87 290 bp and small single-copy region of 17 911 bp separated by a pair of inverted repeats of 25 803 bp. A total of 113 genes were predicted including 79 protein-coding genes, 30 tRNA genes and four rRNA genes. Phylogenomic analysis revealed that *C. chinense* chloroplast genome was most closely related to *Capsicum annuum* var. *glabrusculum* (American bird pepper), a wild progenitor of *C. annuum*.

### ARTICLE HISTORY

Received 15 January 2016  
Accepted 16 January 2016

### KEYWORDS

*Capsicum chinense*;  
chloroplast; genome; NGS





*Capsicum chinense* is one of the five domesticated pepper species belonging to the Solanaceae family. Its fruits are known for distinctive pungency causing severe heat and nominated as the hottest pepper in the world with one million of the Scoville heat units (SHU) (Bosland & Baral 2007). Studies have mainly been focused on pungency-related compounds, capsaicinoids, using metabolic approach in *C. chinense* (Menichini et al. 2009; Wahyuni et al. 2013). Previously, two complete chloroplast genome sequences of *Capsicum annuum* L. were reported: the cultivated type and its wild progenitor (Jo et al. 2011; Zeng et al. 2016). However, complete chloroplast sequence of *C. chinense* has not yet been reported and only several partial chloroplast genome sequences were reported. In this study, we reported complete sequence of chloroplast genome of *C. chinense* to provide basic genetic information and help understand genomic diversity of pepper species.

The genomic DNA was extracted from leaves of *C. chinense* cultivated in the field of Pepper and Breeding Institute, Suwon, Korea. An Illumina pair-end (PE) genomic library of 300-bp insert was constructed by following the PE standard protocol (Illumina, USA) and sequenced using an Illumina HiSeq 2000 platform by THERAGEN ETEX BIO INSTITUTE (<http://www.theragenetex.com>), Suwon, Korea. High quality reads (Phred scores of 20 or more) of 0.5 Gbp were obtained and assembled using CLC genome assembler (ver 4.06 beta, CLC Inc, Aarhus, Denmark), by following the assembly method in Kim et al. (2015a,b). Contigs representing chloroplast

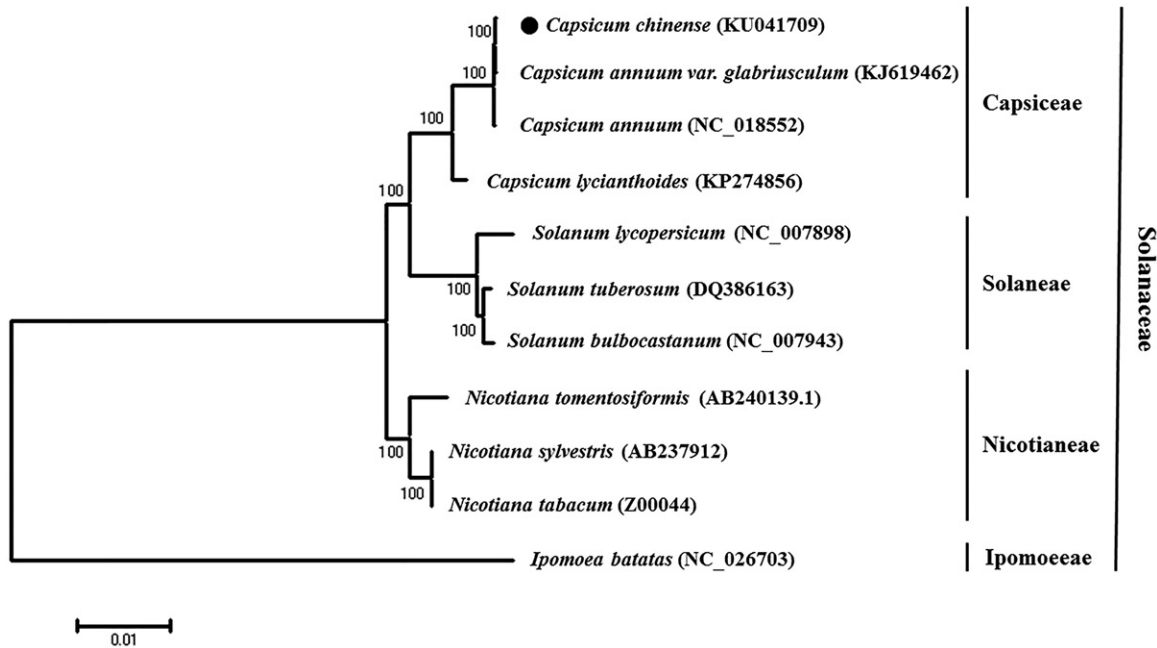
genome were retrieved, ordered and joined into a single draft sequence by comparison of *C. annuum* chloroplast genome (Jo et al. 2011) as a reference. The draft sequence was confirmed and manually corrected by PE read mapping. The draft sequence was annotated using DOGMA (Wyman et al. 2004).

