Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

The complete chloroplast genome sequence of *Aspidopterys concava* (Wall.) A. Juss. (Malpighiaceae)

Songdian Huang[a], Shinan Liu[b], Xiaochun Liang[a] and Yonghua Qin[a,c]

[a]Guangxi Nanning Arboretum, Nanning, China; [b]Forestry College, Guangxi University, Nanning, China; [c]Guangxi Forestry Inventory and Planning Institute, Nanning, China

ABSTRACT

Aspidopterys concava is related to a group of important medicinal plants in Malpighiaceae in southeast Asia. Here, we report the first chloroplast genome fully sequenced and annotated for *Aspidopterys concava*. The genome size was 160,441 bp and contained a large single-copy (LSC) region of 71,434 bp, a small single-copy (SSC) region of 53,544 bp, and a pair of inverted repeats (IRs) regions of 8943 bp. Total GC content was 37.9%. It contained 125 genes in total, comprising 82 protein-coding genes, 37 transfer RNA genes, and six ribosomal RNA genes. Phylogenetic analysis showed that *A. concava* was the most closely related to *A. obcordata* from the same genus.

Aspidopterys concava (Wall.) A. Juss. 1840 is a species in family Malpighiaceae distributed across China (Guangxi), the Indo-China Peninsula, Malaysia, Indonesia, and the Philippines (Qin 2010). This woody liana is well adapted to limestone substrates in harsh mountain terrain. According to the survey, residents of the Hechi and Chongzuo areas from Guangxi province (China) eat *A. concava* as a wild vegetable, underscoring its ecological and economic value. Moreover, a related species *A. obcordata* is used in Dai traditional medicine (Li et al. 2019; Gong et al. 2020), making *A. concava* an excellent candidate for bioprospecting for bioactive compounds. Here, we report the first fully assembled and annotated complete chloroplast genome for *A. concava*.

Fresh young leaves of *A. concava* were collected from seedlings growing in the Nanning Arboretum, Guangxi, China (108°17′30.46″ E longitude, 22°43′17.42″ N latitude). Tissues were immediately frozen in liquid nitrogen and stored at −80 °C before DNA extraction with a modified CTAB method (Doyle and Doyle 1987). A specimen was deposited at the herbarium of Guangxi Forestry Inventory and Planning Institute (Mr. Qin, gx_qinyonghua@126.com) under the voucher number Qin 200421005. A library was generated using TIANSeq Fast DNA Library Prep Kit NG102 (Illumina, Inc., San Diego, CA) and sequenced on the Illumina HiSeq TM 4000 Platform using paired end runs with 2 × 150 cycles. Briefly, genomic DNA was sonicated to ∼350 bp fragments. Data QC was conducted using Fastp (version 0.21.0) to parse 21.6 Gb clean data. The 144,832,380 reads at 150 bp length were *de novo* assembled using SPAdes 3.6.1 with K-mer sizes of 55, 87,

and 121 (Bankevich et al. 2012). *De novo* assembled chloroplast contigs were concatenated into larger contigs using Sequencher 5.3.2 (Gene Codes Inc., Ann Arbor, MI). Automatic annotation of the chloroplast genomes was generated by CpGAVAS and a circular representation of both sequences was drawn using the online tool OGDR (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html) (Lohse et al. 2007; Liu et al. 2012).

The chloroplast genome of *A. concava* is 160,441 bp in length with a typical quadripartite structure. It contains a large single-copy (LSC, 71,434 bp) region, a small single-copy (SSC, 53,544 bp) region, and a pair of inverted repeats (IRs, 8943 bp) regions. The GC-content of the whole cp genomes is 37.9%, and in LSC, SSC, and IR regions are 39.3%, 39.8%, and 38.0%, respectively. The cpDNA contains 125 genes including 82 protein-coding genes (PCGs), 37 transfer RNA genes (tRNAs), and six ribosomal RNA genes (rRNAs).

