MITOCHONDRIAL DNA PART B
2022, VOL. 7, NO. 6, 1160–1162
https://doi.org/10.1080/23802359.2022.2087562

Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

Chloroplast genome structure and phylogenetic analysis of *Glycosmis parviflora* (Sims) Little 1948, a folk medicinal plant featured in Lingnan Region, China

Aimin Chen[a,b], Fang Li[b], Xuena Xie[b], Rong Huang[b], Enwei Tian[b] and Zhi Chao[a,b,c]

[a]Department of Pharmacy, Zhujiang Hospital, Southern Medical University, Guangzhou, People's Republic of China; [b]Faculty of Medicinal Plants and Pharmacognosy, School of Traditional Chinese Medicine, Southern Medical University, Guangzhou, People's Republic of China; [c]Guangdong Provincial Key Laboratory of Chinese Medicine Pharmaceutics, Guangzhou, China

ABSTRACT

Glycosmis parviflora is the most widely spread and the most morphologically varied species of Chinese *Glycosmis*, and its roots and leaves serve as folk medicines. We sequenced the complete chloroplast (cp) genome of *G. parviflora*. The cp genome obtained was a circular DNA molecule of 159,825 bp in length, containing one large and one small single copy region (LSC and SSC) of 87,517 and 18,352 bp separated by a pair of 26,978 bp inverted repeat regions (IRs). The overall GC content of the cp genome was 38.40%. The phylogenetic analysis revealed that *Glycosmis* was strongly supported as a monophyletic group belonging to Clauseneae, and *G. parviflora* was closely related to *G. pentaphylla*. The results will provide the basis for the further study of molecular markers and phylogeny of *G. parviflora*.

ARTICLE HISTORY
Received 14 April 2022
Accepted 6 June 2022

KEYWORDS
Glycosmis parviflora; complete chloroplast genome; phylogeny

Glycosmis is a group of Rutaceaous plants and evergreen glabrous shrub with orange berries (Teja et al. 2021). *G. parviflora* (Sims) Little 1948 is the most widespread species of the genus in China (Zhang and Hartley 2008). It is a commonly used medicinal plant in Lingnan Region (which is a loosely defined geographic areas south to Nangling Mountains, mainly covering Guangdong Province and Guangxi Zhuang Autonomous Region) of China. In traditional Chinese medicine, the root of *G. parviflora* is used to eliminate phlegm and relieve cough, and its leaves can dissipate stasis and disperse swelling (Wang 2014). It is also used in other folk medicines (Teja et al. 2021), and the compounds it contains have been highly explored for anticancer activity (Knölker and Reddy 2008). Widely distributed in Lingnan Region, *G. parviflora* shows great variations in the length of its inflorescences, ranging from 3 cm to 14 cm, and in the morphologies of the leaves, from a simple leaf, unifoliate leaf to a compound leaf with 2–6 leaflets; the leaflets also vary in the shapes, which can be elliptic, oblong or lanceolate (Mou et al. 2012). To identify molecular markers for species and provide information for its phylogenetic position, we sequenced, assembled and annotated the complete chloroplast (cp) genome of *G. parviflora*.

Fresh leaves of *G. parviflora* were collected from the medicinal plant garden of Southern Medical University (113°19′43.35"E, 23°11′20.58"N), Guangzhou, China. The collections were in accordance with regulations including our university and regional, national, or international ones. The voucher specimen (Chao Zhi 20200405016) was identified by Professor Zhi Chao and deposited in the herbarium of the School of Traditional Chinese Medicine, Southern Medical University (contact Zhi Chao, chaozhi@smu.edu.cn). The DNA of *G. parviflora* was extracted by the modified CTAB method (Yang et al. 2014). Sequencing was performed on BGISEQ-500 platform in high output mode with 150 bp paired-end reads at Beijing Genomics Institute (BGI, Shenzhen, China). The cp genome of *G. parviflora* was assembled by using GetOrganelle v.1.7.1 (Jin et al. 2020) with *G. mauritiana* and *G. pentaphylla* as references (Accession No. NC_032686 and NC_032687). The assembled genome was annotated using GeSeq (Annotation of Organellar Genomes) (https://chlorobox.mpimp-golm.mpg.de/geseq.html) (Tillich et al. 2017) and Plastid Genome Annotator (PGA) software (Qu et al. 2019). The annotated sequence has been deposited in GenBank (Accession No. MW714375).

The complete cp genome of *G. parviflora* showed a typical circular tetramerous structure of 159,825 bp in length. Its quadripartite structure is composed of a large single-copy (LSC) region of 87,517 bp and a small single-copy (SSC) region of 18,352 bp, separated by a pair of inverted repeats (IRs) of 26,978 bp. The complete cp genome encodes 124 genes, including 85 protein-coding genes (PCGs), 30 transfer RNA genes, 8 ribosomal RNA genes and 1 pseudogene(*ycf1*). Among these genes, 92 genes are single copy, while 8 PCGs, 4 tRNA genes, and 4 rRNA genes in IR regions are duplicated. Moreover, there are ten genes containing one intron, and three genes containing two introns. The GC content of the

CONTACT Zhi Chao chaozhi@smu.edu.cn School of Traditional Chinese Medicine, Southern Medical University, Guangzhou 510515, China
Supplemental data for this article is available online at https://doi.org/10.1080/23802359.2022.2087562.

