



The complete chloroplast genome of *Agave amaniensis* (Asparagales: Asparagaceae: Agavoideae)

Bochao Xu^a, Shibe Tan^b, Xu Qin^c, Xing Huang^b , Jingen Xi^b, Helong Chen^b, Jianfeng Qin^c, Tao Chen^c and Kexian Yi^{b,d,e} 

^aSchool of Life Sciences, Hainan University, Haikou, PR China; ^bEnvironment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, PR China; ^cGuangxi Subtropical Crops Research Institute, Nanning, PR China; ^dMinistry of Agriculture and Rural Affairs, Key Laboratory of Integrated Pest Management on Tropical Crops, Haikou, PR China; ^eHainan Key Laboratory for Monitoring and Control of Tropical Agricultural Pests, Haikou, PR China

ABSTRACT

Agave amaniensis Trel. & W. Nowell (1933) has long been used for phytosteroid production, which is also one of the parents of the famous *Agave* hybrid cultivar 11648 for sisal fiber production. However, its systematic position and phylogenetic relationship remains unknown at the chloroplast (cp) genome level. Therefore, we have sequenced and assembled the cp genome of *A. amaniensis* via Illumina sequencing. The cp genome is 157,282 bp in length with a GC content of 37.84%. A large single-copy region of 85,899 bp, a small single-copy region of 18,233 bp, and inverted repeat regions of 26,575 bp were found in the cp genome. Based on the annotation, 86 protein-coding genes, eight rRNAs, and 38 tRNAs were identified in the cp genome with total lengths of 78,981 bp, 9050 bp, and 2867 bp, respectively. The phylogenetic tree indicates that *A. amaniensis* is closely related with *A. H11648*, *A. angustifolia*, and *A. americana*.

ARTICLE HISTORY

Received 10 May 2022
Accepted 1 August 2022

KEYWORDS

Agave amaniensis;
chloroplast genome;
phylogenetic tree

Background

Agave plants are widely cultivated in the tropical areas of the world for food, beverage, fiber, and medicine production (Huang et al. 2019). Among the 166 cultivated *Agave* species, *Agave amaniensis* Trel. & W. Nowell (1933) has long been used for phytosteroid production (Indrayanto et al. 1993; Gil-Vega et al. 2006). This species also serves as one of the parents of the famous *Agave* hybrid cultivar 11648, which is cultivated for sisal fiber production worldwide (Huang et al. 2018). However, the systematic position and phylogenetic relationship of *A. amaniensis* remains unknown at the chloroplast (cp) genome level. Therefore, we have sequenced and assembled the cp genome of *A. amaniensis* via Illumina sequencing to facilitate future studies on *Agave* cps.







Methods

The leaves of *A. amaniensis* were collected from a two-year-old plant grown in the germplasm garden (22.90°N, 108.33°E) of the Guangxi Subtropical Crops Research Institute, Nanning, China. Several leaves were processed as specimens and stored in the Herbarium of Environment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences (voucher no. EPPI-jm2020012,

<https://eppi.catas.cn/>, Xing Huang, hxalong@gmail.com). Further, DNA was extracted from the remaining leaves using the modified CTAB method and stored at -80°C until submission to Biozeron Biotech (Shanghai, China) for sequencing (Doyle and Doyle 1987). Paired-end sequencing was performed using Illumina HiSeq 2500 (San Diego, CA). The raw data obtained were used for cp genome assembly using the NOVOPlasty software, followed by gap-filling using GapCloser (Luo et al. 2012; Dierckxsens et al. 2017). The complete cp genome was annotated using GeSeq and CPGAVAS2 (Tillich et al. 2017; Shi et al. 2019). Thereafter, the nucleotide sequences of protein-coding genes were extracted from the cp genome sequence. The merged protein-coding sequence was further aligned with those of other species using the MAFFT software (Kato and Standley 2013). Sequence alignment was imported into the MEGA7 software to construct a maximum likelihood phylogenetic tree with 1000 bootstrap replicates (Kumar et al. 2016).