The complete chloroplast sequence of *C. chinense* had a length of 156 807 bp (GenBank accession no. KU041709) and consisted of four typical chloroplast parts, such large single copy (LSC) of 87 290 bp and small single copy (SSC) of 17 911 bp, separated by a pair of inverted repeats of 25 803 bp (IRa and IRb). Total GC contents were 37.7%, which was consistent with that of reported chloroplast genome of cultivated *C. annuum* (Jo et al. 2011). A total of 113 genes were predicted, including 79 protein-coding genes, 30 tRNA genes and four rRNA genes of which 28 genes were duplicated in IR regions. The order and contents of genes in *C. chinense* chloroplast genome were identical to those of *C. annuum* chloroplast genome.

Phylogenomic analysis was performed with previously reported chloroplast genome sequences of 10 species belonging to the Solanaceae family using a neighbor-joining analysis of MEGA 6.0 (Tamura et al. 2013) with 1000 bootstrap replicates. As expected, the 10 species were classified into three groups in our phylogenetic tree according to their tribes, Capsiceae, Solaneae, and Nicotianeae (Figure 1). In the phylogenetic tree, *C. chinense* was included in Capsiceae tribe group, in which *C. chinense* was much closer to *C. annuum* var. *glabrusculum* (American bird pepper), a wild

**CONTACT** Tae-Jin Yang  [tjyang@snu.ac.kr](mailto:tjyang@snu.ac.kr)  Department of Plant Science, Plant Genomics and Breeding Institute, and Research Institute for Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Republic of Korea; Jae Bok Yoon  [yoonyb2@snu.ac.kr](mailto:yoonyb2@snu.ac.kr)  Research and Development Unit, Pepper and Breeding Institute, Business Incubator, College of Agricultural and Life Sciences, Seoul National University, Suwon 16614, Republic of Korea

\*These authors contributed equally to this work.



**Figure 1.** Phylogenetic relationship of 11 complete chloroplast genome sequences belonging to the Solanaceae family based on the neighbour-joining analysis of MEGA6.0. Numbers around nodes indicate the bootstrap value from 1000 replicates. The chloroplast sequence of *Ipomoea batatas* (sweet potato) was set as the outgroup.

progenitor of *C. annuum* (Zeng et al. 2016), than other *Capsicum* species.

### Disclosure statement

The authors report no conflict of interest. The authors alone are responsible for the content and writing of the paper. This research was supported by "Cooperative Research Program for Agriculture Science & Technology Development (Project no. PJ01114501)" Rural Development Administration, Republic of Korea.

### References

- Bosland PW, Baral JB. 2007. 'Bhut Jolokia' – the world's hottest known chile pepper is a putative naturally occurring interspecific hybrid. *Hort Sci.* 42:222–224.
- Jo YD, Park J, Kim J, Song W, Hur CG, Lee YH, Kang BC. 2011. Complete sequencing and comparative analyses of the pepper (*Capsicum annuum* L.) plastome revealed high frequency of tandem repeats and large insertion/deletions on pepper plastome. *Plant Cell Rep.* 30:217–229.
- Kim K, Lee SC, Lee J, Lee HO, Joh HJ, Kim NH, Park HS, Yang TJ. 2015a. Comprehensive survey of genetic diversity in chloroplast genomes and 45S nrDNAs within *Panax ginseng* species. *PLoS One.* 10: e0117159.
- Kim K, Lee SC, Lee J, Yu Y, Yang K, Choi BS, Koh HJ, Waminal NE, Choi HI, Kim NH, et al. 2015b. Complete chloroplast and ribosomal sequences for 30 accessions elucidate evolution of *Oryza* AA genome species. *Sci Rep.* 5:15655.
- Menichini F, Tundis R, Bonesi M, Loizzo MR, Conforti F, Statti G, De Cindio B, Houghton PJ, Menichini F. 2009. The influence of fruit ripening on the phytochemical content and biological activity of *Capsicum chinense* Jacq. cv Habanero. *Food Chem.* 114:553–560.
- 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.
- Wahyuni Y, Ballester AR., Tikunov Y, De Vos RC, Pelgrom KT, Maharajaya A, Sudarmonowati E, Bino RJ, Bovy AG. 2013. Metabolomics and molecular marker analysis to explore pepper (*Capsicum* sp.) biodiversity. *Metabolomics.* 9:130–144.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.
- Zeng FC, Gao CW, Gao LZ. 2016. The complete chloroplast genome sequence of American bird pepper (*Capsicum annuum* var. *glabriusculum*). *Mitochondrial DNA.* 27:1:724–726. DOI: 10.3109/19401736.2014.913160.