The complete chloroplast genome of *Arcangelisia gusanlung* H.S.Lo (Menispermaceae), an important traditional medicine from South China

Guihua Wen, Sheng Chen, Lingling Fan, Lei Zeng and Bingqiang Xu

*Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China; College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China; Guizhou Institute of Biology, Guiyang, China; Key Laboratory of Forest Cultivation, Conservation and Utilization, Guangdong Academy of Forestry, Guangzhou, China*

**ABSTRACT**

*Arcangelisia gusanlung* H.S.Lo is widely used as a folk medicine by the Dai and Li peoples. Here, we report the first complete chloroplast (cp) genome sequence for this species based on Illumina paired-end sequencing data. The cp genome was 162,509 bp in length with a small single-copy (SSC) region of 20,852 bp, a large single-copy (LSC) region of 91,449 bp, and two separated inverted region of 25,104 bp. In total, 129 unique genes were identified from this genome, including 84 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The GC contents of this genome is 37.8%. Phylogenetic analysis based on 13 complete cp genomes showed a strong sister relationship with *Tinospora cordifolia* (Willd.) Miers and *Tinospora sinensis* (Lour.) Merr. This complete genome of *A. gusanlung* will provide valuable information to elucidate the mechanism of speciation of *Arcangelisia* Becc.

*Arcangelisia* Becc. is a small genus in Menispermaceae, comprised of only four species; *Arcangelisia gusanlung* H.S.Lo is the only recorded plant of the genus in China (Luo et al. 2008). This species has large woody vines and is widely distributed in the south of China including the provinces of Hainan, Guangdong, and Guangxi. The stems and roots of this plants are used in Hainan as a traditional antipyretic medicine much like the traditional Chinese medicinal plant *Phellodendron chinense* Schneid. (Yu et al. 2014). Pharmacological research shows that *A. gusanlung* has anti-inflammatory, analgesia, antipyresis, antitussive, expectorant, and anti diarrheal effects (Hu et al. 2013) that maybe associ-ated with the 15 protoberberine-type alkaloids obtained from this plant. Belonging to a isoquinoline alkalioid class, proto-berberine alkaloids encompass a diverse class of secondary metabolites with many pharmacologically active members, such as berberine and palmatine (Yu et al. 2014). However, until now, genomic data for this important Chinese folk medicine are scarce. Here, we report the first complete chloroplast (cp) genome sequence for *A. gusanlung* and verified its phylogenetic position in Menispermaceae.

The fresh leaves of *A. gusanlung* were collected from the South China Botanical Garden, Chinese Academy of Sciences (CAS) (113.373785 E; 23.187841 N; Wen & Cheng GSL01). The voucher specimen (accession number: IBSC819035) was deposited at the herbarium of South China Botanical Garden (Fei-Yan Zeng, zengfeiy@scib.ac.cn). The total genomic DNA was extracted from the silica-gel dried leaves following the modified Tiangen Kit (Tiangen Biotech, Beijing, China). The extracted DNA was stored in the Plant Resources Conservation and Sustainable Utilization Laboratory in South China Botanical Garden. Then, the genomic library (paired-end, PE = 150 bp) was sequenced on an Illumina HiSeq X Ten platform at Beijing Genomics Institute (Shenzhen, China). The clean data were quality-controlled by using SOAPnuke 2.X (Chen et al. 2018). In total, 17,109,640 reads (2,566,446,000 bp) were obtained and used to assemble the cp genome after filtering and trimming the low-quality reads. The complete cp genome assembly was executed on NOVOPlasty 2.6.3 (Dierckxsens et al. 2017) with the default k-mer of 39-59. The cp genome of *Tinospora cordifolia* (Willd.) Miers (GenBank accession NC_042153) was used as the reference for assembling. Genes were annotated using GeSeq and visually checked in Geneious version 11.0.3 (Kearse et al. 2012; Tillich et al. 2017). Finally, the complete annotated cp genome was submitted to GenBank (accession MW829779).

The complete cp genome of *A. gusanlung* was 162,509 bp in length with the typical quadripartite structure of angiosperms, including a small single-copy (SSC) region of 20,852 bp, a large single-copy (LSC) region of 91,449 bp, and a pair of inverted repeats (IRs) of 25,104 bp. The genome contained 129 genes, including 37 tRNA genes, eight rRNA genes, and 84 protein-coding genes. The overall GC content in the cp genome of *A. gusanlung* was 37.80%, the corresponding values for the SSC, LSC, and IR regions were 32.60, 35.90, and 43.50%, respectively.

To identify the phylogenetic position of *A. gusanlung* in Menispermaceae, the cp genome sequences of *A. gusanlung*,...
10 other species of Menispermaceae and two outgroups (Berberis amurensis Rupr. and Nandina domestica Thunb.) were aligned using MAFFT 7.0 (Katoh and Standley 2013). A maximum-likelihood (ML) tree was constructed by using IQ-TREE (Nguyen et al. 2015) based on the best model of GTR + F + R2 and 1000 bootstrap replicates (Figure 1). The phylogenetic tree indicates that A. gusanlung forms a well supported clade with two species of Tinospora. This Arcangelisia + Tinospora clade is sister to a well supported clade containing all other Menispermaceae species for which cp genome sequence data are publicly available. The plastome data presented here provides a fundamental genomic resource for future conservation and pharmacological studies of A. gusanlung as well as phylogenetic relationships within Menispermaceae.

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

Funding
This work was supported by the National Traditional Chinese Medicine Resources Survey Project of the State Administration of Traditional Chinese Medicine [GZY-KJS-2018-004] and the Guangdong Provincial Forestry Bureau Traditional Chinese Medicine Protection Project of Lingnan.

Data availability statement
The genome sequence data that support the findings of this study are openly available in GenBank of NCBI (https://www.ncbi.nlm.nih.gov/) under the accession no. MW829779. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA758265, SRR15662106, and SAMN21019820, respectively.

References
Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, Li Y, Ye J, Yu C, Li Z, et al. 2018. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. Gigascience. 7(1):1–6.
Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.
Hu H-z, Dong Z, Zhu Y, Chen G-b. 2013. Pharmacology and toxicology of extract from Arcangelisia gusanlung. Chin Herb Med. 5(2):109–115.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.
Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12):1647–1649.
Luo XR, Chen T, Gilbert MG. 2008. Menispermaceae. In: Wu ZY, Reihan PA, Hong DY, editors. Flora of China 7. St. Louis, Beijing: Science Press Beijing, Missouri Botanical Garden Press; p. 1–31.
Nguyen LT, Schmidt HA, Haeseler von A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.
Tillich M, Lehwark P, Pellizzier T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq: versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–W11.
Yu LL, Li RT, Ai YB, Wei L, Deng ZS, Zou ZM. 2014. Protoberberine isoquinoline alkaloids from Arcangelisia gusanlung. Molecules. 19(9):13332–13341.