


The complete chloroplast genome sequence of the *Dioscorea persimilis* Prain et Burkill (Dioscoreaceae)

Tianxu Cao^{a,b}, Qianglong Zhu^b , Xin Chen^b, Putao Wang^b, Nan Shan^b, Qionghong Zhou^b and Yingjin Huang^{a,b}

^aMinistry of Education of the P.R. China, Key Laboratory of Crop Physiology, Ecology and Genetic Breeding, Jiangxi Agricultural University, Nanchang, China; ^bCollege of Agronomy, Jiangxi Agricultural University, Nanchang, China

ABSTRACT

Dioscorea persimilis belongs to *Dioscorea* genus, which is considered as one of the most popular food and traditional folk medicine in China. The complete chloroplast genome of *D. persimilis* was determined in this study. The total genome size was 153,219 bp in length, containing a pair of inverted repeats (IRs) of 25,477 bp, which were separated by large single copy (LSC) and small single copy (SSC) of 83,448 bp and 18,817 bp, respectively. The GC content is 37.01%. A total of 129 genes were predicted including 84 protein-coding genes, eight rRNA genes and 37 tRNA genes. Phylogenetic tree analysis of 24 species in the genus *Dioscorea* indicated that *D. persimilis* was closer to Chinese yam, but has remote phylogenetic relationship with Guinea yam.

ARTICLE HISTORY

Received 9 November 2019
Accepted 7 December 2019

KEYWORDS

Dioscorea persimilis;
Chloroplast genome;
phylogenetic relationship



Dioscorea persimilis Prain et Burkill, belongs to the family Dioscoreaceae and commonly known as ‘Huai-shan-yao’, is one of the important cash crop and traditional folk medicine in southern China. The rhizome of *D. persimilis* contain polysaccharides, saponins and allantoin, which have great pharmacological activities on immunomodulation, antioxidation, blood sugar and blood fat reducing (Yuan 2008). Moreover, *D. persimilis* have been approved as one of the most popular vegetable in China. It is significant for further cultivation, breeding, and utilization to clarify the phylogenetic relationships of species in *Dioscorea* (Zhou et al. 2018). Their phylogenetic relationship in *Dioscorea* that contains approximately 600 species has not been better understood (Zhao et al. 2018). The chloroplast genome is smaller in size and has very low recombination, which makes it suitable for species identification and phylogenetic studies (Zhu et al. 2016). Therefore, we characterized the complete chloroplast genome sequence of *D. persimilis* to confirm its phylogenetic position and promote its utilization.

Sample of *D. persimilis* (no. JY236) was collected in Sanming Academy of Agricultural Sciences (26°14′41″N, 117°37′8″E, Fujian, China) and stored in Jiangxi agricultural university (Nanchang, China). The total genomic DNA of *D. persimilis* was isolated from the young leaves using the Plant Genomic DNA Kit (TaKaRa, Beijing, China). The DNA sample was sent to BioMarker (Beijing, China) for constructing a 150 bp pair-end library and being sequenced by illumina HiSeq2500. About 4 Gb of sequence data was obtained after sequencing and base quality control, 0.5 Gb of clean pair-end reads was randomly extracted using Seqtk. The *D. persimilis*

chloroplast genome was assembled by using SPAdes (v 3.12.0) (Bankevich et al. 2012), BlastN (v2.7.1), and GapCloser (v1.12-r6). The paired-end reads were qualitatively assembled with using the Plasmidspades.py in SPAdes. Chloroplast genome sequence contigs were retrieved, ordered and incorporated into a single scaffold sequence by comparison with the chloroplast genome of *Dioscorea alata* (NC_039707.1) using BlastN. The gaps in the chloroplast single draft sequence of were closed by using GapCloser. Chloroplast genome annotation was annotated using CPGAVA2 (Shi et al. 2019) and manually corrected by Sequin and IGV (Robinson et al. 2011).

The complete chloroplast genome of *D. persimilis* (MN585218) is 153,219 bp in length with 37.01% GC contents. The genome was composed of two inverted repeat regions (IRa and IRb) of 25,477 bp, which were divided by a small single copy (SSC) region of 83,448 bp and a large single copy (LSC) region of 18,817 bp. There is a total of 129 genes, including 84 protein-coding genes, eight rRNA genes and 37 tRNA genes. Five of the protein-coding genes, eight of the tRNA genes and four rRNA genes are duplicated within the IRs.

To reveal the phylogenetic position of *D. persimilis*, a phylogenetic analysis was conducted on the complete chloroplast genomes of *D. persimilis* and other 23 species downloaded from the genome database in NCBI. The phylogenetic tree was constructed by Neighbor-Joining method using MAFFT v7.407 (Nakamura et al. 2018) and MEGA v10.0.4 (Kumar et al. 2018). The results demonstrated that *D. persimilis* have closer relationship with *D. alata* and *D. polystachya* that were also known as Chinese yam, but has

CONTACT Yingjin Huang  yjhuang_cn@126.com  College of Agronomy, Jiangxi Agricultural University, No.1101 Zhimin Street Qingshanhu District, Nanchang, Jiangxi Province 330045, China

© 2020 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.

