

Complete chloroplast genome sequence of *Adenostemma lavenia* (Asteraceae) and phylogenetic analysis with related species

Mingze Xia^{a,b} and Yan Li^c

^aKey Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, China; ^bUniversity of Chinese Academy of Sciences, Beijing, China; ^cSchool of Pharmacy, Weifang Medical University, Weifang, China

ABSTRACT

Adenostemma lavenia is a perennial medical herb in the family Asteraceae. Here, we sequenced and analyzed the complete chloroplast genome of *A. lavenia*. The complete chloroplast genome size is 150,063 bp with a GC content of 37.63%. The *A. lavenia* chloroplast genome is a typical quadripartite structure, including a large single-copy region (LSC) of 82,017 bp and a small single-copy region (SSC) of 18,142 bp separated by a pair of inverted repeats (IRs) of 24,952 bp each. A total of 114 unique genes, including 29 tRNA genes, four rRNA genes, and 81 protein-coding genes were found in the chloroplast genome. Phylogenetic analysis revealed that *A. lavenia* is more closely related with *Chromolaena odorata*.

ARTICLE HISTORY

Received 11 February 2021
Accepted 11 June 2021

KEYWORDS

Adenostemma lavenia;
chloroplast genome;
phylogenetic analysis

Chloroplast genomes are closed circular DNA molecules in most angiosperms, which show a typically quadripartite structure of a large single-copy (LSC) region and a small single-copy (SSC) region separating by a pair of inverted repeats (IRs) (Palmer 1985). Chloroplast genomes are useful in the development of DNA barcodes for identification (Daniell et al. 2016), and can improve the resolution of phylogenetic relationships in large and complex plant lineages (Doorduyn et al. 2011; Dong et al. 2018).

Comprising ca. 25,000–35,000 species in at least 1600 genus with the distribution of all continents except Antarctica, Asteraceae is considered the largest flowering plant family (Funk et al. 2009). *Adenostemma lavenia* (L.) O. Kuntze 1891, a perennial herb of the family Asteraceae, is widely distributed in tropical and temperate region of Asia and Pacific Islands (Cheng et al. 1979; Irmanida et al. 2020). In traditionally, the whole plant has been used as a medical herb to treat bronchitis, tonsillitis, pneumonia, lung congestion, hepatitis, fever and malaria (Cheng et al. 1979; Yang et al. 2007; Irmanida et al. 2020). Although this species possesses a wide distribution range and a large number of biological resources (Cheng et al. 1979), the related study of molecular biology is rare.

Fresh leaves of *A. lavenia* was sampled from Enshi City, Hubei Province (Geographic coordinates 30°28'N, 109°49'E) and quickly dried in silica gel. A specimen and DNA were deposited at the Herbarium of Weifang Medical University (Yan Li and liyan715@mails.ucas.ac.cn) under the voucher accession number LY2020001. Total genomic DNA was extracted from approximately 10 mg of silica-dried leaf tissue by the modification of CTAB method (Doyle and Doyle 1987).

The extracted DNAs of all the individuals were then sent to Novogene (Beijing, China) for genomic library construction and Illumina sequencing. Paired-end reads of 2 × 150 bp for all samples were generated in a single lane on an Illumina HiSeq2500 sequencer (San Diego, CA, USA). We sequenced 10.21 Gb raw data, then used Trimmomatic v. 0.33 (Bolger et al. 2014) to control the quality of reads and obtained 10.14 Gb clean data. The chloroplast genome was assembled *de novo* by using GetOrganelle with the specific parameters (the assembling type: embplant_pt; the maximum number of extending rounds: 15; the K-mer length: 21, 45, 65, 85 and 105) (Jin et al. 2020) and visualized assembly result in bandage v. 0.8.1 (Wick et al. 2015). Annotation was performed through the online program GeSeq (<http://chlorobox.mpimp-golm.mpg.de/geseq.html>) (Tillich et al. 2017) and then manually adjusted for start/stop codons and intron/exon borders in SEQUIN Version 15.50 (<https://www.ncbi.nlm.nih.gov/Sequin/>) after BLAST searches. The chloroplast genome composition of *A. lavenia* was displayed using Chloroplot (Zheng et al. 2020). Sequence length of complete genomes and junction sites of LSC, IR, and SSC regions among *A. lavenia* and six other Asteraceae species were compared by using the program IRscope (Amiryousefi et al. 2018).

The complete genome size of *A. lavenia* is 150,063 bp in length, containing the large single-copy (LSC, 82,017), small single-copy (SSC, 18,142) and two inverted repeat (IR, 24,952) regions. Overall GC contents of chloroplast genomes were 38%. A total of 114 unique genes, including 29 tRNA genes, four rRNA genes, and 81 protein-coding genes were found in chloroplast genomes. In addition, comparative analysis revealed that the chloroplast genome size of *A. lavenia* is the

