

The complete chloroplast genome sequence of *Sargentodoxa cuneata*: genome structure and genomic resources

Xuejun Cui^a, Xingong Wang^b, Yanjie Wang^a, Qingjun Yuan^c, Ye Shen^c and Linde Liu^a

^aSchool of Life Sciences, Ludong University, Yantai, China; ^bDepartment of Traditional Chinese medicine, Shandong College of Traditional Chinese Medicine, Yantai, China; ^cState Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China

ABSTRACT

Sargentodoxa cuneata is used as traditional Chinese medicine. In this study, we report its complete chloroplast genome by Illumina pair-end sequencing. The total chloroplast (cp) genome size was 158,094 bp in length, containing a pair of inverted repeats of 26,132 bp, separated by large single-copy and small single-copy regions of 86,508 bp and 19,322 bp, respectively. The chloroplast genome of *S. cuneata* encodes 113 different genes, including 79 protein-coding genes, 30 transfer RNAs, and 4 ribosomal RNAs. A total of 84 perfect chloroplast microsatellites were analyzed in the *S. cuneata*. The majority of the SSRs in this chloroplast genome are mononucleotides (66.67%). The reconstructed phylogeny revealed that *S. cuneata* was sister to the remaining Lardizabalaceae.

ARTICLE HISTORY

Received 24 August 2020
Accepted 21 September 2020

KEYWORDS

Chloroplast genome;
Sargentodoxa cuneata;
molecular markers; SSR

Sargentodoxa, a monotypic genus of the Lardizabalaceae has often been placed in its own family, Sargentodoxaceae (Hoot et al. 1995). It consists of the single species, *Sargentodoxa cuneata* (Oliver) Rehder and E. H. Wilson, mostly confined to subtropical China (Wang et al. 2009). *S. cuneata* is used as traditional Chinese medicine or ethnic medicine (called Hongteng or Xuoteng). Wild resources of this species have been seriously deteriorated due to years of over-harvesting, indicating an urgent need for reasonable conservation strategies. It is necessary to develop genomic resources for *S. cuneata* to provide intragenic information for its utilization and conservation. Because of the small and relatively constant size, conserved genome structure, and uniparental inheritance, chloroplast genomes provide a valuable genetic resource for phylogenetic analysis and species conservation (Dong et al. 2017; Dong et al. 2018). In this study, we sequenced and analyzed the chloroplast genome of *S. cuneata* based on the next-generation sequencing method (Li et al. 2018). The objective of this study was to gain a comprehensive understanding of the chloroplast genome of *S. cuneata* by describing its genome structure and features.

Sample of *S. cuneata* was collected from Xinyang, Henan province of China (31°49'24.2"N, 114°06'50.6"E). The voucher specimen (Specimen accession number: HNRC010701) was deposited at the herbarium of the National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences. Total genomic DNA was extracted using

mCTAB protocol (Li et al. 2013). The DNA from silica dried tissue was fragmented to construct 350 bp insert library following the manufacturer's manual (Illumina Inc., San Diego, CA, USA). Approximately 4 Gb of raw data were generated with paired-end 150 bp read length. The chloroplast genome was assembled with GetOrganelle (Jin et al. 2020). Plastomes were annotated with Plann (Huang and Cronk 2015). The annotated genomic sequence had been submitted to GenBank with the accession number MT898426.

The size of the chloroplast genome of *S. cuneata* is 158,094 bp. The cp genome exhibits a quadripartite structure, which includes a pair of inverted repeats (IRa and IRb: 26,132 bp), and separated large single-copy (86,508 bp) and small single-copy (19,322 bp) regions. The GC content of the chloroplast DNA is 38.2%. The chloroplast genome of *S. cuneata* encodes 113 different genes, including 79 protein-coding genes, 30 transfer RNAs (tRNA), and 4 ribosomal RNAs (rRNA). Fifteen distinct genes contain a single intron while *ycf3* and *clpP* each contain two. Simple sequence repeats in the *S. cuneata* chloroplast genomes were detected using GMAT (Wang and Wang 2016) with the minimum repeats of mono-, di-, tri-, tetra-, penta- and hexa-nucleotides being set to 10, 5, 4, 3, 3, and 3, respectively. A total of 84 perfect cp microsatellites were analyzed in the *S. cuneata*. The majority of the SSRs in this chloroplast genome were mononucleotides (66.67%) and almost all of the mononucleotides (88.09%) are composed of A/T.

