

The complete chloroplast genome sequence of chaya (*Cnidoscolus aconitifolius*) and phylogenetic analysis

Yufeng Qin^a, Mimi Li^{b,c} , Lianxiang Zhong^a, Qiulan Wei^a, Yufei Xiao^a, Xiaoning Zhang^a  and Zihai Qin^a

^aGuangxi Forestry Research Institute, Nanning, China; ^bInstitute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China; ^cThe Jiangsu Provincial Platform for Conservation and Utilization of Agricultural Germplasm, Nanjing, China

ABSTRACT

Cnidoscolus aconitifolius is a leafy green healthy vegetable and medicinal plant belongs to the family Euphorbiaceae. In the present study, we sequenced the complete chloroplast genome of *C. aconitifolius*, which is 158,658 bp in length and consisted of two copies of inverted repeat (IR) of 26,982 bp separated by a large single copy (LSC) of 87,022 bp and a small single copy (SSC) of 17,672 bp. The GC content of *C. aconitifolius* was 36.3%. A total of 130 genes were predicted, including 86 protein-coding genes, 36 tRNAs and 8 rRNAs. The plastid phylogenomic analysis support *C. aconitifolius* is closely related to *Manihot esculenta*.

ARTICLE HISTORY

Received 7 July 2021
Accepted 14 November 2021

KEYWORDS

Euphorbiaceae; chaya;
chloroplast genome;
Cnidoscolus aconitifolius; plastid

Cnidoscolus aconitifolius (Mill.) Johnst. (family Euphorbiaceae, subfamily Crotonoideae, tribe Manihoteae), generally known as 'chaya' or tree spinach, is a perennial shrub species (Breckon 1979). It is naturally distributed in Mexico and Central America and recently widely introduced into the USA, Europe and China for potential use as a leafy green healthy vegetable and medicinal plant due to its edible and therapeutic impacts (Kuti and Kuti 1999). Chaya has been reported to contain protein, vitamins, calcium, iron and with medicinal benefits especially linked to diabetes mellitus, atherosclerosis, hypertension (Kuti and Torres 1996). However, genetic and genomic information on chaya is scarce. The plastid genomic information reported in this study is valuable for genetic improvement through genetic engineering in chaya.

The fresh mature and healthy leaves of *C. aconitifolius* were sampled at Guangxi Forestry Research Institute, Guangxi Province, China (22°55'30"N, 108°21' E) and voucher specimen was saved at Guangxi Forestry Research Institute with collection number 20210317003 (contact: Zihai Qin, 75455621@qq.com). Total genomic DNA was extracted from leaves by modified hexadecyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987). A DNA library was prepared using NEB Next[®] Ultra DNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's instructions with an insert size of 300 bp and subsequently sequenced on a Illumina HiSeq X-ten platform (San Diego, USA) at Novogene Co., Ltd. (Beijing, China). The Paired-end (PE) reads were directly assembled into plastid genome using NOVOPlasty 4.3.1 (Dierckxsens et al. 2017) with default parameters using *Manihot esculenta* (NC_010433) as a reference. The genome annotation was performed by GeSeq

(Tillich et al. 2017) and adjusted by manual in Geneious 11.1.5 (Kearse et al. 2012). The complete plastid sequence was deposited to the National Center for Biotechnology Information (NCBI) under accession number MZ045411.

The plastid genome of *C. aconitifolius* was 158,658 bp in length and consisted of two copies of IR of 26,982 bp separated by a LSC of 87,022 bp and a SSC of 17,672 bp. The total GC content of *C. aconitifolius* was 36.3%, of which was 33.9%, 30.3%, and 42.2% in LSC, SSC and IRs regions, respectively. A total of 130 genes were predicted, including 86 protein-coding genes (PCGs), 36 transfer RNA genes (tRNA) and 8 ribosomal RNA genes (rRNA). Seven PCGs (*rpoC1*, *ndhB*, *ndhA*, *petD*, *rps16*, *petB*, *rpl2*) and five tRNA (*trnK-UUU*, *trnI-GAU*, *trnA-UGC*, *TrnV-UAC*, *trnL-UAA*) had one intron, while two PCGs (*clpP* and *ycf3*) had two introns.

