Complete Chloroplast Genome Sequencing and Genetic Relationship Analysis of *Capsicum chinense* Jacq

Sebastin Raveendar¹, Kyung Jun Lee¹, Myoung-Jae Shin¹, Gyu-Taek Cho¹, Jung-Ro Lee¹, Kyung-Ho Ma¹, Gi-An Lee^{1*}, Jong-Wook Chung^{2*}

¹National Agrobiodiversity Center, National Institute of Agricultural Sciences, RDA, Jeonju 54874, Korea

²Department of Industrial Plant Science and Technology, Chungbuk National University, Cheongju 28644, Korea

ABSTRACT *Capsicum chinense* is one of the five domesticated pepper species belonging to the Solanaceae family. *Capsicum* sp. have been used as model systems in comparative and evolutionary genomics because their superior availability of chloroplast genome in the solanaceae family. Similarly, molecular markers derived from the complete chloroplast genome can provide effective tools for species identification and phylogenetic resolution. So far however, only partial taxonomic and phylogenetic analyses have been carried out for the genus. Thus, the complete chloroplast genome sequence of a cultivated pepper (*C. chinense*) has been reported here. The total length of the chloroplast genome is 156,936 bp, with 37.7% overall GC content. A pair of inverted repeats (IRs) of 25,847 bp was separated by a small single copy (SSC) region of 17,912 bp and a large single copy (LSC) region of 87,330 bp. The chloroplast genome harbors 113 known genes, including 79 protein-coding genes, four ribosomal RNA genes, and 30 transfer RNA (tRNA) genes. In all, 21 of these genes are duplicated in the inverted repeat regions, 15 genes and six tRNA genes contain a single intron, while two genes have two introns. Analysis revealed 117 simple sequence repeat (SSR) loci, which are mostly located in the intergenic regions. The complete chloroplast genome reported here enriches our knowledge of the genetic complement of *C. chinense*, and contributes to our understanding of the genetic relationships within the genus *Capsicum*.

Keywords *Capsicum chinense*, Chloroplast, DNA sequencing, Hot pepper

INTRODUCTION

Capsicum L. (Solanaceae) is one of the most economically important vegetable crops with versatile applications as a food and source of spice, as well as for ornamental and pharmaceutical purposes (Qin *et al.* 2014). The tropical regions of South America are believed to be the origin of peppers which are now grown worldwide (Greenleaf 1986). Today, the *Capsicum* genus consists of about 25 wild species and five cultivated species which include *C. annuum* L., *C. baccatum* L. var. *pendulum* (Willd.) Eshbaugh., *C. chinense* Jacq., *C. frutescens* L., and *C. pubescens* Ruiz & Pav. (Kumar *et al.* 2006). Species identification in

Capsicum is traditionally based on phenotypic characteristics and hybridization studies (Heiser and Smith 1953; Pickersgill 1988; Hunziker 1998; Onus and Pickersgill 2004). However, species identification based on morphological characteristics is often difficult, since most of these features are also influenced by environmental factors. In the last few decades, genotypic molecular markers have been identified, and are gaining increasing importance in resolving phylogenetic relationships.

Capsicum chinense is well known for its distinctive pungency causing severe heat sensation and has been recorded as the hottest pepper in the world (Bosland and Baral 2007). Characterization of genebank accessions is an

Received August 25, 2017; Revised September 8, 2017; Accepted September 8, 2017; Published December 1, 2017

*Corresponding author Gi-An Lee, gkntl1@korea.kr, Tel: +82-63-238-4882, Fax: +82-63-238-4859

*Corresponding author Jong-Wook Chung, jwchung73@chungbuk.ac.kr, Tel: +82-43-261-2524, Fax: +82-43-271-0413

important step for germplasm conservation, maintenance, and breeding studies. Several methodologies have been widely applied to analyze genetic variability in *Capsicum* species diversity studies. These include using phenotypic markers such as morphological and agronomic descriptors and genotypic markers such as restriction fragment length polymorphism (RFLP) (Lefebvre *et al.* 1993), amplified fragment length polymorphism (AFLP) (Aktas *et al.* 2009), random amplified polymorphic DNA (RAPD) (Adetula 2006; Moses and Umaharan 2012), microsatellite or simple sequence repeat (SSR) (Portis *et al.* 2007; Stigel *et al.* 2009; Pacheco-Olvera *et al.* 2012), random amplified microsatellite polymorphism (RAMPO) (Rai *et al.* 2013) and direct amplification of minisatellite DNA (DAMD-PCR) (Ince *et al.* 2009).