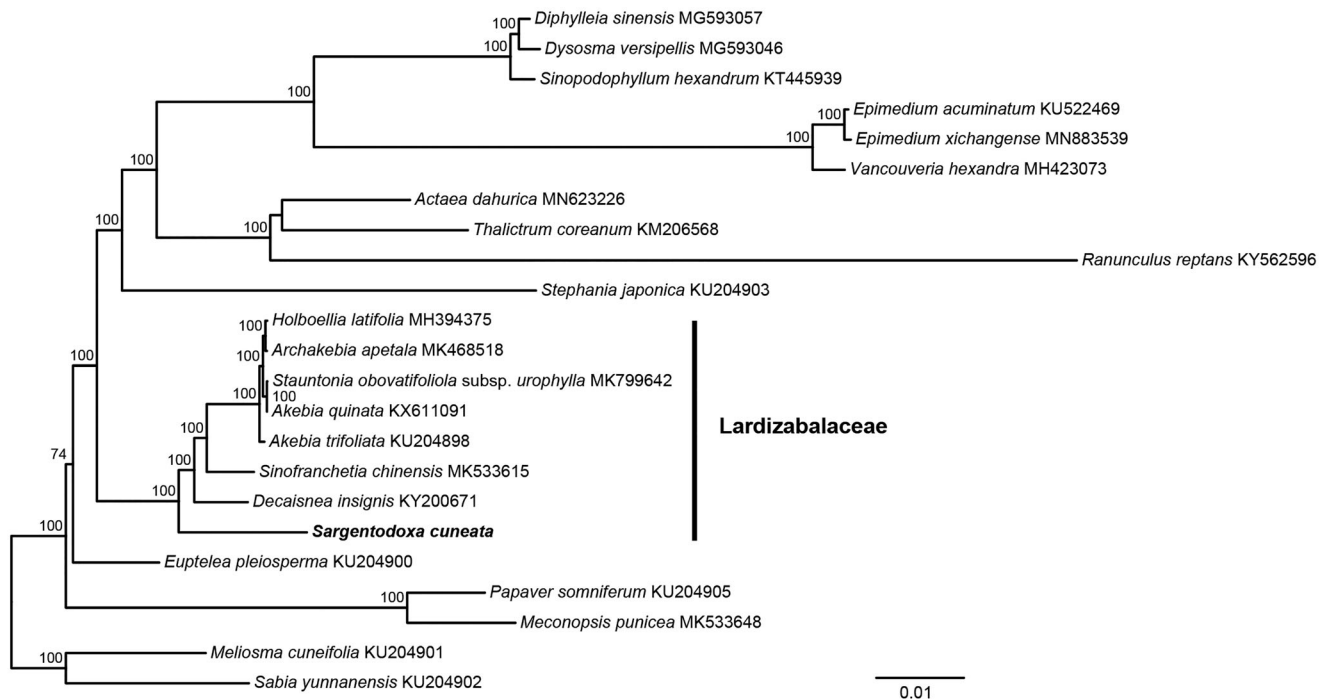


Figure 1. Maximum-likelihood tree of Ranunculales based on the complete chloroplast genome sequences. Bootstrap support values >50% are given at the nodes.

To estimate phylogenetic relationships of *S. cuneata* with other Lardizabalaceae species. Phylogenetic analysis was performed using the whole chloroplast genome sequence. The chloroplast genome sequences were aligned using MAFFT v7 (Kato and Standley 2013). Ambiguous alignment regions were trimmed by Gblocks 0.91 b (Castresana 2002). The maximum-likelihood (ML) analyses were performed in RAxML v.8.1.24 (Stamatakis 2014). The statistical support for the branches was calculated by rapid bootstrap analyses with 1000 replicates. The reconstructed phylogeny revealed that *S. cuneata* was sister to the remaining Lardizabalaceae (Figure 1). The whole chloroplast genome sequences provided sufficient genetic information for species identification, phylogeny analysis, and conservation genetics.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [81573533].

Data availability statement

The chloroplast genome sequence of the *S. cuneata* was submitted to GenBank of NCBI (<https://www.ncbi.nlm.nih.gov>). The accession number from GenBank is MT898426. The raw data have been deposited in SRA under accession no. PRJNA662211 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA662211>).

References

- Castresana J. 2002. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 17(4): 540–552.
- Dong W, Xu C, Li W, Xie X, Lu Y, Liu Y, Jin X, Suo Z. 2017. Phylogenetic resolution in *Juglans* based on complete chloroplast genomes and nuclear DNA sequences. *Front Plant Sci.* 8:1148.
- Dong W, Xu C, Wu P, Cheng T, Yu J, Zhou S, Hong D-Y. 2018. Resolving the systematic positions of enigmatic taxa: manipulating the chloroplast genome data of Saxifragales. *Mol Phylogenet Evol.* 126:321–330.
- Hoot SB, Culham A, Crane PR. 1995. Phylogenetic relationships of the Lardizabalaceae and Sargentodoxaceae: chloroplast and nuclear DNA sequence evidence. In Jensen U, Kadereit JW, editors. *Systematics and evolution of the Ranunculiflorae*. Vienna (Austria): Springer; p. 195–199.
- Huang DI, Cronk QCB. 2015. Plann: a command-line application for annotating plastome sequences. *Appl Plant Sci.* 3(8):1500026.
- Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology.* 21(1):241.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Li J, Wang S, Jing Y, Wang L, Zhou S. 2013. A modified CTAB protocol for plant DNA extraction. *Chin Bull Bot.* 48:72–78.
- Li W, Liu Y, Yang Y, Xie X, Lu Y, Yang Z, Jin X, Dong W, Suo Z. 2018. Interspecific chloroplast genome sequence diversity and genomic resources in *Diospyros*. *BMC Plant Biol.* 18(1):210.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9): 1312–1313.
- Wang H-F, Kirchoff BK, Qin H-N, Zhu Z-X. 2009. Reproductive morphology of *Sargentodoxa cuneata* (Lardizabalaceae) and its systematic implications. *Plant Syst Evol.* 280(3–4):207–217.
- Wang X, Wang L. 2016. GMATA: an integrated software package for genome-scale SSR mining, marker development and viewing. *Front Plant Sci.* 7:1350.