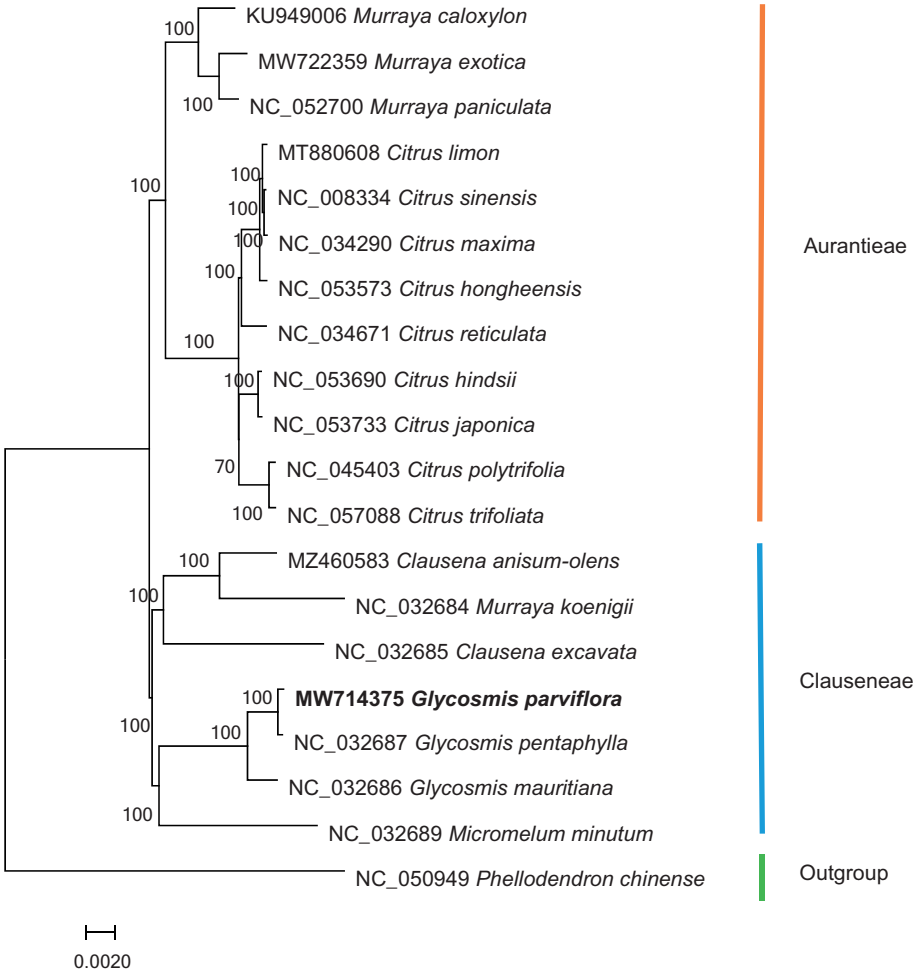

Figure 1. The maximum-likelihood (ML) tree based on 20 chloroplast genomes, with the best fitting substitution model (GTR + F+R2). Bootstrap support values were indicated at each node. *Phellodendron chinense* was used as outgroup.

overall chloroplast genome, IR regions, LSC and SCC are 38.40, 42.9, 36.7, 32.9%, respectively.

In order to gain insight into its phylogenetic position, a phylogenetic analysis was performed using the complete cp genomes of *G. parviflora* and other 19 species of Rutaceae. The cp genomes of the 19 species were downloaded from GenBank (species names and accession numbers shown in Figure 1). All the cp genomes were aligned with MAFFT (Rozewicki et al. 2019), and then edited manually by MEGA X (Kumar et al. 2018). The maximum likelihood (ML) tree was inferred in RAxML (Stamatakis 2014) with the best fitting substitution model (GTR + F+R2) determined by the Akaike information criteria (AIC). The bootstrap support was calculated with 1000 replications.

In the phylogenic tree with *Phellodendron chinense* as an outgroup species, two clades were clearly recognized, which corresponded to Clauseneae and Aurantieae repectively. *Glycosmis* was a strongly supported monophyletic group (PP = 100) in Clauseneae, and was the sister group to *Micromelum*. *G. parviflora* was closely related to *G. pentaphylla* (Figure 1).

Author contributions

Z C conceived the study, reviewed and revised the drafts of the paper. AM C analyzed data, wrote and revised the manuscript. F L performed the genome assembly and annotation. XN X conducted the molecular experiments, and assisted with data analysis. R H assisted in manuscript revision. EW T collected plant materials and assisted with the molecular experiments. All authors read and approved the final manuscript.

Disclosure statement

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

Funding

This research was partly supported by National Natural Science Foundation of China [Grant No. 81373905], and Guangdong Natural Science Foundation [Grant No. 2014A030313321].

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov/ under the accession no. MW714375. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA820442, SRX14636471, and SAMN27008371, respectively.

References

Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21(1):241.

Knölker HJ, Reddy KR. 2008. Biological and pharmacological activities of carbazole alkaloids. Alkaloids - Chem Biol. 65:181–193.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 35(6):1547–1549.

Mou FJ, Ma SZ, Xu QH, Zhang DX. 2012. Floral morphology and anatomy of Glycosmis (Rutaceae) in China and their systematic implications. J. Plant Sci. 30(4):327–336.

Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods. 15:50.

Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. 2019. MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. 47(W1):W5–W10.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313.

Teja PK, Patel P, Bhavsar D, Bindusri C, Jadhav K, Chauthe SK. 2021. Traditional uses, phytochemistry, pharmacology, toxicology and formulation aspects of Glycosmis species: a systematic review. Phytochemistry. 190:112865.

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.

Wang GQ. 2014. Nationwide compilation of traditional and folk medicinal materials. 3rd ed. Vol. 2. Beijing: People's Medical Publishing House; p. 132.

Yang JB, Li DZ, Li HT. 2014. Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. Mol Ecol Resour. 14(5):1024–1031.

Zhang DX, Hartley TG. 2008. GLYCOSMIS. In Flora of China. Vol. 11. St. Louis (MO): Missouri Botanical Garden Press; p. 80–83.