Recently, whole genome sequencing strategies have been used to explore many plant species at the molecular level in rapid and cost-effective ways (Cao *et al.* 2011). Fragments of DNA from the chloroplast (cp) genome have been widely used for phylogenetic reconstruction and species-level identification because of their relatively stable genomic structure and higher evolutionary rate relative to mitochondrial genomes (Dong *et al.* 2012). The complete cp genome sequences of cultivated peppers have been reported (Jo *et al.* 2011; Zeng *et al.* 2016; Shim *et al.* 2016; Park *et al.* 2016; Kim *et al.* 2016). The main objective of this study was to sequence the complete cp genome and examine the genetic architecture of *C. chinense* with a view to resolving its internal relationships.

MATERIALS AND METHODS

Sampling and DNA extraction

The cultivated pepper (Accession No: IT247196) seeds were collected from the National Agrobiodiversity Center, Rural Development Administration, Republic of Korea. Seeds were germinated and grown in a greenhouse, fresh leaves were collected from 40-day-old seedlings and DNA was extracted to construct cp DNA library.

Chloroplast genome sequencing and assembly

An Illumina paired-end cp DNA library (average insert

size of 500 bp) was constructed using the Illumina TruSeq library preparation kit following the manufacturer's instructions. The library was sequenced with 2×300 bp on the MiSeq instrument at LabGenomics (<http://www.labgenomics.co.kr/>). Assembly of complete cp genome sequences was performed by *de novo* assembly of the low coverage whole genome sequence (WGS), via a bio-informatics pipeline (<http://www.phyzen.com>). Briefly, prior to *de novo* cp genome assembly, low quality sequences (quality score < 20; Q20) were filtered out, and the remaining high quality reads were assembled using the CLC Genome Assembler (version beta 4.6, CLC Inc. Aarhus, Denmark) with a 200 - 600-bp overlap size. Cp contigs were selected from the initial assembly by performing a BLAST (ver. 2.2.31) search against known cp sequences. The selected contigs were oriented to construct the complete cp genome structure. Ambiguous nucleotides or gaps were corrected manually to build the complete cp genome.

Chloroplast genome annotation

The web-based program Dual OrganellarGenoMe Annotator (DOGMA, <http://dogma.cccb.utexas.edu/>) was used to annotate the assembled genome using default parameters to predict protein coding, transfer RNA (tRNA) genes and ribosomal RNA (rRNA) genes. Subsequently, BLASTN (ver. 2.2.31) was used to further identify intron-containing gene positions by searching a published cp genome sequence (GenBank accession NC_018552). A cp gene map was constructed using the OrganellarGenome-DRAW software (OGDRAW, <http://ogdraw.mpimp-golm.mpg.de>).

Discovery of SSRs and SNPs

MicroSatellite identification tool (MISA) software (<http://pgrc.ipk-gatersleben.de/misa/>) was used to find the SSR markers present in the cp genome. This software allows the localization and identification of both perfect and compound microsatellites with 1 to 6 nucleotides in the basic repeat unit. To be considered a microsatellite using this approach, a sequence must be present for a minimum of 10 repeat units for mononucleotide motifs, six repeat units for dinucleotide motifs and five repeat units for all motifs of larger size. MISA has been used for SSR identification in

many species including *Eucalyptus* (Ceresini *et al.* 2005) and barley (Thiel *et al.* 2003). To identify SNP and InDel variants in the cp genome, we used msaTovcf program locally modified with an in-house script (Page *et al.* 2016).

Comparative phylogeny

A new multiple sequence alignment was performed by adding 13 species belonging to the Solanaceae family. The resulting alignment was imported in MEGA 6.0 for

inferring maximum parsimony (MP) analysis, setting 1000 bootstrap replications, Tree-Bisection-Regrafting (TBR) algorithm and 10 random addition replicates. All positions containing gaps and missing data were eliminated. Phylogenetic tree was generated by a maximum likelihood (ML) analysis using MEGA 6.0 with 1000 bootstrap replicates (Tamura *et al.* 2013).

