The complete chloroplast genome of *Tilia endochrysea* Handel-Mazzetti 1926 (Malvaceae)

Huanli Wang\(^{a,b}\), Xi Huang\(^{a,b}\), Lingjun Yan\(^{a,b}\) and Shijie Tang\(^{a,b}\)

\(^{a}\)Jiangsu Key Laboratory for the Research and Utilization of Plant Resources, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing Botanical Garden, Memorial Sun Yat-Sen, Nanjing, China; \(^{b}\)Jiangsu Provincial Germplasm Repository of Indigenous Landscape Tree, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing Botanical Garden, Memorial Sun Yat-Sen, Nanjing, China

**ABSTRACT**

*Tilia endochrysea* Handel-Mazzetti 1926, distributed in the southern margin of Chinese Subtropical Zone, is the only one species in genus *Tilia* had dehiscent fruit. The whole chloroplast genome of *T. endochrysea* was 162,838 bp with 36.46% GC content, including a large single-copy (LSC) region of 91,264 bp, a small single-copy (SSC) region of 20,414 bp, and a pair of inverted repeats (IRs) of 51,160 bp. The chloroplast genome encodes 130 genes, including 85 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The phylogenetic tree shows *T. endochrysea* was the basal taxon in *Tilia* and closely related to *Craigia yunnanensis* (Smith et Evans 1921).

*Tilia* is the only genus that is intermittently distributed in temperate and subtropical regions of the Northern Hemisphere. Genus *Tilia* includes 23 species of deciduous trees with excellent ornamental and economic values. It is a distinct genus characterized by the paddle-shape bract of inflorescence. However, the taxonomy and evolutionary relationships among *Tilia* species are difficult to determine and controversial because of frequent hybridization and polyploidization.

*Tilia endochrysea*, endemic to China, is the only species in section *Endochrysea* H. T. Chang with dehiscent fruits (Pigott 2012). The thin and fragile fruit wall opens by splitting along the lines down the center of each locule to release the seeds. *T. endochrysea* is distributed mainly in mountains of South China, and is one of the few species in this genus that grows in low latitude area (Pigott 2012). The taxonomic characteristics of *T. endochrysea* are relatively clear, and it is considered to be the original species of genus *Tilia* based on morphological and fossil evidence (Pigott 2012). Genome mining of *T. endochrysea* will help in determining the origin and differentiation of *Tilia*.

The fresh leaves of *T. endochrysea* were collected from a mature plant in Nanling National Nature Reserve (113°03′34.93″E, 24°53′25.34″N, Ruyuan, Guangdong). A specimen was preserved in Nanjing Botanical Garden, Memorial Sun Yat-Sen (http://www.cnbg.net/management/xxzx, Huanli Wang, wanghuanli@cnbg.net) under voucher number NGJBJ-Tilia-0005. Total genomic DNA was isolated using a modified cetyltrimethylammonium bromide (CTAB) protocol. Paired-end (2 × 350 bp) sequencing was performed on an Illumina HiSeq 2000 platform (BIOZERON Co., Ltd., Shanghai, China). Approximately, 40 GB raw data (291,608,234 Clean Reads) were generated and deposited in Sequence Read Archive (SRA) under accession number SRR16573320. The complete *T. endochrysea* chloroplast genome was assembled using NOVOPlasty Version 3.3 (Dierckxsens et al. 2016) and SPAdes-3.13.0 package (Antipov et al. 2016) with the complete chloroplast genome of *Tilia mandshurica* Maximowicz 1856 (GenBank number NC028589) as the template. The Geseq web server (https://chlorobase.mpimp-golm.mpg.de/gesenq.html) was used to annotate the chloroplast genome automatically, then it was manually adjusted and visualized in Geneious Prime v.2019.1.3 (Kearse et al. 2012). A final circular chloroplast genome map was drawn using OGDRAW (https://chlorobase.mpimp-golm.mpg.de/OGDraw.html).

