

Characterization of the complete mitochondrial genome of *Ageratum conyzoides*

Ze-Ping Luo and Li-Wei Pan

College of Chemistry and Bio-engineering, Hechi University, Hechi, PR China

ABSTRACT

As a pharmaceutical plant with multi-bioactivity, *Ageratum conyzoides* appears to be a valuable agricultural resource. In this study, the complete mitochondrial (mt) genome of *A. conyzoides* was sequenced through Illumina sequencing method, and the mt genome was recovered after *de novo* assembly and annotation. The results showed that the 219,198 bp mt genome has a total of 52 genes, including 30 protein-coding genes, 3 rRNA genes and 19 tRNA genes. The overall GC content of this mitogenome is 45.4%. By phylogenetic analysis using maximum-likelihood (ML) method, *A. conyzoides* showed the closest relationship with *Diplostephium hartwegii* in the family of Asteroideae.

ARTICLE HISTORY

Received 5 July 2019
Accepted 3 August 2019

KEYWORDS

Ageratum conyzoides;
mitochondrial genome;
Illumina sequencing;
phylogenetic analysis

Ageratum conyzoides L., Asteraceae, is an annual herbaceous plant with a long history of traditional medicinal uses in several countries of the world and also has bioactivity with insecticidal and nematocidal activity (Ming 1999). In this study, we finished and analysed the mitochondrial (mt) genome of *A. conyzoides* based on the next-generation sequencing method.

Plant materials of *A. conyzoides* sequenced in this study were acquired from medical plants garden in Guiyang University of Traditional Chinese Medicine (26°57'N, 106°72'W). This specimen and its total genomic DNA and were stored in the Key Laboratory of Miao Medicine, Guiyang University of Traditional Chinese Medicine with accession NO. AGC-20170608-265710672. Sequencing was done on an Illumina HiSeq2500 platform. The clean reads were assembled by SPAdes version 3.11.1 (Bankevich et al. 2012) with default settings.

We used four mt genomes from Asteroideae plants as seeds to collect assembled mt fragments that facilitated by using Exonerate alignment (Slater and Birney 2005) through comparison between published protein-coding genes and the total assembled contigs. To fill the gap, Price (Ruby et al. 2013) and MITObim version 1.8 (Hahn et al. 2013) were applied and Bandage (Wick Ryan et al. 2015) was used to construct the circular assembly path guided by these positions of long scaffolds. The complete sequence was primarily annotated by Plann (Huang and Cronk 2015) and Exonerate combined with manual correction. All tRNAs were confirmed using the tRNAscan-SE search server (Lowe and Eddy 1997). Other protein-coding genes were verified by BLAST search on the NCBI website, and manual correction for start and stop

codons was conducted. This complete mt genome sequence was submitted to GenBank under the accession numbers of MN075945.

The mt genome of *A. conyzoides* is a typical circular structure of 219,198 bp with a GC content of 45.4%. This mt genome contains a total of 52 genes, including 30 protein-coding genes, three rRNA genes (*rrn5*, *rrn18*, and *rrn26*), 19 complete native mt tRNA genes.

Another 15 published complete mt sequences in asterids were collected from the Genbank database. Whole genome-wide alignments using HomBlocks (Bi et al. 2018) under Gblocks trimming method (Dereeper et al. 2008), resulting in 25,724 characters in total, including almost all whole or partial PCGs and rRNA genes. ModelFinder (Kalyanamoorthy et al. 2017) was used to identify the most appropriate substitution model. Phylogenetic analyses of the concatenated sequences were performed using the maximum-likelihood (ML) method implemented in IQ-TREE version 1.6.6 (Nguyen et al. 2014) according to the optimal substitution models under the rapid bootstrap algorithm (1000 replicates). To test support for the branch points of each gene trees, nonparametric branch support tests based on the Shimodaira–Hasegawa-like approximate likelihood ratio test (SH-like aLRT) procedure were also performed in the same run. As shown in Figure 1, the phylogenetic positions of these 16 mt genomes were successfully resolved with high bootstrap supports across almost all nodes. *A. conyzoides*, belonging to the Eupatorieae, exhibited the closest relationship with *Diplostephium hartwegii* in the family of Asteroideae.