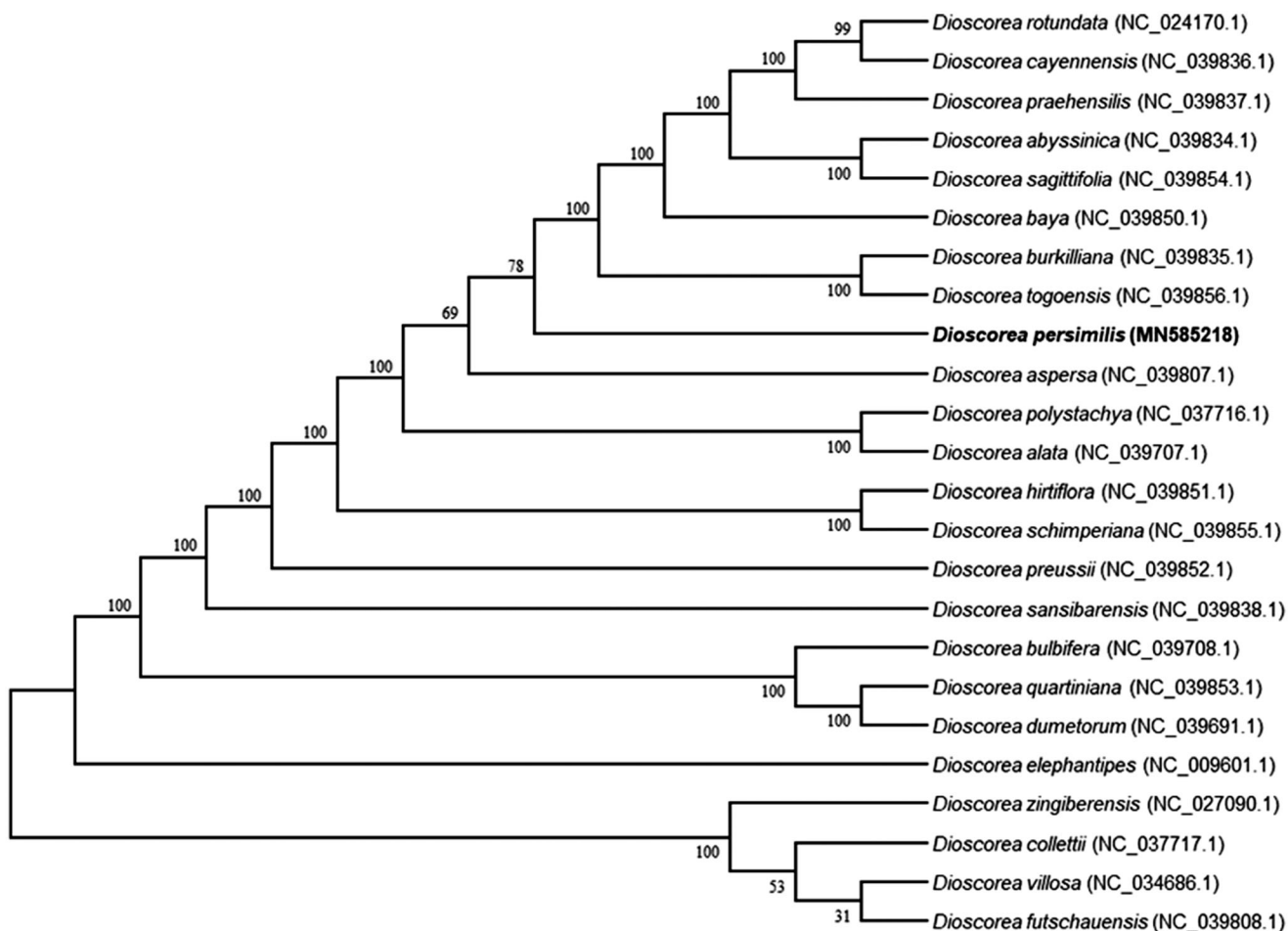


Figure 1. Phylogenetic tree showing relationship of *Dioscorea persimilis* with other 23 species in the genus *Dioscorea*. Phylogenetic tree was constructed based on the complete chloroplast genomes using Neighbor-Joining method (NJ) with 1,000 bootstrap replicates. Numbers in each the node indicated the bootstrap support values.

remote phylogenetic relationship with Guinea yam, *D. rotundata*, that is a staple food in Guinea (Figure 1). These findings further support the conclusions of the previous researches (Wu et al. 2014; Zhao et al. 2018).

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding

This work was supported by the earmarked fund for Jiangxi Agriculture Research System [JXARS-19] and Jiangxi Provincial Key Research and Development Project [20192BBF60005].

ORCID

Qianglong Zhu  <http://orcid.org/0000-0001-5355-8815>

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.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 35(6):1547–1549.
- Nakamura T, Yamada KD, Tomii K, Katoh K. 2018. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics.* 34(14):2490–2492.
- Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer [10.1038/nbt.1754]. *Nat Biotechnol.* 29(1):24–26.
- 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.
- Wu ZG, Li XX, Lin XC, Jiang W, Tao ZM, Mantri N, Fan CY, Bao XQ. 2014. Genetic diversity analysis of yams (*Dioscorea* spp.) cultivated in China using ISSR and SRAP markers. *Genet Resour Crop Evol.* 61(3):639–650.
- Yuan SL. 2008. Research advances on chemical compositions and bioactivities of *Dioscorea opposita* thunb. (Chinese yam). *Food Res Dev.* 29(3):176–179.
- Zhao Z, Wang X, Yu Y, Yuan S, Jiang D, Zhang Y, Zhang T, Zhong W, Yuan Q, Huang L. 2018. Complete chloroplast genome sequences of *Dioscorea*: characterization, genomic resources, and phylogenetic analyses. *PeerJ.* 6:e6032.
- Zhou W, Li B, Li L, Ma W, Liu Y, Feng S, Wang Z. 2018. Genome survey sequencing of *Dioscorea zingiberensis*. *Geno Nat Res Council Canada.* 61(8):567–574.
- Zhu Q, Gao P, Liu S, Amanullah S, Luan F. 2016. Comparative analysis of single nucleotide polymorphisms in the nuclear, chloroplast, and mitochondrial genomes in identification of phylogenetic association among seven melon. *Breed Sci.* 66(5):711–719.