Fig. 1. Complete chloroplast genome map of the cultivated pepper (*Capsicum chinense*). Genes drawn inside the circle are transcribed clockwise, while those outside are counterclockwise and marked with two arrows. Differential functional gene groups are color-coded. The GC content variation is shown in the middle circle.

RESULTS

Chloroplast genome assembly

Illumina sequencing generated 10,056,140 paired-end reads, with an average fragment length of 280 bp, which were then analyzed to reveal 3,017,719,480 bp of sequence. After removal of low-quality reads (Q20), the remaining high quality reads were mapped to the reference cp genome of *C. annuum* L. (GenBank accession NC_018552). As nucleic DNA was not excluded during DNA extraction, a total of 128,000 cleaned reads were mapped with an average coverage of 173× on the cp genome. The cp reads extracted from the Illumina dataset were assembled into a total of four contigs. A complete cp genome was constituted without gaps, from overlaps between contigs (Fig. 1). The whole genome sequencing approach successfully allowed micro-reads to be assembled correctly using a reference-guided method. The fully annotated cp genome sequence of the cultivated pepper *C. chinense* has been deposited in the GenBank database under accession number KX913217.

Features of the chloroplast genome

The complete cp genome of *C. chinense* is 156,936 bp in length. It has the typical angiosperm plastome structure, including a pair of inverted repeats (IRA and IRB) of 25,847 bp that separate the large single-copy (LSC) region of 87,330 bp from the small single-copy (SSC) region of 17,912 bp. The overall GC content of the cp genome is

37.7%, while the LSC and SSC are 35.8 and 32.0%, respectively. The cp genome contains 79 protein coding genes, 30 tRNA genes and four rRNA genes, totaling 113 unique genes (Table 1). Of these, 21 genes are duplicated in the inverted repeat regions, 15 genes and six tRNA genes contain one intron, while two genes (*ycf3* and *rps12*) have two introns.

Discovery of SSRs and SNPs

Cp SSRs (cpSSRs) analyzed with MISA Perl script revealed a total of 117 potential SSRs motifs which are located primarily in the intergenic region (Supplementary Table S1). The majority of the cpSSRs in this cp genome are tri-nucleotides (58.11%) and di-nucleotides (36.75%). Only six tetra-nucleotides (5.12%) were present in the *C. chinense* cp genome. SNP and InDel variants extracted from a multi-FASTA alignment with reference cp genome sequence of *C. annuum* (JX270811) revealed a total of 174 mutations (82 SNPs and 92 InDels) and 69 of these variants involving more than one nucleotide (Supplementary Tables S2 and S3). Among the detected variants, 35 SNPs and 7 InDels were observed in the genic region of the cp genome.

Phylogenetic analysis

The phylogenetic relationship between *C. chinense* and other Solanaceae family members was determined by collecting 13 published cp genome sequences from GenBank of the NCBI database (Fig. 2). Resulting alignment and ML phylogenetic tree using MEGA 6.0 strongly indicated that the *C. chinense* was 129-bp larger than the reported *C. chinense* cp genome (KU041709.1) and much closer to *C. annuum* var. *glabriusculum*, a wild progenitor of *C. annuum*, than other *Capsicum* species. All the nodes in the phylogenetic tree received high bootstrap (100%).

DISCUSSION

Cp sequences have been used to identify different plant species with DNA barcoding methods (Walsh and Hoot 2001; Martine *et al.* 2006). Furthermore, massively parallel

Table 1. General features of the *Capsicum chinense* chloroplast genome.

Features	Chloroplast
Genome size (bp)	156,936
GC content (%)	37.7
Total number of genes	113
Protein coding genes	79
No. of rRNA genes	4
No. of tRNA genes	30
No. of gene duplications in IR regions	21
Total introns	17
Single intron (gene)	15
Double intron (gene)	2
Single intron (tRNA)	6

Fig. 2. Maximum likelihood (ML) phylogram of the 13 Solanaceae species using whole chloroplast genome sequences. Numbers above each node indicate the ML bootstrap support values.

