

MITOGENOME ANNOUNCEMENT



The complete chloroplast genome sequence of *Casimiroa edulis*

Dejun Yang^a, Qiong Qiu^a, Linhong Xu^a, Yumei Xu^a and Yi Wang^{a,b} 

^aInstitute of Tropical Forestry, Yunnan Academy of Forestry, Kunming, People's Republic of China; ^bLaboratory of Forest Plant Cultivation and Utilization, Yunnan Academy of Forestry, Kunming, People's Republic of China

ABSTRACT

The first complete chloroplast genome (cpDNA) sequence of *Casimiroa edulis* was determined from Illumina HiSeq pair-end sequencing data in this study. The cpDNA is 160,176 bp in length, contains a large single-copy region (LSC) of 87,536 bp and a small single-copy region (SSC) of 18,576 bp, which were separated by a pair of inverted repeats (IR) regions of 27,032 bp. The genome contains 131 genes, including 86 protein-coding genes, 8 ribosomal RNA genes, and 37 transfer RNA genes. The overall GC content of the whole genome is 38.2%, and the corresponding values of the LSC, SSC, and IR regions are 36.5, 33.0, and 42.9%, respectively. Further, phylogenomic analysis showed that *C. edulis*, *Phellodendron amurense*, and *Zanthoxylum bungeanum* clustered in a clade in family Rutaceae.

ARTICLE HISTORY

Received 20 October 2019
Accepted 27 October 2019

KEYWORDS

Casimiroa edulis;
chloroplast; Illumina
sequencing; phylogenetic
analysis

Casimiroa edulis is the species of the genus *Casimiroa* within the family Rutaceae and is widely distributed in the tropical and subtropical regions of Mesoamerica including Mexico (Ibrahim et al. 2019), it is an evergreen tree that reaches up to 18 m and is of great value in economy as it produces an

edible fruit popularly known as white sapote (Agustin et al. 2017). With the sweet flavor and medicinal properties, *C. edulis* has been cultivated around the world (Sohn and Epstein 2019). The extracts of *C. edulis* showed anticonvulsant, antihypertensive activities (Magos et al. 1998).

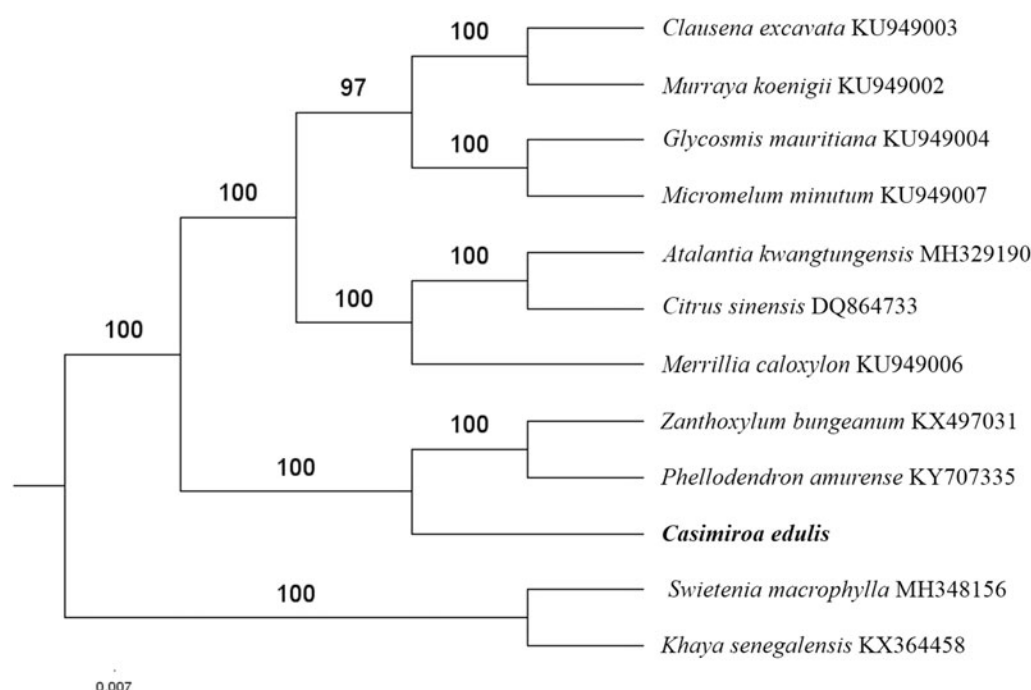


Figure 1. The maximum-likelihood tree based on the 10 chloroplast genomes of *Rutaceae*. The bootstrap value based on 1000 replicates is shown on each node.