To determine the phylogenetic location of *A. concava*, it was aligned with 14 other representative species of Malpighiales using MAFFT v7 (Katoh and Standley 2013). A phylogenetic reconstruction was estimated with a maximum-likelihood (ML) analysis of a whole cp genome dataset using RAxML-HPC BlackBox v.8.1.24 with 1000 bootstrap replicates at the CIPRES Science Gateway (Stamatakis et al. 2008). The results showed that *A. concava* and the medicinal *A. obcordata* are strongly supported as sister in a clade containing *Banisteriopsis caapi* (Figure 1). The plastome of *A. concava* was similar to that of *A. obcordata* in length, but the latter had fewer genes including 79 protein-coding genes, 30 transfer RNAs, and four ribosomal RNAs.
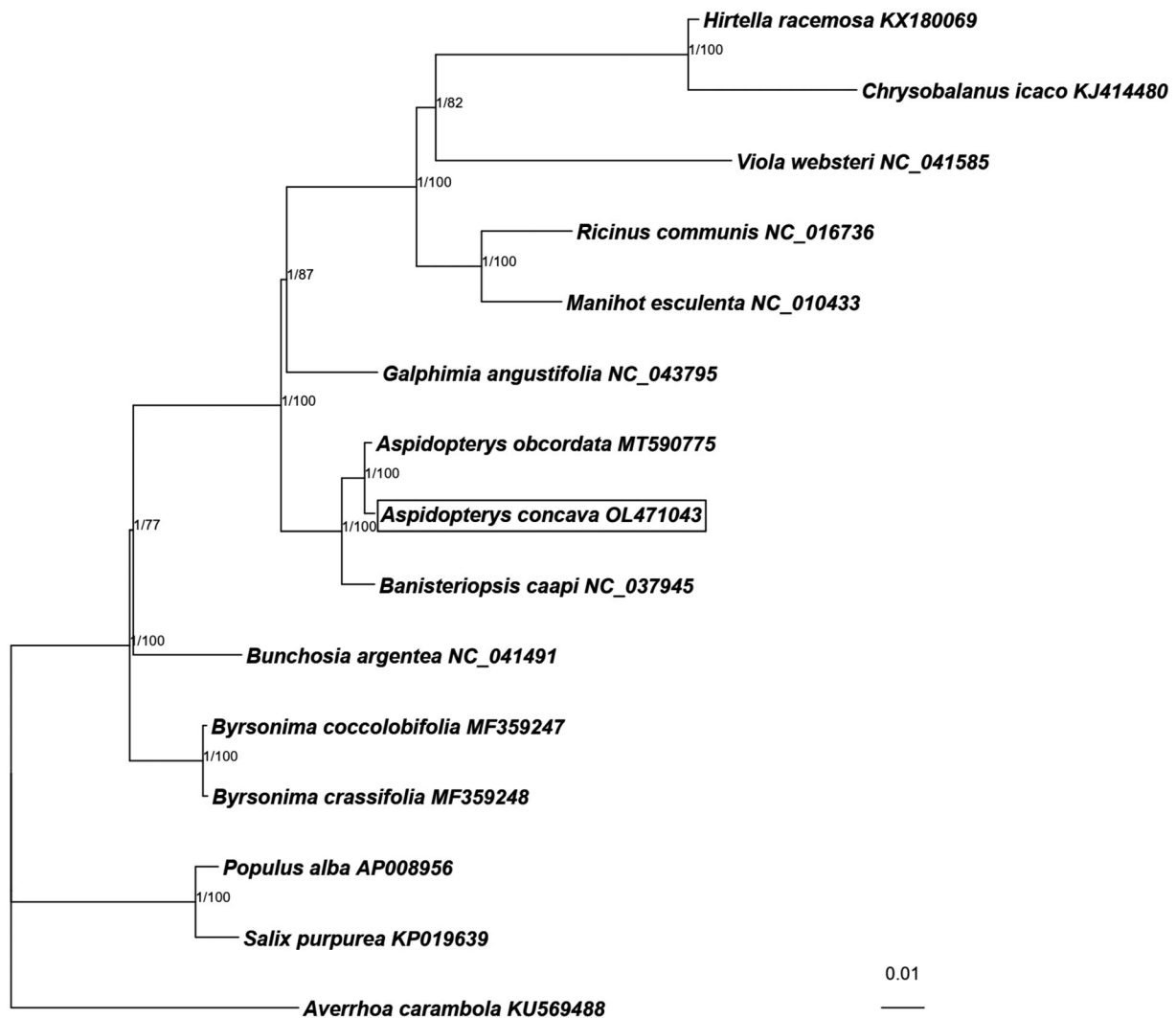
Figure 1. A maximum-likelihood tree reconstructed using 15 complete chloroplast genomes. Bootstrap support values ($N = 1000$) are shown at each node and the scale bar indicates the number of nucleotide substitutions per site. GenBank accession number used are: *Aspidopterys concava* (OL471043, in this study), *Aspidopterys obcordata* (MT590775), *Viola websteri* (NC_041585), *Populus alba* (AP008956), *Ricinus communis* (NC_016736), *Manihot esculenta* (NC_010433), *Averrhoa carambola* (KU569488), *Hirtella racemosa* (KX180069), *Chrysobalanus icaco* (KJ414480), *Byrsonima coccolobifolia* (MF359247), *Byrsonima crassifolia* (MF359248), *Bunchosia argentea* (NC_041491), *Galphimia angustifolia* (NC_043795), *Banisteriopsis caapi* (NC_037945), and *Salix purpurea* (KP019639).

Acknowledgements

Author contributions

Songdian Huang was involved in the conception, design and data analysis of the work and write the original draft. Shinan Liu and Xiaochun Liang were involved in analysis and interpretation of the data. Yonghua Qin was involved in the experiment conception or design, data analysis of the work and revising the paper. All authors have read and agreed to the published version of the manuscript.

Disclosure statement

No potential conflict interest was reported by the authors.

Funding

Data availability statement

Chloroplast genome sequence can be accessed via accession number OL471043 in NCBI GenBank. The associated BioProject and BioSample are PRJNA799145 and SAMN25145001, respectively, while three SRA numbers are SRR17798691, SRR17798692, and SRR17798693.

References

Bankevich A, Nurk S, Antipov D, Gurevich A-A, Dvorkin M, Kulikov A-S, Lesin V-M, Nikolenko S-I, Pham S, Prjibelski A-D, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.

Doyle JJ, Doyle J-L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19:11–15.

Gong Y-X, Jiang N, Zhang Z-R, Tian Y-H, Yu W-B. 2020. Complete plastome sequence of *Aspidopterys obcordata* (Malpighiaceae), a traditional Dai medicinal plant from Xishuangbanna, Yunnan, China. Mitochondrial DNA B Resour. 5(3):3533–3535.

Katoh K, Standley D-M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Li Y, Ma G, Lv Y, Su J, Li G, Chen X. 2019. Efficacy of Obcordata A from *Aspidopterys obcordata* on kidney stones by inhibiting NOX4 expression. Molecules. 24(10):1957.

Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet. 52(5–6): 267–274.

Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. 2012. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and Genbank submission of completely sequenced chloroplast genome sequences. BMC Genomics. 13:715.

Qin HN. 2010. China checklist of higher plants. In the Biodiversity Committee of Chinese Academy of Sciences, editor. Catalogue of Life China. 2010 Annual Checklist China. CD-ROM; Science Press, Beijing, China.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst Biol. 57(5):758–771.