

Complete chloroplast genome sequence of *Peganum harmala* (Zygophyllaceae)

Pei-Pei Jiao^{a,b*}, Wei Si^{c*}, Wen-Rui Qu^a, Shan-He Zhang^a, Tian-Ge Yang^d and Zhi-Hua Wu^d 

^aXinjiang Production & Construction Corps Key Laboratory of Protection and Utilization of Biological Resources in Tarim Basin, College of Life Science, Tarim University, Alar, PR China; ^bCollege of Life Science and Technology, Huazhong Agricultural University, Wuhan, PR China; ^cSecurity Department, Tarim University, Alar, PR China; ^dHubei Provincial Key Laboratory for Protection and Application of Special Plant Germplasm in Wuling Area of China, College of Life Sciences, South-Central University for Nationalities, Wuhan, PR China

ABSTRACT

Peganum harmala L. is a perennial herbaceous plant belonging to the family of Zygophyllaceae, and is grows in semi-arid climates, such as Xinjiang, Gansu, Ningxia, Qinghai, and Inner Mongolia in China, and also Middle East and North Africa. This species is of high medicinal value. The complete chloroplast genome was reported in this study. The chloroplast genome with a total size of 159,957 bp consists of two inverted repeats (IR, 26,550 bp) separated by a large single-copy region (LSC, 88,098 bp) and a small single-copy region (SSC, 18,759 bp). Further annotation revealed the chloroplast genome contains 113 genes, including 79 protein-coding genes, 30 *tRNA* genes, and four *rRNA* genes. A total of 90 simple sequence repeats (SSRs) were identified in the chloroplast genome. This information will be useful for study on the evolution and genetic diversity of *Peganum harmala* in the future.

ARTICLE HISTORY



Received 31 January 2021
Accepted 2 March 2021

KEYWORDS

Peganum harmala;
chloroplast genome; evolution

Peganum harmala L. is a perennial herbaceous plant belonging to the family of Zygophyllaceae, is grows in semi-arid climates, such as Xinjiang, Gansu, Ningxia, Qinghai, and Inner Mongolia in China and also Middle East and North Africa (Mahmoudian et al. 2002; Yonezawa et al. 2011). It is widely and frequently used as traditional folk medicine in Northwest and Northeast China, and also in Iranian (Niroumand et al. 2015). The alkaloids extracted from *P. harmala* seeds or whole herb have many pharmacological activities including insecticidal (Jbilou and Sayah 2007), antibacterial (Sodaeizadeh et al. 2010; Shaheen and Issa 2020), antihyperglycemic (Waki et al. 2007; Wang et al. 2015), antipsoriatic (Zeng et al. 2013), antitumor, and antileukemic effects (Zaker et al. 2007; Xue et al. 2015; Liu et al. 2012). Some of the pharmacologic properties have been confirmed by different studies in modern phytotherapy, harmine for example, was reported to show a unique therapeutic promise for human diabetes and anticancer therapy (Wang et al. 2015; Ding et al. 2019; Zhang et al. 2020). However, *P. harmala* harvested from Xinjiang region shows a higher genuineness of Chinese herbs for its significant higher harmine and harmaline contents (Sun et al. 2017). Due to the characteristics of uniparentally inheritance, highly conserved gene content and quadripartite organization, the chloroplast genome is widely used for used for research on plant phylogeny (Wu et al. 2014). In this study, to obtain the new insight into the phylogeny of *P. harmala*, we sequenced, assembled, and annotated the accurate chloroplast genome.