The complete chloroplast genome of *T. endochrysea* (GenBank accession number OK624380) is a circular molecule 162,838-bp long with 36.46% GC content. The genome has a typical quadripartite structure consisting of a large single-copy (LSC) region of 91,264 bp and a small single-copy (SSC) region of 20,414 bp that are separated by two inverted repeats (IRs and IRb) of 25,580 bp. The chloroplast genome contains 130 genes, including 85 protein-coding genes, 37 tRNA, and eight rRNA. Of the 130 genes, 17 are duplicated in the IRs. Therefore, there are 113 unique genes in the chloroplast genome. Among them, 16 genes contain intron and three (ycf3, clpP, rps12) had two introns.

**CONTACT** Shijie Tang sjtang2021@163.com

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
To understand the phylogenetic position of *T. endochrysea* in the family Malvaceae (Angiosperm Phylogeny Group 2016), a maximum-likelihood tree (bootstrap replications 1000) inferred by the best-fit model of HKY + F + I was constructed using IQ-TREE v2.0.3 (Minh et al. 2020). We downloaded 23 published chloroplast genomes, including 22 species in Malvaceae and *Bixa orellana* Linnaeus 1753 in family Bixaceae as the outgroup. The phylogeny showed that *T. endochrysea* was the earliest differentiated species in *Tilia* and closely related to *Craigia yunnanensis* (Figure 1). The branch length of *T. endochrysea* and the other five *Tilia* species were extremely short, implying low sequence divergence and a slow rate of evolution in *Tilia* plastid genomes (Cai et al. 2015). Interestingly, the phylogenetic analysis showed that genera *Tilia* and *Craigia* were clustered with Sterculiaceae species in the Cronquist system, whereas genus *Corchorus* species were on an independent branch. Notably, *Tilia, Craigia,* and *Corchorus* were considered as the three genera of Tiliaceae in the Cronquist system, which was a paraphyletic group based on genomic evidence. These findings show there are inconsistencies in the phylogenetic relationships derived from morphology or genomics. To our knowledge, this is the first circular chloroplast genome of *T. endochrysea* with detailed annotation information to be published. We plan to collect and test additional materials to better understand the phylogenetic classification and evolutionary history of genus *Tilia*.

**Authors contributions**

Huanli Wang and Shijie Tang contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Huanli Wang, Lingjun Yan, and Xi Huang. Huanli Wang prepared the first draft of the manuscript. Huanli Wang and Shijie Tang revised the manuscript. All authors read and approved the final manuscript. All authors are responsible for the content of this work.

**Acknowledgements**

We are very grateful to Jiangsu Provincial Germplasm Repository of Indigenous Landscape Tree for providing plant materials during the experiments. We thank Margaret Biswas, PhD, from Liwen Bianji (Edanz) (www.liwenbianji.cn/) for editing the English text of a draft of this manuscript.

**Ethical statement**: The leaves of *T. endochrysea* were collected with the permission granted by Nanling National Nature Reserve (China).

**Disclosure statement**

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

**Funding**

This work was supported by the Natural Science Foundation of Jiangsu Province under Grant Number [BK20170619], National Natural Science Foundation of China under Grant Number [31700477], and Innovation and Extension of Forestry Science and Technology of Jiangsu Province under Grant Number [LYKJ[2019]06].

**ORCID**

Huanli Wang http://orcid.org/0000-0002-4698-010X
Shijie Tang http://orcid.org/0000-0003-1580-3036

**Data availability statement**

The whole chloroplast genome sequence data that support the findings of this study are available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov/ under the accession no. OK624380. The associated BioProject,
SRA, and Bio-Sample numbers are PRJNA774451, SRR16573320, and SAMN22565305, respectively.

References

Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A, Pevzner PA. 2016. PlasmidSPAdes: assembling plasmids from whole genome sequencing data. Bioinformatics. 32(22):3380–3387.

Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc. 181(1):1–20.

Cai J, Ma P, Li H, Li D. 2015. Complete plastid genome sequencing of four Tilia species (Malvaceae): a comparative analysis and phylogenetic implications. PLOS One. 10(11):e0142705.

Dierckxsens N, Mardulyn P, Smits G. 2016. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.

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.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 37(5):1530–1534.

Pigott CD. 2012. Lime-trees and basswoods: a biological monograph of the genus Tilia. Cambridge: Cambridge University Press.