A total of 12 plastomes were used for the phylogenomic analysis. Two species of *Euphorbia* – *E. lathyris* (MT241376) and *E. smithii* (MN646684) – were defined as outgroups. The complete plastid sequences were aligned using MAFFT 7.409 (Katoh and Standley 2013). The plastid phylogenomic analysis was generated by maximum likelihood (ML) in RAxML with 1000 bootstrap replicates (Stamatakis 2014) under the GTR + G model. Our phylogenetic analyses revealed that the tribe Manihoteae is a monophyletic clade (bootstrap support 100%) that is sister to the tribe Micrandreae (Figure 1). The result is similar with the previous work reported by Wurdack et al. (2005) based on the plastid *rbcL* and *trnL-F* sequences indicating that the cp genome is a powerful tool to resolve the deep phylogenetic relationships within the family Euphorbiaceae. Within the tribe Manihoteae, *C. aconitifolius* was most closely related to *Manihot esculenta*.

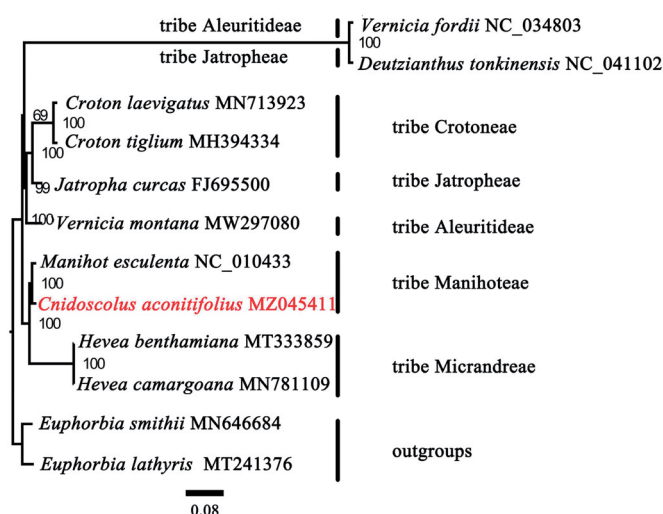


Figure 1. Maximum likelihood phylogenetic tree of *Cnidoscopus aconitifolius* with 1000 bootstrap replicates.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was financial supported by the Open program of Guangxi Key Laboratory of Traditional Chinese Medicine Quality Standards [Gui Zhong Zhong Kai 202002].

ORCID

Mimi Li <http://orcid.org/0000-0002-7883-0325>

Xiaoning Zhang <http://orcid.org/0000-0003-2105-0560>

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/>) under the accession no. MZ045411. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA743565, SRR15036486 and SAMN20055435, respectively.

References

- Breckon G. 1979. Studies in *Cnidoscopus* (Euphorbiaceae), I: *Jatropha tubulosa*, *Jatropha Liebmanni* and allied taxa from central Mexico. *Brittonia*. 31(1):125–148.
- Dierckx N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res*. 45(4):e18.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19:11–15.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28: 1647–1649.
- Kuti J, Kuti H. 1999. Proximate composition and mineral content of two edible species of *Cnidoscopus* (tree spinach). *Plant Food Human Nutr*. 54:275–283.
- Kuti J, Torres E. 1996. Potential nutritional and health benefits of tree spinach. In Janick J, editor. *Progress in new crops*. 3rd series. Arlington (VA): ASHS Press.
- Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9): 1312–1313.
- 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–W11.
- Wurdack KJ, Hoffmann P, Chase MW. 2005. Molecular phylogenetic analysis of uniovulate Euphorbiaceae (Euphorbiaceae sensu stricto) using plastid *rbcl* and *trnL-F* DNA sequences. *Am J Bot*. 92:1397–1420.