The materials of *P. harmala* in this study were collected from Aheqi County, Kizilsu Kirghiz Autonomous Prefecture, Xinjiang province of China (78°39.366'E, 40°57.837'N, 1876 m above sea level). The voucher specimen (TD-00562, *Peganum harmala* L.) was stored in the Germplasm Bank of Wild Species (<http://www.genobank.org>) and the herbarium of Tarim University. The complete genomic DNA was extracted using CTAB method (Doyle JJ and Doyle JL 1987) and sequenced using the Illumina NovaSeq 6000 platform at Majorbio Company (Shanghai, China). The assembly of chloroplast genome was described as follows in brief. First, the clean reads were quality-controlled by FastQC version 0.11.9 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Then, the whole chloroplast genome was assembled using GetOrganelle version 1.7.3 (Jin et al. 2020). Finally, to check the accuracy of final assembly, the slimmed assembly graph and selected target assembly graph were further visualized by Bandage version 0.8.1 (Wick et al. 2015). Gene annotation was performed using CPGAVAS2 (<http://47.96.249.172:16019/analyzer/annotate>) (Shi et al. 2019), and PGA (<https://github.com/quxiaojian/PGA>) (Qu et al. 2019). Annotations of protein-coding sequences were confirmed using BLASTx in NCBI. The complete chloroplast genome was 159,957 bp (MW307830) and composed of two IRs of 26,550 bp each, which divide into a large single copy (LSC) region of 88,098 bp and a small single copy (SSC) region of 18,759 bp, the average GC content was 37.53%. The chloroplast genomes encoded 113 functional genes, including 79

CONTACT Zhi-Hua Wu  zhwu@scuec.edu.cn  Hubei Provincial Key Laboratory for Protection and Application of Special Plant Germplasm in Wuling Area of China, College of Life Sciences, South-Central University for Nationalities, Wuhan, PR China

*These authors have contributed equally to this work

© 2021 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.

