Characterization and phylogenetic relationships analysis of the complete chloroplast genome of *Capsicum annuum* (Solanaceae)

Mei Bie\(^{a,b}\), Chengdong Han\(^{b}\), Xuanyao Wang\(^{b}\), Wei Xiao\(^{b}\) and Kai Song\(^{c}\)

\(^{a}\)College of Computer Science and Technology, Jilin University, Changchun, Jilin, P. R. China; \(^{b}\)Institute of Education, Changchun Normal University, Changchun, Jilin, P. R. China; \(^{c}\)School of Life Science, Changchun Normal University, Changchun, Jilin, P. R. China

**ABSTRACT**

*Capsicum annuum* is one of the oldest domesticated crops in the Americas, which is also the most widely grown spice crop in the world. The complete chloroplast genome of *C. annuum* has been assembled and annotated in this paper. Its length was 156,781 bp, containing a large single-copy region of 87,367 bp, a small single-copy region of 17,850 bp, and a pair of IR regions of 25,782 bp in each. The whole chloroplast genome of *C. annuum* contains 135 genes, including 89 protein-coding genes (PCGs), 38 transfer RNA genes (tRNAs), and 8 ribosome RNA genes (rRNAs). The overall nucleotide composition is: A of 30.8%, T of 31.5%, C of 19.1% and G of 18.6%, with a total GC content of the chloroplast genome 37.7% and AT of 62.3%. Phylogenetic relationship analysis was based on 10 plant species using the maximum-likelihood (ML) methods, which showed that the position of *C. annuum* clustered with *C. galapagoense*.

*Capsicum annuum* is an economically important genus of the Solanaceae family (Moscone et al. 2007). It is used as a vegetable, spice and medicine in the world (Ahn et al. 2016). It has several pharmacological and physiological effects, including analgesic, anti-inflammatory, antioxidant, and anti-obesity properties (Luo et al. 2011). *Capsicum annuum* contains pungent compounds called capsaicinoids. These include capsaicin, dihydrocapsaicin, nordihydrocapsaicin, and trace amounts of other compounds. Capsaicin (trans-8-methyl-N-vanillyl-6-nonenamide) is the major pungent agent, responsible for about 70% of the burn in *C. annuum* (Tomi and Anthony 2019). In this paper, we assembled and annotated the complete chloroplast genome of *C. annuum* and discussed the phylogenetic relationship with other plant species, which provides a valuable resource for genetic resources and gene databases.

The sample of *C. annuum* was collected from the Cucumber-picking garden in Changchun district of Jilin province (Changchun, Jilin, China, 125.41E; 43.92N). The fresh plant tissue and chloroplast (cp) DNA of *C. annuum* were extracted using the modified CTAB method and stored at the Institute of Education, Changchun Normal University (No. IE-CNU-02). The cp DNA was purified and fragmented using the NEB Next Ultra\(^{TM}\) II DNA Library Prep Kit (NEB, BJ, and CN) and was sequenced. Quality and adapters control was performed and removed low-quality reads and adapters using the NGS QC Toolkit software (Patel and Jain 2012). The chloroplast genome of *C. annuum* was assembled and annotated using the MitoZ software (Meng et al. 2019).

The complete physical map of *C. annuum* chloroplast genome was generated using the OGDRAW version 1.3.1 (Greiner et al. 2019). The annotated complete cp genome of *C. annuum* sequence was deposited in the GenBank with the accession No. NC_018552.1 (Jo et al. 2011).

The complete chloroplast genome of *C. annuum* was a circular in shape with 156,781 bp length, containing a large single-copy region (LSC) of 87,367 bp, a small single-copy region (SSC) of 17,850 bp, and a pair of inverted repeat regions (IRs) of 25,782 bp in each. We annotated and found the cp of *C. sativus* comprised 135 genes, including 89 protein-coding genes (PCGs), 38 transfer RNA genes (tRNAs), and 8 ribosomal RNA genes (rRNAs). The overall nucleotide composition is: 30.8% of A, 31.5% of T, 19.1% of C, and 18.6% of G, with a total G+C content of 37.7% and A+T of 62.3%.

To confirm the phylogenetic position and phylogenetic relationship of *C. annuum* with 10 plant species, the Maximum-Likelihood (ML) method was used. The phylogenetic tree was reconstructed using RaxML version 8.0 (Stamatakis 2014) with the GTR+G+I model. Phylogenetic relationship obtained with the ML approach were identical to those obtained using the Bayesian analysis, which was analyzed by MrBayes version 3.2.5 (Ronquist and Huelsenbeck 2003) and based on the most appropriate model. The phylogenetic tree was performed using MEGA X software (Kumar et al. 2018) by 2000 bootstrap replicates and edited using iTOL version 4.0 (https://itol.embl.de/) (Letunic and Bork 2019). In the phylogenetic tree (Figure 1), the result showed that the chloroplast genome of *C. annuum* is clustered with...
Solanaceae species of *Capsicum galapagoense* (NC_033524.1) in the phylogenetic relationship.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**References**

Ahn YK, Karna S, Jun TH, Yang EY, Lee HE, Kim JH, Kim JH. 2016. Complete genome sequencing and analysis of *Capsicum annuum* varieties. Mol Breed. 36(10):140.

Greiner S, Lehwalk P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 47(W1):W59–W64.

Jo Y D, 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(2):217–229.

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.

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOl) v4: recent updates and new developments. Nucleic Acids Res. 47(W1):W256–W259.

Luo XJ, Peng J, Li YJ. 2011. Recent advances in the study on capsaicinoids and capsinoids. Eur J Pharmacol. 650(1):1–7.

Meng G, Li Y, Yang C, Liu S. 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res. 47(11):e63–e63.

Moscone EA, Scaldaferrro MA, Grabiele M, Cecchini NM, Sánchez GY, Jarret R, Daviña JR, Ducasse DA, Barboza GE, Ehrendorfer F. 2007. The evolution of chili peppers (*Capsicum* – Solanaceae): a cytogenetic perspective. Acta Hortic. 745:137–170.

Patel RK, Jain M. 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One. 7(2):e30619.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19(12):1572–1574.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313.

Tomi L O, Anthony J A. 2019. Comparison of nutritional, antioxidant vitamins and capsaicin contents in *Capsicum annuum* and *C. frutescens*. Int J Veg Sci. 1:1–18.