MITOCHONDRIAL DNA PART B
2022, VOL. 7, NO. 7, 1384–1386
https://doi.org/10.1080/23802359.2022.2101398

Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

OPEN ACCESS

The complete chloroplast genome and phylogenetic analysis of *Salvia karwinskii* (Lamiaceae)

Guiping Zhao[a,b], Yifei Pei[b], Dade Yu[b], Furong Xu[a] and Xiwen Li[a,b]

[a]College of Traditional Chinese Medicine, Yunnan University of Chinese Medicine, Kunming, China; [b]Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China

ABSTRACT

Salvia karwinskii Benth. 1835 is a perennial herb in the Lamiaceae family native in Mexico and Central America. The complete chloroplast (cp) genome of *S. karwinskii* was sequenced using the Illumina platform and assembled for the first time. The complete plastid genome of *S. karwinskii* was 150,907 bp in length including a large single-copy (LSC) region of 82,205 bp, a small single-copy (SSC) region of 17,538 bp, and a pair of inverted repeat (IR) regions of 25,582 bp. The total GC content of this genome was 38.05%, and that of LSC, SSC, and IR regions was 36.22%, 31.77%, and 43.14%, respectively. The cp genome contained 114 unique genes, including 80 protein-coding genes, 30 tRNA genes, and four rRNA genes. The maximum-likelihood phylogenetic tree was constructed with 38 complete cp genomes, supporting a close relationship between *S. karwinskii* and a 10 species lineage, all of which belong to the subg. *Calosphace* of *Salvia*. The cp genome of *S. karwinskii* provides a foundation for further studies on genetic diversity and improving the classification system of *Salvia*.

ARTICLE HISTORY
Received 25 May 2022
Accepted 11 July 2022

KEYWORDS
Salvia karwinskii; complete chloroplast genome; phylogenetic analysis

Salvia is the largest genus in the family Lamiaceae, containing appropriately 1000 species of shrubs, herbaceous perennials and annuals (Li et al. 2013; Cui et al. 2020). Many species of *Salvia* are popular garden plants because they typically bloom for a long period of time and grow well even in harsh conditions (Hu et al. 2020; Zhou et al. 2021). *Salvia karwinskii* Benth. 1835 is a tall and evergreen perennial shrub belonging to the genus *Salvia*. It is widely distributed throughout Mexico and Central America, including Guatemala, Honduras, El Salvador, and Nicaragua. *S. karwinskii* produces watermelon-pink blooms that form lower-loose and upper-dense racemes. It has great ornamental value, with a long blooming time from May to December every year. However, its research in genetics and evolution is extremely rare. In this study, we sequenced the chloroplast (cp) genome of *S. karwinskii* and examined its phylogenetic position within the genus *Salvia*. It is expected to lay the foundation for further breeding studies of *S. karwinskii*.

Fresh leaves of *S. karwinskii* were collected from Guatemala Botanical Garden, Guatemala (90°30′ N, 44°02′ W), and identified by Xiwen Li. The specimen was deposited at herbarium of the Institute of Chinese Materia Medica, China Academy of Chinese Medicinal Sciences, Beijing, China (http://www.icmm.ac.cn/, Xiwen Li, xwli@icmm.ac.cn) under the voucher number SZ20190920. The total genomic DNA was extracted by the modified cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987). The total DNA was used to generate a library with an average insertion size of 350 bp. The cp genome sequencing was performed on the Illumina Hiseq 1500 platform (Illumina Inc., San Diego, CA) with the paired-end 150 bp strategy. The complete cp genome of *S. miltiorrhiza* (NC020431) was used as the reference genome for extracting cp genome reads (Qian et al. 2013). It was assembled by SOAPdenovo (version 2.04) (Luo et al. 2012). The complete cp genome of *S. karwinskii* (accession number MT156372) was submitted to GenBank after being annotated by Plann (Huang and Cronk 2015).

The complete cp genome of *S. karwinskii* was 150,907 bp in length including a large single-copy (LSC) region of 82,205 bp, a small single-copy (SSC) region of 17,538 bp, and a pair of inverted repeat (IR) regions of 25,582 bp. The total GC content of this cp genome was 38.05%, while 36.22%, 31.77%, and 43.14% in the LSC, SSC, and IR regions, respectively. Besides, 114 unique genes were obtained in the cp genome, including 80 protein-coding genes, 30 tRNA, genes and four rRNA genes.