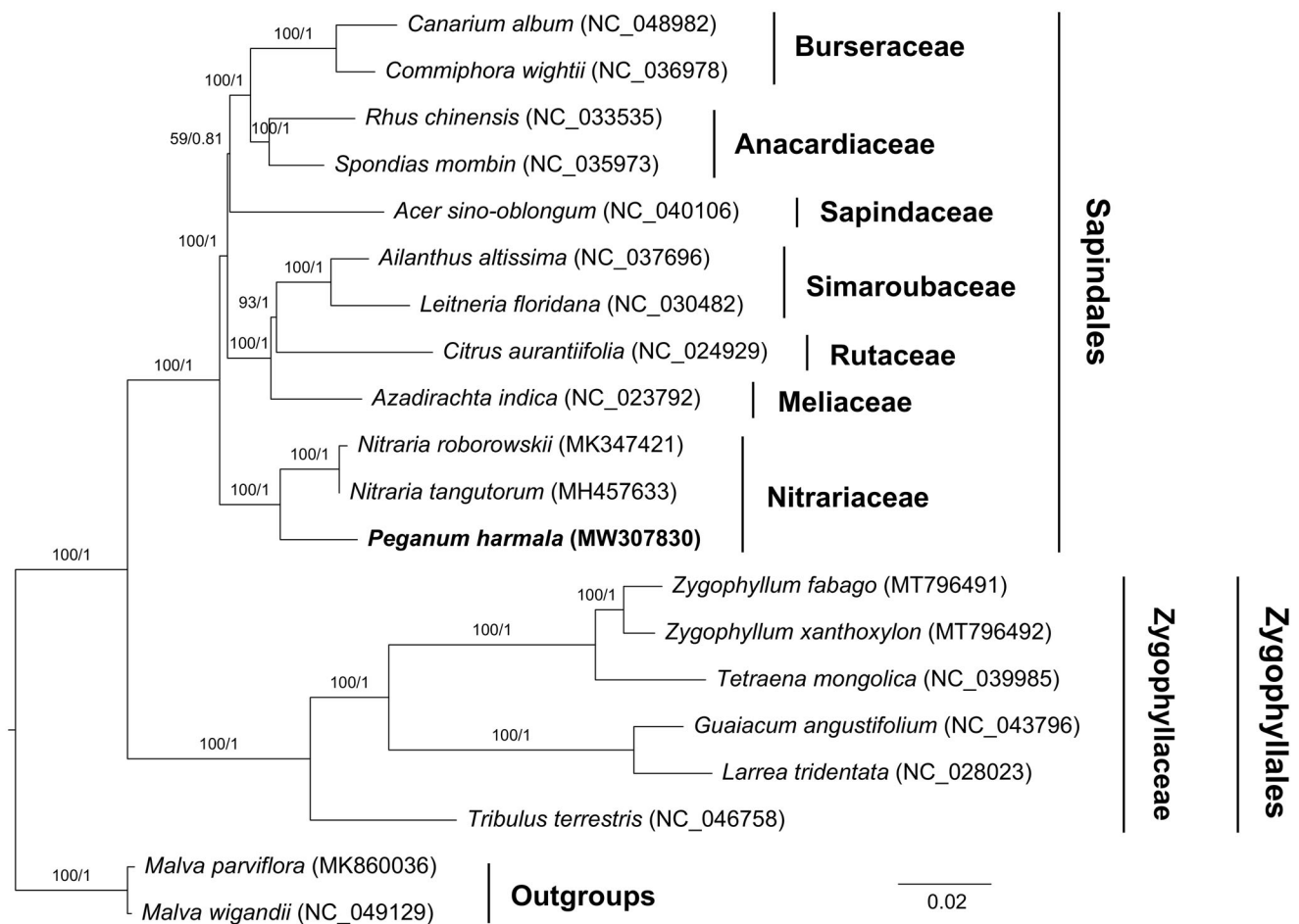


Figure 1. Phylogenetic tree reconstructed by maximum-likelihood (ML) and Bayesian inference (BI) analysis based on the 79 chloroplast protein-coding genes from 20 species.

protein-coding genes, 30 *tRNA* genes, and four *rRNA* genes. A total of 90 SSR markers ranging from mononucleotide to hexa-nucleotide repeat motif were identified in *P. harmala* chloroplast genome.

To explore the phylogenetic relationship of *P. harmala* within Zygophyllaceae, additional 11 species from Sapindales and six species from Zygophyllaceae were studied. With the *Malva parviflora* and *Malva wigandii* as the outgroups, the phylogenetic trees were built from the 79 protein-coding gene by maximum-likelihood (ML) and Bayesian inference (BI) (Figure 1). The ML tree was generated using IQ-TREE version 2.1.2 (Nguyen et al. 2015) based on the best model of TVM + F + R3 and 1000 bootstrap replicates, and BI analysis was performed in MrBayes version 3.2.7 (Ronquist et al. 2012). This result showed that the analyzed *P. harmala* was closer to the species of *Nitraria tangutorum* and *Nitraria roborowskii*.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was financially supported by the National Natural Sciences Foundation of China [31660069], The Innovative Team Building Plan for

Key Areas of Xinjiang Production and Construction Corps [2018CB003], Xinjiang Production & Construction Corps Key Laboratory of Protection and Utilization of Biological Resources in Tarim Basin [BRZD2003], Hubei Provincial Natural Science Foundation of China [2019CFB214].

ORCID

Zhi-Hua Wu  <http://orcid.org/0000-0002-2895-7256>

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/>] (<https://www.ncbi.nlm.nih.gov/>) under the accession no. MW307830. The associated 'BioProject', 'SRA', and 'Bio-Sample' numbers are PRJNA692776, SRR13451271, and SAMN17369979, respectively.

References

- Ding Y, He J, Huang J, Yu T, Shi X, Zhang T, Yan G, Chen S, Peng C. 2019. Harmine induces anticancer activity in breast cancer cells via targeting TAZ. *Int J Oncol.* 54(6):1995–2004.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem Bull.* 19:11–15.
- Jbilou R, Sayah F. 2007. Effects of *Peganum harmala* (Zygophyllaceae) seed extracts on the development of *Tribolium castaneum* (Coleoptera: Tenebrionidae). *JTI.* 27(3–4):199–209.