Results

In total, 7.2 Gb of raw data were generated from Illumina sequencing, which were deposited to the SRA database (accession no. PRJNA705737). The assembled cp genome of *A. amaniensis* was submitted to the GenBank database

CONTACT Xing Huang  hxalong@gmail.com  Environment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, PR China; Tao Chen  15607718198@wo.cn  Guangxi Subtropical Crops Research Institute, Nanning, PR China; Kexian Yi  yikexian@126.com  Environment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, PR China

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

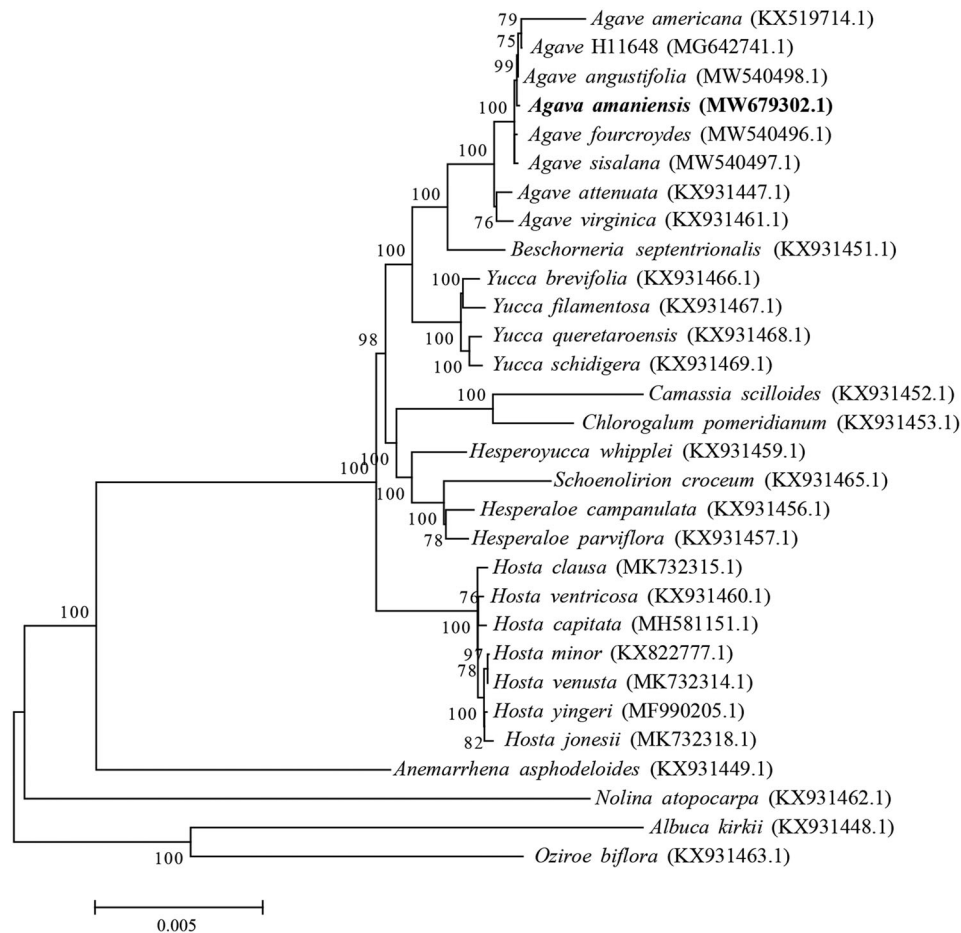


Figure 1. The maximum-likelihood phylogenetic tree of 30 species, including 27 Agavoideae species and three other species (*Albuca kirkii*, *Nolina atopocarpa*, and *Oziroë biflora*) as outgroup. The nucleotide sequences of protein-coding genes were extracted and merged from each cp genome for sequence alignment.