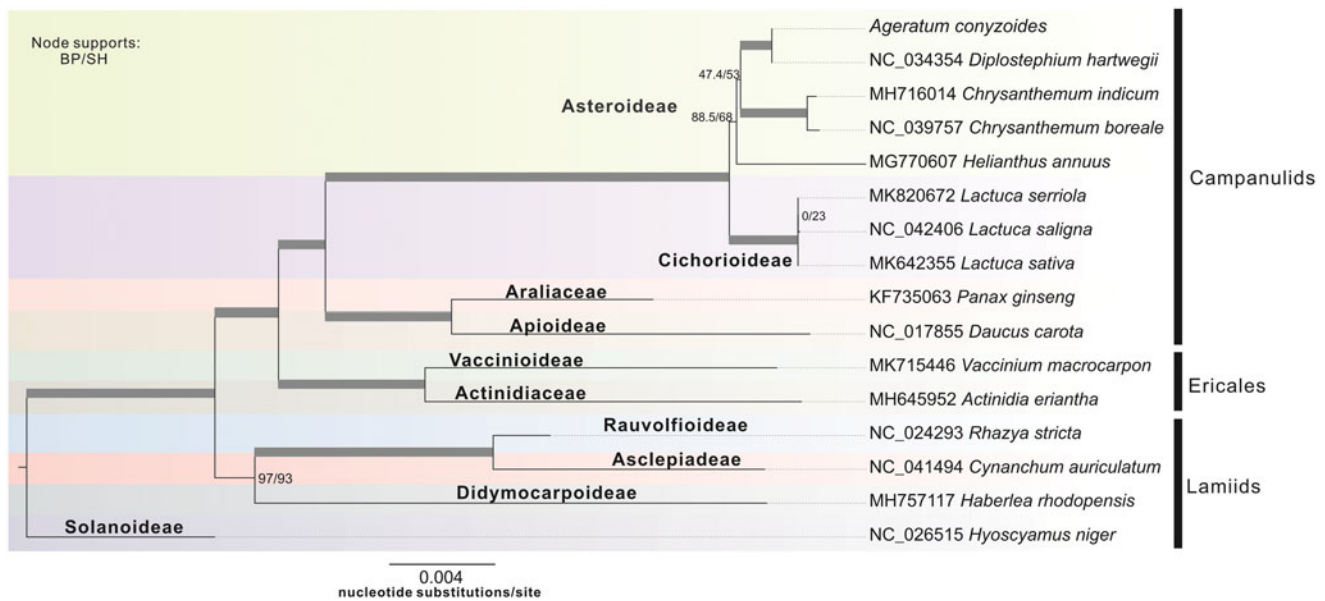


Figure 1. Phylogenetic tree yielded by ML analysis of 16 higher plant mt genomes. ML consensus tree is shown with bootstrap supports indicated by numbers besides branches. Branches harvesting both full bootstrap and SH-like aLRT values are indicated in bold. The scale indicates substitutions per site.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Bi G, Mao Y, Xing Q, Cao M. 2018. HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching. *Genomics.* 110(1):18–22.
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, et al. 2008. Phylogeny. fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36: W465–W469.
- Hahn C, Bachmann L, Chevreaux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* 41(13): e129–e129.
- Huang DI, Cronk QC. 2015. Plann: a command-line application for annotating plastome sequences. *Appl Plant Sci.* 3(8):1500026.
- Kalyaanamoorthy S, Minh BQ, Wong Thomas KF, Haeseler A. V, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 14(6):587.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25(5): 955–964.
- Ming LC. 1999. *Ageratum conyzoides*: a tropical source of medicinal and agricultural products. In Janick J, editor. *Perspectives on New Crops and New Uses.* Alexandria: ASHS Press; p. 469–473.
- Nguyen LT, Schmidt HA, Haeseler AV, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Ruby JG, Bellare P, DeRisi JL. 2013. PRICE: software for the targeted assembly of components of (Meta) genomic sequence data. *G3 (Bethesda).* 3(5):865–880.
- Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics.* 6(1):31.
- Wick Ryan R, Schultz Mark B, Schultz J, Holt Kathryn E. 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics.* 31(20):3350–3352.