Table 2. Genes present in the *Capsicum chinense* chloroplast genome.

Group of genes	Name of genes
Photosystem I	<i>psaA, psaB, psaC, psaI, psaJ, ycf3^{y)}, ycf4</i>
Photosystem II	<i>STpsbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, PpsbN, psbT, psbZ</i>
Cytochrome b6/f	<i>petA, petB^{z)}, petD^{z)}, petG, petL, petN</i>
ATP synthase	<i>atpA, atpB, atpE, atpF^{z)}, atpH, atpI</i>
Rubisco	<i>rbcL</i>
NADH oxidoreductase	<i>ndhA^{z)}, ndhB^{z)x)}, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
Large subunit ribosomal proteins	<i>rpl2^{z)x)}, rpl14, rpl16^{z)}, rpl20, rpl22, rpl23^{x)}, rpl32, rpl33, rpl36</i>
Small subunit ribosomal proteins	<i>rps2, rps3, rps4, rps7^{x)}, rps8, rps11, rps12^{y)x)w)}, rps14, rps15^{x)}, rps16^{z)}, rps18, rps19^{x)}</i>
RNA polymerase	<i>rpoA, rpoB, rpoC1^{z)}, rpoC2</i>
Unknown function protein coding gene	<i>ycf1^{x)}, ycf2^{x)}, ycf15^{x)}</i>
Other genes	<i>ccsA, cemA, clpP1, clpP2, matK</i>
Ribosomal RNAs	<i>rrn16^{x)}, rrn23^{x)}, rrn4.5^{x)}, rrn5^{x)}</i>
Transfer RNAs	<i>trnA-UGC^{z)x)}, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-UCC^{z)}, trnG-GCC, trnH-GUG, trnI-CAU^{x)}, trnI-GAU^{z)x)}, trnK-UUU^{z)}, trnL-UAA^{z)}, trnL-UAG, trnL-CAA^{x)}, trnM-CAU, trnM-CAU, trnN-GUU^{x)}, trnP-UGG, trnQ-UUG, trnR-ACG^{x)}, trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-UAC^{z)}, trnV-GAC^{x)}, trnW-CCA, trnY-GUA</i>

^{z)}Gene containing a single intron; ^{y)}Gene containing two introns; ^{x)}Two gene copies in IRs; ^{w)}*Trans*-splicing gene.

sequencing technologies have efficiently increased the phylogenetic resolution of land plants at low taxonomic levels (Parks *et al.* 2009). In recent years, cp DNA has been considered a promising approach for the study of plant

evolution, as many published chloroplast genome sequences are available in the NCBI database. In this study, Illumina re-sequencing and assembly of the *C. chinense* cp genome has an indicated length of 156,936 bp which is

129-bp larger than the reported *C. chinense* cp genome (156,807 bp). Total GC contents were 37.7%, which is consistent with reported cp genomes of *C. chinense* and other cultivated peppers (Jo *et al.* 2011; Park *et al.* 2016;).

Cp structural rearrangements and gene loss/gain events occur frequently in some angiosperms. The organization and gene architecture of the *C. chinense* cp genome exhibited similar characteristics, with a total of 113 genes as reported in the previous study. However, in the present study only 21 genes were duplicated in IR regions. As shown in Fig. 1 and Table 2, the genome organization appeared to be more conserved among unique genes, as discovered previously in the *Capsicum* species (Park *et al.* 2016; Zeng *et al.* 2016). The conserved gene sequences in the cp genome indicated that maintenance of gene clusters might be essential for normal gene expression. The sequence data generated using the Illumina platform covered a greater depth (173×) than has previously been reported, thus, the cp assembly described here with its new insights supports the suggestion of a previous report recommending the significance of greater genome coverage (Wu *et al.* 2014).

Microsatellite SSRs are widely used as markers in phylogenetic investigations because of their polymorphic nature within the species (Xue *et al.* 2012). A total of 117 potential cpSSRs were identified in the *C. chinense* cp genome with most of them being intergenic. These results are in agreement with trends of SSR density observed in small genomes (Morgante *et al.* 2002). Similarly, pairwise comparison of the *C. chinense* cp genome with the reference cp sequence revealed a total of 174 mutations (82 SNPs and 92 InDels) with 69 of these variants involving more than one nucleotide, indicating that they can be used as molecular markers to study the genetic diversity and genetic structure of the *Capsicum* species.

