

The complete chloroplast genome sequence of *Xylia xylocarpa*

Yunfen Geng, Yunqing Li, Jinfeng Zhang and Yi Wang 

Laboratory 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 *Xylia xylocarpa* was determined from Illumina HiSeq pair-end sequencing data in this study. The cpDNA is 161,288 bp in length, contains a large single copy region (LSC) of 89,186 bp and a small single copy region (SSC) of 19,354 bp, which were separated by a pair of inverted repeats (IR) regions of 26,370 bp. The genome contains 131 genes, including 86 protein-coding genes, eight ribosomal RNA genes, and 37 transfer RNA genes. The overall GC content of the whole genome is 36.6%, and the corresponding values of the LSC, SSC, and IR regions are 34.1%, 30.8%, and 42.8%, respectively. Further phylogenomic analysis showed that *X. xylocarpa* clustered in a unique clade in *Caesalpinioideae* subfamily.

ARTICLE HISTORY

Received 24 December 2019
Accepted 7 January 2020

KEYWORDS

Xylia xylocarpa; chloroplast;
Illumina sequencing;
phylogenetic analysis

Xylia xylocarpa belongs to the genus of *Xylia* within the Fabaceae and distributes naturally in India, Indo-China, Myanmar, and Thailand (Wattanakupakin et al. 2015). It is a fast-growing tree with a denser wood and yields a hard and durable wood used for heavy construction (Josue 2004). The mature seed can provide a variety of nutrients such as protein, fat, and fiber (Siddhuraju et al. 1995). The extracts of *X. xylocarpa* also showed anti-cholinesterase and memory-improving effects (Lam et al. 2016). However, there have been no genomic studies on *X. xylocarpa*.

Herein, we reported and characterized the complete *X. xylocarpa* plastid genome. The GenBank accession number is MN823700. One *X. xylocarpa* individual (specimen number: 201907030) was collected from Puwen, Yunnan Province of China (23°30'38"N, 101°36'47"E). The specimen is stored at Yunnan Academy of Forestry Herbarium, Kunming, China, and the accession number is ZJFEP119. DNA was extracted from its fresh leaves using DNA Plantzol Reagent (Invitrogen, Carlsbad, CA, USA).

Paired-end reads were sequenced using Illumina HiSeq system (Illumina, San Diego, CA). In total, about 26.8 million high-quality clean reads were generated with adaptors trimmed. Aligning, assembly, and annotation were conducted using CLC de novo assembler (CLC Bio, Aarhus, Denmark), BLAST, GeSeq (Tillich et al. 2017), and GENEIOUS v 11.0.5 (Biomatters Ltd, Auckland, New Zealand). To confirm the phylogenetic position of *X. xylocarpa*, other five species of *Caesalpinioideae* subfamily from NCBI were aligned using MAFFT v.7 (Kato 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.

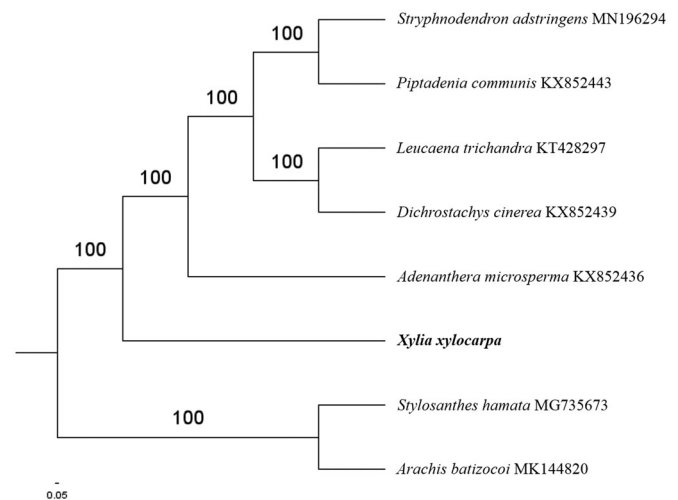


Figure 1. The maximum-likelihood tree based on the six chloroplast genomes of *Caesalpinioideae* subfamily. The bootstrap value based on 1000 replicates is shown on each node.

The maximum-likelihood (ML) bootstrap analysis was conducted using RAxML (Stamatakis 2006); bootstrap probability values were calculated from 1000 replicates. *Stylosanthes hamata* (MG735673) and *Arachis batizocoi* (MK144820) were served as the out-group.

The complete *X. xylocarpa* plastid genome is a circular DNA molecule with the length of 161,288 bp, contains a large single copy region (LSC) of 89,186 bp and a small single copy region (SSC) of 19,354 bp, which were separated by a pair of inverted repeats (IR) regions of 26,370 bp. The overall GC content of the whole genome is 36.6%, and the corresponding values of the LSC, SSC, and IR regions are 34.1%, 30.8%, and 42.8%, respectively. The plastid genome contained 131 genes,

including 86 protein-coding genes, eight ribosomal RNA genes, and 37 transfer RNA genes. Phylogenetic analysis showed that *X. xylocarpa* clustered in a unique clade in *Caesalpinioideae* subfamily (Figure 1). The determination of the complete plastid genome sequences provided new molecular data to illuminate the *Caesalpinioideae* subfamily evolution.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was financially supported by Sub-Projects of the National Key Research and Development Plan [2017YFD0601101-5] and Ten Thousand Industry Leaders Plan Talent of Yunnan Province.

ORCID

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

References

- Josue J. 2004. Some wood properties of *Xylia xylocarpa* planted in Sabah. Sep Bull. 1:1–15.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.
- Lam LMT, Nguyen MTT, Nguyen HX, Dang PH, Nguyen NT, Tran HM, Nguyen HT, Nguyen NM, Min BS, Kim JA, et al. 2016. Anti-cholinesterases and memory improving effects of Vietnamese *Xylia xylocarpa*. Chem Cent J. 10(1):48.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22(21):2688–2690.
- Siddhuraju P, Vijayakumari K, Janardhanan K. 1995. Nutrient and chemical evaluation of raw seeds of *Xylia xylocarpa*: an underutilized food source. Food Chem. 53(3):299–304.
- 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.
- Wattanakulpakin T, Iamtham S, Grubbs KC, Volkaert HA. 2015. Population genetic analysis of *Xylia xylocarpa* (Fabaceae—Mimosoideae) in Thailand. Tree Genet Genomes. 11(1):7.