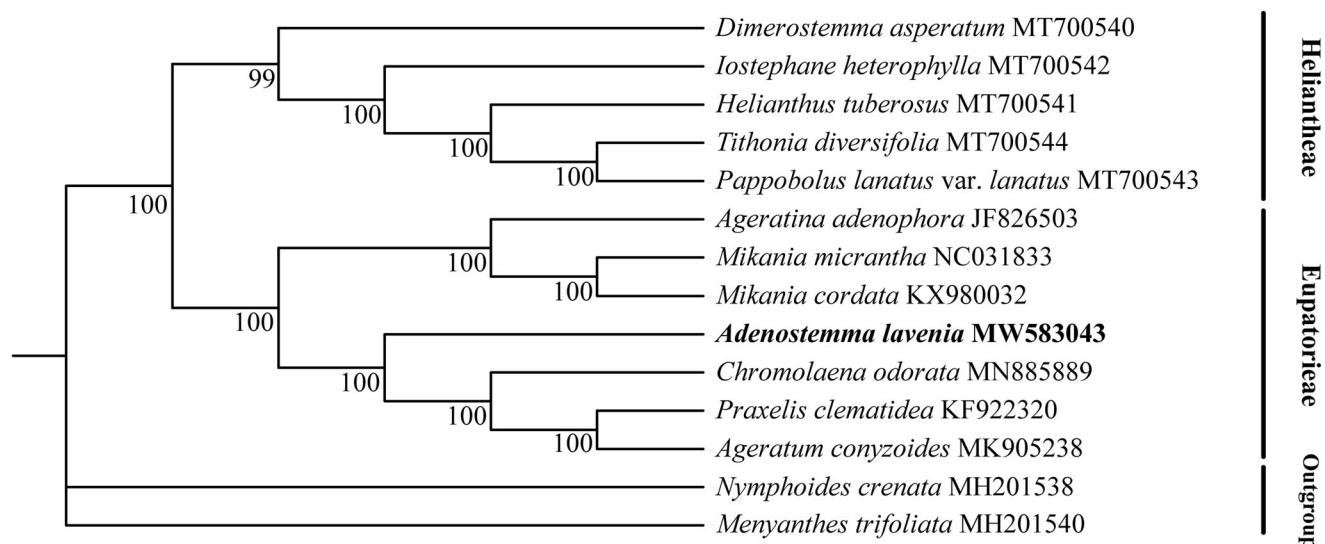


Figure 1. Bayesian inference tree based on the protein sequences of 14 chloroplast genomes. Numbers on the branches are posterior probabilities.

smallest among the compared species of Asteraceae, but the organization and gene content is highly similar among species except *H. tuberosus*. Phylogenetic analysis was performed on the chloroplast genomes sequences of 14 species (including 7 species of Eupatorieae, 5 species of Heliantheae, and 2 species as outgroup of Menyanthaceae). Protein sequences of each species were selected as dataset for phylogenetics analysis, and then compared with MAFFT (Kato and Standley 2013). Bayesian inference (BI) was performed by PhyloSuite v1.2.2 (Zhang et al. 2020). The Monte Carlo Markov chains (MCMC, one cold chain and three hot chains) analysis was run for 1,000,000 generations and trees were sampled every 1000 generations. The result showed that *A. lavenia* is closely related with *Chromolaena odorata*, *Ageratum conyzoides* and *Praxelis clematidea*, and all selected species of Eupatorieae and Heliantheae form a well-supported clade, respectively (Figure 1), supporting phylogenetic relationships of Asteraceae (Fu et al. 2016). This study will be helpful for future studies on chloroplast genomes and phylogeny of *Adenostemma*.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Doctoral Research Initiation Fund of Weifang Medical University under Grant [041064].

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI (<https://www.ncbi.nlm.nih.gov/>) under the accession no. MW583043. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA732125, SRR14626653, and SAMN19312750, respectively.

References

- Amiryousefi A, Hyvönen J, Poccai P. 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics*. 34(17):3030–3031.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15):2114–2120.
- Cheng PC, Charles DH, Norman JD. 1979. Isolation of 11-hydroxyated kauranic acids from *Adenostemma lavenia*. *J Nat Prod*. 42(2):183–186.
- Daniell H, Lin CS, Yu M, Chang WJ. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol*. 17(1):134.
- Dong WP, Xu C, Wu P, Cheng T, Yu J, Zhou SL, Hong DY. 2018. Resolving the systematic positions of enigmatic taxa: manipulating the chloroplast genome data of Saxifragales. *Mol Phylogenet Evol*. 126:321–330.
- Doorduyn L, Gravendeel B, Lammers Y, Ariyurek Y, Chin-A-Woeng T, Vrieling K. 2011. The complete chloroplast genome of 17 individuals of pest species *Jacobaea vulgaris*: SNPs, microsatellites and barcoding markers for population and phylogenetic studies. *DNA Res*. 18(2): 93–105.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19:11–15.
- Fu ZX, Jiao BH, Nie B, Zhang GJ, Gao TG, China PC. 2016. A comprehensive generic-level phylogeny of the sunflower family: implications for the systematics of Chinese Asteraceae. *J Sytemat Evol*. 54(4):416–437.
- Funk VA, Susanna A, Steussy TF, Robinson HE. 2009. Classification of compositae. In: Funk VA, Susanna A, Steussy TF, Bayer RJ, editors. *Systematics, evolution, and biogeography of compositae*. Vienna: International Association for Plant Taxonomy; p. 171–189.
- Irmanida B, Rika IA, Muhammad EP, Auliya I, Miwa M, Mayu S, Akie H, Hiroshi T. 2020. The antiaging effect of active fractions and ent-11 α -hydroxy-15-oxo-kaur-16-en-19-oic acid isolated from *Adenostemma lavenia* (L.) O. Kuntze at the cellular level. *Antioxidants*. 9(8):719.
- Jin JJ, Yu WB, Yang JB, Song Y, Claude WD, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol*. 21(1):1–31.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30 (4):772–780.
- Palmer JD. 1985. Comparative organization of chloroplast genomes. *Annu Rev Genet*. 19:325–354.
- 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):W6–11.
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics*. 31(20): 3350–3352.

- Yang YL, Guo SJ, Ma RJ, Wang ZL. 2007. Chemical composition of the volatile oil in *Adenostemma lavenia* (L.) O. Ktze. *J Trop Subtrop Bot.* 15(4):355–358.
- Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol Ecol Resour.* 20(1):348–355.
- Zheng SY, Poczai P, Hyvönen J, Tang J, Amiryousefi A. 2020. Chloroplast: an online program for the versatile plotting of organelle genomes. *Front Genet.* 11:576124.