A total of 34 cp genomes of *Salvia* species were downloaded from GenBank. The complete cp genomes including both IR regions of these 34 Salvia species and *Salvia karwinskii* were used for phylogenetic analysis, together with *Mentha longifolia*, *M. canadensis* as well as *Glechoma longituba* as outgroups. After alignment with MAFFT (version 7.310) (Rozewicki et al. 2019), the maximum-likelihood (ML) tree was conducted using RAxML (version 8.2.12) (Stamatakis

CONTACT Xiwen Li xwli@icmm.ac.cn Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China; Furong Xu xfrong99@163.com College of Traditional Chinese Medicine, Yunnan University of Chinese Medicine, Kunming 650500, China

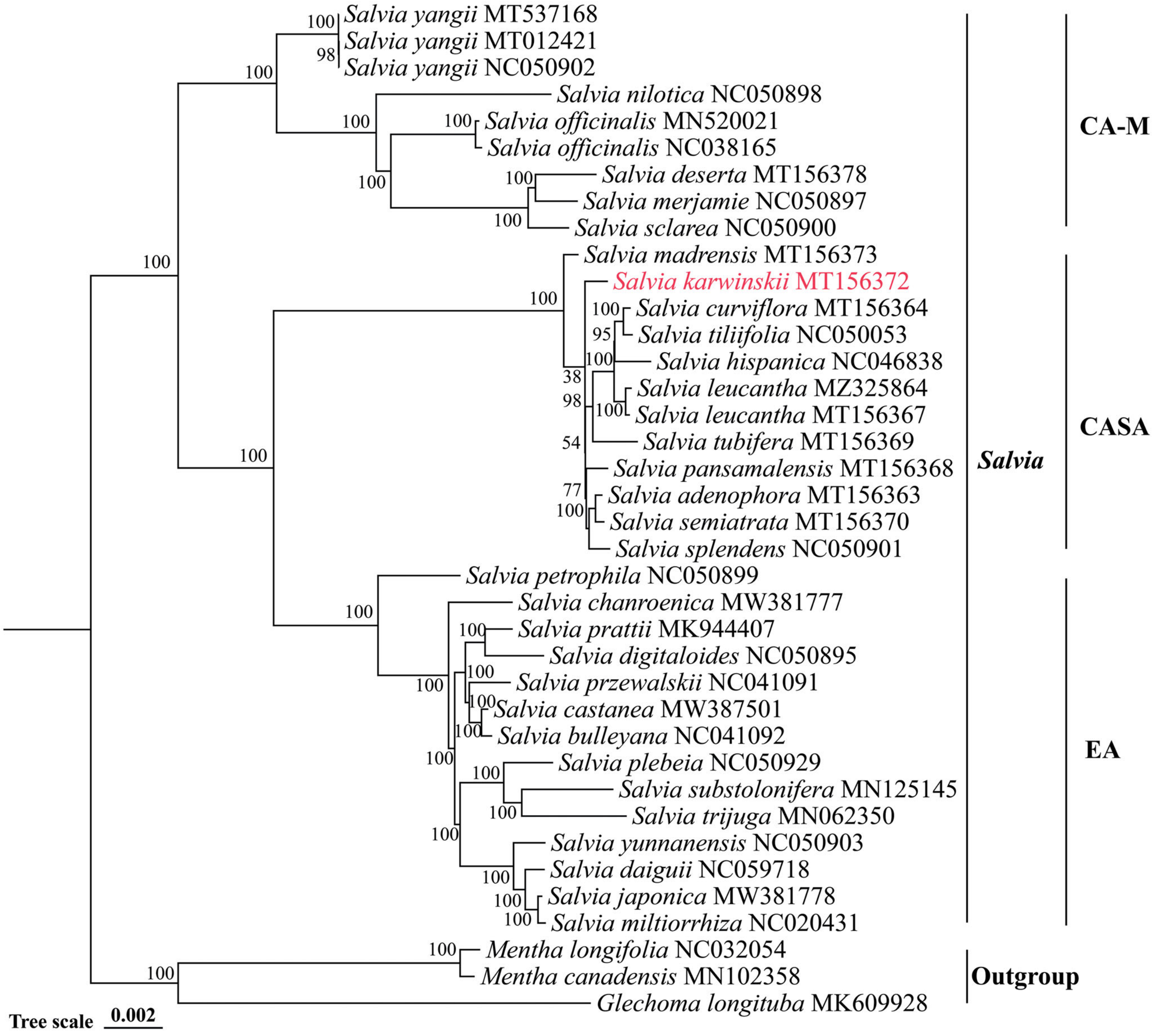

Figure 1. The maximum-likelihood (ML) tree was constructed based on complete chloroplast genome sequences of 35 *Salvia* species. *Mentha canadensis*, *M. longifolia*, and *Glechoma longituba* were used as outgroups. Bootstrap values with 1000 replicates were shown under each branch. CA-M: Central Asia-Mediterranean; CASA: Central and South America; EA: East Asia.