A new multiple sequence alignment was performed by downloading 13 Solanaceae species from NCBI. Resulting alignment and phylogenetic analysis using MEGA 6.0 indicated *C. chinense* was 129-bp larger than the reported *C. chinense* cp genome (KU041709.1). As shown in Fig. 2, the constructed ML tree indicated three major clades, Capsiceae, Solaneae, and Nicotianeae, with very high overall bootstrap values and is in agreement with the results

of an earlier study (Park *et al.* 2016).

In conclusion, Illumina sequencing followed by *de novo* assembly and extensively annotated cp genome of the most important vegetable crop in the Solanaceae family (*Capsicum*) were reported. The cp genome is well-conserved in terms of size, gene arrangement, and coding sequences, within major subgroups of *Capsicum* as reported. We also performed comparative analyses with cp genomes sequenced in other Solanaceae species. These analyses revealed that cpSSR sequences were particularly frequent in the *C. chinense* cp genome which is similar to other small plastomes of land plants. Moreover, the ML tree determined the evolutionary position of the pepper in the Solanaceae family. This new genomic dataset will enable further exploration of the genetic diversity of *C. chinense* in the genus *Capsicum*.

ACKNOWLEDGEMENTS

This study was carried out with the support of the “Research Program for Agricultural Science & Technology Development (Project No. PJ010898)” and was supported by the 2016 Postdoctoral Fellowship Program of the National Institute of Agricultural Sciences, Rural Development Administration, Republic of Korea.

REFERENCES

- Adetula OA. 2006. Genetic diversity of *Capsicum* using random amplified polymorphic DNAs. *Afr. J. Biotechnol.* 5: 120-122.
- Aktas H, Abak K, Sensoy S. 2009. Genetic diversity in some Turkish pepper (*Capsicum annuum* L.) genotypes revealed by AFLP analyses. *Afr. J. Biotechnol.* 8: 4378-4386.
- Bosland PW, Baral JB. 2007. 'Bhut Jolokia' - The world's hottest known Chile pepper is a putative naturally occurring interspecific hybrid. *Hortscience* 42: 222-224.
- Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, Fitz J, *et al.* 2011. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat. Genet.* 43: 956-963.
- Ceresini PC, Silva CLSP, Missio RF, Souza EC, Fischer CN,

