

Complete chloroplast genome sequence of the medicinal plant *Gynostemma pentaphyllum*

Xin-Yi Zeng, Li-Qiu Zou, Xue-Jun Kuang and Ying Li

Institute of Medicinal Plant Development, China Academy of Medical Sciences & Peking Union Medical College, Beijing, China

ABSTRACT

We report the complete chloroplast genome sequence of *Gynostemma pentaphyllum*, a well-known traditional Chinese medicine, which produces triterpenoid saponins similar to *Panax ginseng*. The assembled chloroplast genome (cpDNA) was 157,654 bp in length and structurally divided into four distinct regions, namely, large single copy region (86,794 bp), small single copy region (18,654 bp) and a pair of inverted repeat regions (26,103 bp). A total of 143 genes were annotated, including 87 protein-coding genes, 10 tRNA genes and 46 rRNA genes. Phylogenetic analysis revealed that the chloroplast genome sequence of *G. pentaphyllum* is most closely related to *Cucumis melo*.

ARTICLE HISTORY

Received 13 April 2016
Revised 20 May 2016
Accepted 31 May 2016

KEYWORDS

Chloroplast genome;
Gynostemma pentaphyllum;
phylogenetic relationship

Gynostemma pentaphyllum is a perennial creeping herb of the genus *Gynostemma* in the Cucurbitaceae family and is widely distributed throughout China, India, Myanmar, Korea and Japan (Chen 1995). *G. pentaphyllum* is a traditional Chinese medicinal herb that contains more than 100 dammarane-type saponins (Yin et al. 2004; Liu et al. 2005). These saponins possess important biological functions and are responsible for its beneficial effects against inflammation, hyperlipidemia, cardiovascular disease and cancer (Attawish et al. 2004; Circosta et al. 2005; Wang et al. 2007; Yan et al. 2014). Elucidation of the complete *G. pentaphyllum* cp genome would enrich the chloroplast

genome information of the Cucurbitaceae family and contribute to further research regarding its evolution and molecular identification.

The total DNA was isolated from the mature leaves of *G. pentaphyllum*, which were collected from the base of Good Agricultural Practice of Medicinal Plants and Animals (GAP) for *G. pentaphyllum* in Pingli County, Shanxi Province. Pingli County is one of the major *G. pentaphyllum* producing areas in China. The specimen has been stored in the Institute of Medicinal Plant Development (IMPLAD) and the accession number is Gp_PL01. A Illumina Hiseq 4000 paired-end (PE) genomics library with 150bp read length was constructed

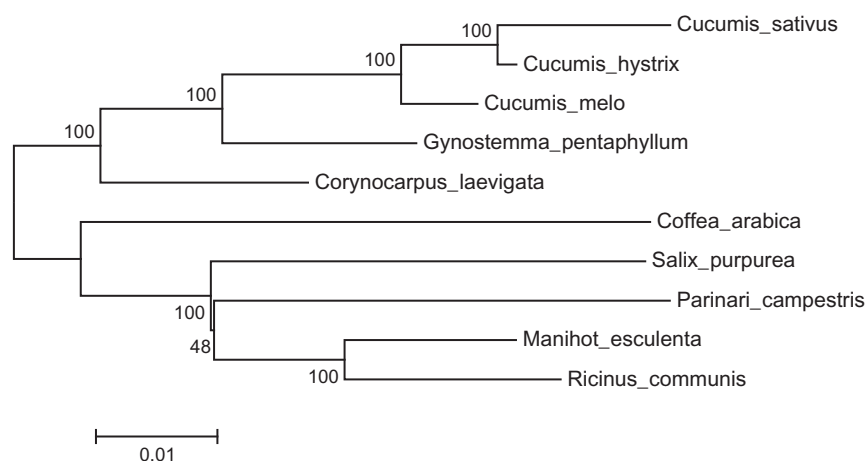


Figure 1. Neighbor-joining tree of the Cucurbitales and related families based on total protein-coding genes. Numbers in the nodes are boot-strap values from 1000 replicates. The chloroplast sequence of *Coffea arabica*, a species of the Gentianales, was set as an outgroup. The GenBank numbers are as follows: *Cucumis sativus*, AJ970307; *Cucumis hystrix*, KF957866; *Cucumis melo*, JF412791; *Gynostemma pentaphyllum*, KX014626; *Corynocarpus laevigata*, HQ207704; *Coffea arabica*, EF044213; *Salix purpurea*, NC_026722; *Parinari campestris*, NC_024067; *Manihot esculenta*, NC_010433; *Ricinus communis*, JF937588.