(accession no. MW679302). The cp genome is 157,282 bp in length with a GC content of 37.84%. A large single-copy region of 85,899 bp, a small single-copy region of 18,233 bp, and inverted repeat regions of 26,575 bp were found. Based on the genome annotation, 86 protein-coding genes, eight rRNAs, and 38 tRNAs were identified in the cp genome with total lengths of 78,981 bp, 9050 bp, and 2867 bp, respectively.

The cp genome sequences of 30 species, 27 Agavoideae species, and three other species (*Albuca kirkii*, *Nolina atopocarpa*, and *Oziroë biflora*) as outgroup were utilized for the construction of the phylogenetic tree (Qin et al. 2021). The results indicate that *A. amaniensis* is closely related with *A. H11648*, *A. angustifolia*, and *A. americana* (Figure 1).

Ethics statement and consent form

The study involved only a cultivated crop without any threatened/endangered species. It was exempted from ethical approval and did not need any permission to carry it out.

Author contributions

Xing Huang, Tao Chen, and Kexian Yi conceived and designed the experiments. Bochao Xu, Shibe Tan, and Xu Qin analyzed the data and drafted the manuscript. Bochao Xu, Shibe Tan, Xu Qin, Jingen Xi, Helong Chen, Jianfeng Qin, and Tao Chen carried out the experiments, including the species identification and sample preparation. Xing Huang, Tao Chen, and

Kexian Yi revised the manuscript. All authors approved the final manuscript and agreed to be accountable for all aspects of the work.

Disclosure statement

The authors declare that they have no competing interests.

Funding

This research was supported by Hainan Provincial Natural Science Foundation of China [322MS112], China Agriculture Research System of MOF and MARA [CARS-16], Central Public-interest Scientific Institution Basal Research Fund for Chinese Academy of Tropical Agricultural Sciences [1630042022005], Guangxi Key R&D Program [GuiKe-AB20297007], Fundamental Research Fund of Guangxi Academy of Agricultural Sciences [Guinongke-2021YT153], and the Innovation Platform for Academicians of Hainan Province.

ORCID

Xing Huang  <http://orcid.org/0000-0003-3950-0080>
Kexian Yi  <http://orcid.org/0000-0003-0496-3157>

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/nucleotide/> under the accession MW679302. The accession numbers of

BioProject, SRA, and Bio-Sample are PRJNA705737, SRS8376943, and SAMN18100000, respectively.

References

- Dierckxsens 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 from small quantities of fresh leaf tissues. *Phytochem Bull.* 19:11–15.
- Gil-Vega K, Diaz C, Nava-Cedillo A, Simpson J. 2006. AFLP analysis of *Agave tequilana* varieties. *Plant Sci.* 170(4):904–909.
- Huang X, Wang B, Xi J, Zhang Y, He C, Zheng J, Gao J, Chen H, Zhang S, Wu W, et al. 2018. Transcriptome comparison reveals distinct selection patterns in domesticated and wild *Agave* species, the important CAM plants. *Int J Genom.* 2018:5716518.
- Huang X, Xiao M, Xi J, He C, Zheng J, Chen H, Gao J, Zhang S, Wu W, Liang Y, et al. 2019. De novo transcriptome assembly of *Agave* H11648 by Illumina sequencing and identification of cellulose synthase genes in *Agave* species. *Genes.* 10(2):103.
- Indrayanto G, Rahayu L, Rahman A, Noeraeni PE. 1993. Effect of calcium, strontium, and magnesium ions on the formation of phytosteroids in callus cultures of *Agave amaniensis*. *Planta Med.* 59(1): 97–98.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33(7):1870–1874.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, et al. 2012. Soapdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience.* 1(1):18.
- Qin X, Yang X, Huang X, Huang X, Peng X, Liu M, Chen T, Yi K. 2021. The complete chloroplast genome of *Agave fourcroydes*. *Mitochondrial DNA B Resour.* 6(8):2326–2327.
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73.
- 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.