- Guillherme IR, *et al.* 2005. Satellyptus: analysis and database of microsatellites from ESTs of *Eucalyptus*. Genet. Mol. Biol. 28: 589-600.
- Dong W, Liu J, Yu J, Wang L, Zhou S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PLoS ONE 7: e35071.
- Greenleaf WH. 1986. Pepper breeding. In: MJ. Bassett (ed.), Breeding vegetable crops. AVI Publishing Co., Westport.
- Heiser CB, Smith PG. 1953. The cultivated capsicum peppers. Economic Botany 7: 214-227.
- Hunziker AT. 1998. Estudios sobre solanaceae. XLVI. Los Ajies silvestres de argentina (Capsicum). Darwiniana 36: 201-203.
- Ince AG, Karaca M, Onus AN. 2009. Development and utilization of diagnostic DAMD-PCR markers for *Capsicum* accessions. Genet. Res. Crop Evol. 56: 211-221.
- Jo YD, Park J, Kim J, Song W, Hur CG, Lee YH, *et al.* 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 T-S, Lee J-R, Raveendar S, Lee G-A, Jeon Y-A, Lee H-S, *et al.* 2016. Complete chloroplast genome sequence of *Capsicum baccatum* var. *baccatum*. Mol. Breed. 36: 1-5.
- Kumar S, Kumar R, Singh J. 2006. Cayenne/American pepper (*Capsicum* species). In: KV. Peter (ed.). Handbook of herbs and spices. Woodhead, Cambridge.
- Lefebvre V, Palloix A, Rives M. 1993. Nuclear RFLP between pepper cultivars (*Capsicum annuum* L.). Euphytica 71: 189-199.
- Martine CT, Vanderpool D, Anderson GJ, Les DH. 2006. Phylogenetic relationships of andromonoecious and dioecious Australian species of *Solanum* subgenus *Leptostemonum* section *Melongena*: Inferences from ITS sequence data. Syst. Bot. 31: 410-420.
- Morgante M, Hanafey M, Powell W. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat. Genet. 30: 194-200.
- Moses M, Umaharan P. 2012. Genetic structure and phylogenetic relationships of *Capsicum chinense*. J. Am. Soc. Hortic. Sci. 137: 250-262.
- Onus AN, Pickersgill B. 2004. Unilateral incompatibility in *Capsicum* (Solanaceae): Occurrence and taxonomic distribution. Ann. Bot. 94: 289-295.
- Pacheco-Olvera A, Hernandez-Verdugo S, Rocha-Ramirez V, Gonzalez-Rodriguez A, Oyama K. 2012. Genetic diversity and structure of pepper (*Capsicum Annuum* L.) from Northwestern Mexico analyzed by microsatellite markers. Crop Sci. 52: 231-241.
- Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, *et al.* 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microbial Genomics 2: e000056.
- Park H-S, Lee J, Lee S-C, Yang T-J, Yoon JB. 2016. The complete chloroplast genome sequence of *Capsicum chinense* Jacq. (Solanaceae). Mitochondrial DNA Part B 1: 164-165.
- Parks M, Cronn R, Liston A. 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biology 7: 84.
- Pickersgill B. 1988. The genus *Capsicum*: a multidisciplinary approach to the taxonomy of cultivated and wild plants. Biologisches Zentralblatt 107: 381-389.
- Portis E, Nagy I, Sasvari Z, Stigel A, Barchi L, Lanteri S. 2007. The design of *Capsicum* spp. SSR assays via analysis of *in silico* DNA sequence, and their potential utility for genetic mapping. Plant Sci. 172: 640-648.
- Qin C, Yu CS, Shen YO, Fang XD, Chen L, Min JM, *et al.* 2014. Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc. Natl. Acad. Sci. U.S.A. 111: 5135-5140.
- Rai VP, Kumar R, Kumar S, Rai A, Kumar S, Singh M, *et al.* 2013. Genetic diversity in *Capsicum* germplasm based on microsatellite and random amplified microsatellite polymorphism markers. Physiol. Mol. Biol. Plants 19: 575-586.
- Shim D, Raveendar S, Lee JR, Lee GA, Ro NY, Jeon YA, *et al.* 2016. The complete chloroplast Genome of *Capsicum frutescens* (Solanaceae). Appl. Plant Sci. 4: 1600002.
- Stigel A, Gyurjan I, Sasvari Z, Lanteri S, Ganai M, Nagy I. 2009. Patterns of molecular evolution of microsatellite loci in pepper (*Capsicum* spp.) revealed by allele sequencing. Plant Syst. Evol. 281: 251-254.
- 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.

- Thiel T, Michalek W, Varshney RK, Graner A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 106: 411-422.
- Walsh BM, Hoot SB. 2001. Phylogenetic relationships of *Capsicum* (Solanaceae) using DNA sequences from two noncoding regions: The chloroplast *atpB-rbcL* spacer region and nuclear *waxy* introns. Int. J. Plant Sci. 162: 1409-1418.
- Wu ZH, Gui ST, Quan ZW, Pan L, Wang SZ, Ke WD, *et al.* 2014. A precise chloroplast genome of *Nelumbo nucifera* (Nelumbonaceae) evaluated with Sanger, Illumina MiSeq, and PacBio RS II sequencing platforms: insight into the plastid evolution of basal eudicots. BMC Plant Biol. 14: 289.
- Xue J, Wang S, Zhou SL. 2012. Polymorphic chloroplast microsatellite loci in *Nelumbo* (Nelumbonaceae). Am. J. Bot. 99: e240-e244.
- Zeng FC, Gao CW, Gao LZ. 2016. The complete chloroplast genome sequence of American bird pepper (*Capsicum annuum* var. *glabriusculum*). Mitochondrial DNA A DNA Mapp. Seq. Anal. 27: 724-726.