CONTACT Li Ying  liying@implad.ac.cn  Institute of Medicinal Plant Development, China Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China

© 2016 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 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

and sequenced. The cp reads from the total data were found by aligning the PE reads to the reference cp genome, *Cucumis sativus* (AJ970307) and *Cucumis melo* (JF412791), which belong to the genus *Cucumis* in the Cucurbitaceae family. Read number analysis indicated that the total DNA contained ~5% cp DNA. The cp reads were used in the assembly, and the resulting cp genome contigs were manually checked to confirm the ligation between the LSR-IR and SSR-IR boundary. Dual Organellar GenoMe Annotator (DOGMA) (Wyman et al. 2004) was used to predict protein-coding genes, transfer RNA (tRNA) genes and ribosome RNA (rRNA) genes. BLASTX was conducted to identify further the positions of genes with intron by searching against the reference cp genome. The complete cp genome sequence, together with gene annotations, was submitted to the GenBank with accession number KX014626.

The complete chloroplast genome of *G. pentaphyllum* was a circular form of 157,654 bp in length, which was divided into four distinct regions, namely, 86,794 bp of large single copy region, 18,654 bp small single copy region and 26,103 bp of a pair of inverted repeat regions. The chloroplast genome contained annotated genes, including 87 protein-coding genes, 10 tRNA genes and 46 rRNA genes.

A neighbor-joining tree was constructed in MEGA 6.0 (Tamura et al. 2013) using the total chloroplast protein-coding sequences of *G. pentaphyllum*, other Cucurbitales plants and other species. The phylogenetic tree showed that the Cucurbitales, which includes *G. pentaphyllum* and the genus *Cucumis*, was clustered to a branch (Figure 1).

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

References

- Attawish A, Chivapat S, Phadungpat S, Bansiddhi J, Techadamrongsin Y, Mitrijit O, Chaorai B, Chavalittumrong P. 2004. Chronic toxicity of *Gynostemma pentaphyllum*. *Fitoterapia*. 75:539–551.
- Chen SK. 1995. A classificatory system and geographical distribution of the genus *Gynostemma* BL. (Cucurbitaceae) *Acta Phytotaxon Sin.* 33:403–410.
- Circosta C, De Pasquale R, Occhiuto F. 2005. Cardiovascular effects of the aqueous extract of *Gynostemma pentaphyllum* Makino. *Phytomedicine*. 12:638–643.
- Liu X, Ye W, Mo Z, Yu B, Wu H, Zhao S, Che C, Hsiao WL. 2005. Three dammarane-type saponins from *Gynostemma pentaphyllum*. *Planta Med.* 71:880–884.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 30:2725–2729.
- Wang QF, Chiang CW, Wu CC, Cheng CC, Hsieh SJ, Chen JC, Hsieh YC, Hsu SL. 2007. Gypenosides induce apoptosis in human hepatoma huh-7 cells through a calcium/reactive oxygen species-dependent mitochondrial pathway. *Planta Med.* 73:535–544.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 20:3252–3255.
- Yan H, Wang X, Niu J, Wang Y, Wang P, Liu Q. 2014. Anti-cancer effect and the underlying mechanisms of gypenosides on human colorectal cancer SW-480 cells. *Plos One*. 9:e95609.
- Yin F, Hu L, Lou F, Pan R. 2004. Dammarane-type glycosides from *Gynostemma pentaphyllum*. *J Nat Prod.* 67: 942–952.