2014) with the GTRGAMMA model and 1000 bootstrap replicates. It was clear from the phylogenetic tree that the main nodes were supported with high bootstrap values (Figure 1). The topology based on entire cp genomes showed that *Salvia* species were monophyletic and clustered into three clades. These clades just corresponded to the three distribution centers of *Salvia*, which were Central and South America, Central Asia-Mediterranean and East Asia (Walker et al. 2004). The ML tree also showed that *S. karwinskii* was embedded within the 'Central and South America' clade. The 'Central and South America' clade could be divided into two subclades: *S. madrensis* and a small lineage including *S. karwinskii* and 10 other subg. *Calosphace* species. Our phylogenetic analysis revealed a close relationship between *S. karwinskii* and a 10 species lineage, all of which belong to the subg. *Calosphace* of *Salvia*. In conclusion, this newly assembled cp genome could serve as a foundation for further better cultivation and utilization of *S. karwinskii*.

Ethical approval

The plant specimen is not designated as endangered species. It requires no specific permissions or licenses. Collection was conducted in accordance with guidelines provided by the author's institution and national and international regulations.

Author contributions

X.W. Li and F.R. Xu designed and conceived this work; D.D. Yu collected the samples and carried out the experiment; G.P. Zhao and Y.F. Pei analyzed the data; G.P. Zhao wrote the first version of the manuscript; all authors read, revised and approved the final manuscript. All authors agree to be accountable for all aspects of the work.

Disclosure statement

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

Data availability statement

The chloroplast genome sequence that supports the findings of this study are openly available in GenBank of National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov) under the accession number MT156372. The associated BioProject, Bio-Sample, and SRA numbers are PRJNA831670, SAMN27760424, and SRR18909709, respectively.

Funding

This work was supported by the National Major Scientific and Technological Special Project for "Significant New Drugs Development" of China [2019ZX09201005-006-001] and the China Academy of Chinese Medical Sciences Innovation Fund [CI2021A03910].

ORCID

Dade Yu http://orcid.org/0000-0002-6635-881X

References

Cui N, Liao BS, Liang CL, Li SF, Zhang H, Xu J, Li XW, Chen SL. 2020. Complete chloroplast genome of *Salvia plebeia*: organization, specific barcode and phylogenetic analysis. Chin J Nat Med. 18(8):563–572.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19(1):11–15.

Hu JL, Zhao M, Hou ZJ, Shang J. 2020. The complete chloroplast genome sequence of *Salvia miltiorrhiza*, a medicinal plant for preventing and treating vascular dementia. Mitochondrial DNA Part B. 5(3): 2460–2462.

Huang DI, Cronk QC. 2015. Plann: a command-line application for annotating plastome sequences. Appl Plant Sci. 3(8):1500026.

Li Q, Li M, Yuan Q, Cui Z, Huang L, Xiao P. 2013. Phylogenetic relationships of *Salvia* (Lamiaceae) in China: evidence from DNA sequence datasets. J Syst Evol. 51(2):184–195.

Luo RB, Liu BH, Xie YL, Li ZY, Huang WH, Yuan JY, He GZ, Chen YX, Pan Q, Liu YJ, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. Gigascience. 1(1):18.

Qian J, Song JY, Gao HH, Zhu YJ, Xu J, Pang XH, Yao H, Sun C, Li X, Li CY, et al. 2013. The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. PLOS One. 8(2):e57607.

Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. 2019. MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. 47(W1):5–10.

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

Walker JB, Sytsma KJ, Treutlein J, Wink M. 2004. *Salvia* (Lamiaceae) is not monophyletic: implications for the systematics, radiation, and ecological specializations of *Salvia* and tribe Mentheae. Am J Bot. 91(7): 1115–1124.

Zhou Y, Zhang H, Ping H, Ding Y, Hu S, Bi G, Li C, Li H, Huang Y, Guo L, et al. 2021. Characterization of the complete chloroplast genome of *Salvia leucantha* (Lamiaceae). Mitochondrial DNA Part B. 6(12): 3406–3408.