Casimiroa edulis also was used as sedative, hypnotic in Mexico (Awaad et al. 2012). However, there have been no genomic studies on *C. edulis*.

Herein, we reported and characterized the complete *C. edulis* plastid genome (MN539263). One *C. edulis* individual (specimen number: 201807055) was collected from Puwen, Yunnan Province of China (23°34'11"N, 101°17'12"E). The specimen is stored at Yunnan Academy of Forestry Herbarium, Kunming, China and the accession number is YAFH0012757. DNA was extracted from its fresh leaves using DNA Plantzol Reagent (Invitrogen, Carlsbad, CA, USA).

Paired-end reads were sequenced by using Illumina HiSeq system (Illumina, San Diego, CA). In total, about 25.1 million high-quality clean reads were generated with adaptors trimmed. Aligning, assembly, and annotation were conducted by CLC de novo assembler (CLC Bio, Aarhus, Denmark), BLAST, GeSeq (Tillich et al. 2017), and GENEIOUS version 11.0.5 (Biomatters Ltd, Auckland, New Zealand). To confirm the phylogenetic position of *C. edulis*, other nine species of family *Rutaceae* from NCBI were aligned using MAFFT version 7 (Katoh and Standley 2013). The Auto algorithm in the MAFFT alignment software was used to align the eight complete genome sequences and the G-INS-i algorithm was used to align the partial complex sequences and maximum likelihood (ML) bootstrap analysis was conducted using RAXML (Stamatakis 2006); bootstrap probability values were calculated from 1000 replicates. *Swietenia macrophylla* (MH348156) and *Khaya senegalensis* (KX364458) were served as the out-group.

The complete *C. edulis* plastid genome is a circular DNA molecule with the length of 160,176 bp, contains a large single-copy region (LSC) of 87,536 bp and a small single-copy region (SSC) of 18,576 bp, which were separated by a pair of inverted repeats (IR) regions of 27,032 bp. The overall GC content of the whole genome is 38.2%, and the corresponding values of the LSC, SSC, and IR regions are 36.5, 33.0, and 42.9%, respectively. The plastid genome contained 131 genes, including 86 protein-coding genes, 8 ribosomal RNA genes, and 37 transfer RNA genes. Phylogenetic analysis showed that *C. edulis*, *Phellodendron amurense*, and *Zanthoxylum bungeanum* clustered in a clade in family *Rutaceae* (Figure 1). The determination of the complete

plastid genome sequences provided new molecular data to illuminate the *Rutaceae* evolution.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Yunnan Key Research and Development Project in Forestry [2018BB007] and Construction project of Xishuangbanna high-efficiency cultivation test and demonstration base for valuable timber tree plantation.

ORCID

Yi Wang  <http://orcid.org/0000-0003-3089-8184>

References

- Agustin JA, Soto M, Famiani F, Cruz-Castillo JG. 2017. *In situ* characterization of fruits and seeds of a number of white sapote (*Casimiroa edulis* Llave & Lex.) accessions in Mexico. Hortscience. 52:1849–1852.
- Awaad AS, Al-Jaber NA, Soliman GA, Al-Outhman MR, Zain ME, Moses JE, El-Meligy RM. 2012. New biological activities of *Casimiroa edulis* leaf extract and isolated compounds. Phytother Res. 26(3):452–457.
- Ibrahim N, El Hawary S, Mohammed M, Ali S, Kandil Z, Refaat E. 2019. Chemical compositions and hypoglycemic activities of the protein and mucilage of *Casimiroa edulis* (Llave & Lex) seeds and fruits. J Appl Pharm Sci. 9:84–91.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.
- Magos GA, Vidrio H, Reynolds WF, Enríquez RG. 1998. Pharmacology of *Casimiroa edulis* IV. Hypotensive effects of compounds isolated from methanolic extracts in rats and guinea pigs. J Ethnopharmacol. 64(1): 35–44.
- Sohn JC, Epstein ME. 2019. *Atemelia casimiroae* sp. nov. (Lepidoptera: Yponomeutoidea: Praydidae): a new pest species on white sapote (*Casimiroa edulis*) in California, United States, reveals unique morphology of spinneret and ventral eversible gland. Zool Anz. 279:52–58.
- Stamatakis A. 2006. RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22(21):2688–2690.
- Tillich M, Lehwark P, Pellizzer 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.