- Jin J-J, Yu W-B, Yang J-B, Song Y, de Pamphilis CW, Yi T-S, Li D-Z. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21(1):241.
- Liu JL, Wu QC, Wang ZH, Hong Q, Wang JH, Han B, Fan CB, Tian M. 2012. Harmine-inducing apoptosis of Jurkat Leukemia cells and its mechanism *in vitro*. *CJGMCM.* 27(2):242–244.
- Mahmoudian M, Jalilpour H, Salehian P. 2002. Toxicity of *Peganum harmala*: review and a case report. *Iranian J Pharmacol Ther.* 5:1–4.
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Niroumand MC, Farzaei MH, Amin G. 2015. Medicinal properties of *Peganum harmala* L. in traditional Iranian medicine and modern phytotherapy: a review. *J Tradit Chin Med.* 35(1):104–109.
- 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(1):50.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 61(3):539–542.
- Shaheen HA, Issa MY. 2020. *In vitro* and *in vivo* activity of *Peganum harmala* L. alkaloids against phytopathogenic bacteria. *Sci Hortic-Amsterdam.* 264:108940.
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73.
- Sodaeizadeh H, Rafieolhossaini M, Van Damme P. 2010. Herbicidal activity of a medicinal plant, *Peganum harmala* L. and decomposition dynamics of its phytotoxins in the soil. *Ind Crop Prod.* 31:358–394.
- Sun XH, Hu HH, Lu JR. 2017. Comparison of contents of two alkaloids in *Peganum harmala* L. *J Qian Med Coll Natl.* 30(1):1–5.
- Waki H, Park KW, Mitro N, Pei L, Damoiseaux R, Wilpitz DC, Reue K, Saez E, Tontonoz P. 2007. The small molecule harmine is an antidiabetic cell-type-specific regulator of PPAR γ expression. *Cell Metab.* 5(5):357–370.
- Wang P, Alvarez-Perez JC, Felsenfeld DP, Liu HT, Sivendran S, Bender A, Kumar A, Sanchez R, Scott DK, Garcia-Ocaña A, et al. 2015. A high-throughput chemical screen reveals that harmine-mediated inhibition of DYRK1A increases human pancreatic beta cell replication. *Nat Med.* 21(4):383–388.
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics.* 31(20):3350–3352.
- Wu ZH, Gui ST, Quan ZW, Pan L, Wang SZ, Ke WD, Liang DQ, Ding Y. 2014. A precise chloroplast genome of *Nelumbo nucifera* (Nelumbonaceae) evaluated with Sanger, Illumina MiSeq, and PacBio RS II sequencing platforms: insight into the plastid evolution of basal eudicots. *BMC Plant Biol.* 14:289.
- Xue XQ, He WT, Zhang HL. 2015. Research progress of Harmel antitumor effect. *J Chin Med.* 5:628–630.
- Yonezawa T, Hasegawa S-i, Asai M, Ninomiya T, Sasaki T, Cha B-Y, Teruya T, Ozawa H, Yagasaki K, Nagai K, et al. 2011. Harmine, a β -carboline alkaloid, inhibits osteoclast differentiation and bone resorption *in vitro* and *in vivo*. *Eur J Pharmacol.* 650(2–3):511–518.
- Zaker F, Oody A, Arjmand A. 2007. A study on the antitumoral and differentiation effects of *Peganum harmala* derivatives in combination with ATRA on leukaemic cells. *Arch Pharm Res.* 30(7):844–849.
- Zeng CL, Ma GZ, Teng L, Dai XY, Yu FS, Wang CH. 2013. Effects of alkaloids from *Peganum harmala* on mice models of psoriasis. *Chin Tradit Patent Med.* 35(2):219–225.
- Zhang L, Li DC, Yu SL. 2020. Pharmacological effects of harmine and its derivatives: a review. *Arch Pharm Res.* 43(12):1259–1275.