


The complete chloroplast genome sequence of the *Dioscorea esculenta* (Lour.) Burkill (Dioscoreaceae)

Xin Chen^a, Lijuan Cai^b, Yu Zhang^a, Wenzheng Su^a, Bicong Li^a, Qinghong Zhou^a and Qianglong Zhu^a 

^aDepartment of Horticulture, College of Agronomy, Jiangxi Agricultural University, Nanchang, P.R. China; ^bNanchang Business College, Jiangxi Agricultural University, Nanchang, P.R. China

ABSTRACT

Dioscorea esculenta (Lour.) Burkill is an essential tuber crop with pharmacological effects in the family Dioscoreaceae. The complete chloroplast genome of *D. esculenta* was determined in this study. The total genome size is 153,437 bp in length and demonstrates a typical quadripartite structure containing a large single copy (LSC, 83,628 bp) and a small single copy (SSC, 18,893 bp), separated by a pair of inverted repeats (IRa, IRb) of 25,458 bp. The GC content of the complete chloroplast genome sequence is 37.07%. A total of 131 genes were predicted including 86 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Phylogenetic tree analysis of 25 species belonging to the genus *Dioscorea* indicated that *D. esculenta* and *D. sansibarensis* were clustered into one branch.

ARTICLE HISTORY

Received 3 August 2020
Accepted 30 September 2020

KEYWORDS

Dioscorea esculenta (Lour.)
Burkill; chloroplast genome;
lesser yam

Dioscorea esculenta (Lour.) Burkill (Dioscoreaceae) is an edible tuber crop and belongs to the family Dioscoreaceae, commonly known as lesser yam (Olayemi and Ajaiyeoba 2007). The tuber of *D. esculenta* contains rich starchprotein and soluble sugar, thus it has been as one of the staple foods in most parts of West Africa, and has been also used as an crude drug for anti-inflammatory, anti-stress, anti-spasmodic, and antifatigue for its medical elements of allantoin, saponin and dioscin (Murugan and Mohan 2012; Lee et al. 2017). *Dioscorea* is highly species-rich (approximately 600 species) and morphologically complex genus, and their phylogenetic relationship has not been uncovered better (Magwé-Tindo et al. 2019; Darkwa et al. 2020). The chloroplast genome has smaller size and low recombination, which makes it suitable for species identification and molecular phylogenetic analyses (Govindarajulu et al. 2015). Here, we analyzed the characters of the complete chloroplast genome sequence for *D. esculenta*, to confirm its phylogenetic position and evolutionary relationship between the *D. esculenta* and other species in Dioscoreaceae family.

Dioscorea esculenta (YAM-256) was planted in the Yulan experimental station of Jiangxi Agricultural University (28°45'27"N, 115°50'20"E), Nanchang, China. Total DNA was extracted from fresh leaves and subjected to construct a genomic library and sequenced by HiSeq X Ten (BGI, Shenzhen, China). About 3 Gb of clean sequence data with ~5× coverage to its nuclear genome was obtained, and 0.4 Gb of clean pair-end reads with a length of 150 bp was randomly extracted using Seqtk. The lesser yam chloroplast graft genome was *de*

novo assembled using SPAdes (v 3.14.1) (Bankevich et al. 2012), BlastN (v2.7.1), and GapCloser (v1.12-r6). Specific method is as follows: Firstly, these reads were assembled with using the Plasmidspades.py in SPAdes. Secondly, Contigs representing the chloroplast genome were then retrieved, ordered and incorporated into a single draft sequence by BlastN against the chloroplast genome of *D. polystachya* (NC_037716.1). Thirdly, the gaps in the chloroplast single draft sequence of were closed by using GapCloser. Finally, the complete genome sequence was annotated using CPGAVA2 (Shi et al. 2019) and manually corrected by Sequin and IGV (Thorvaldsdóttir et al. 2013).

The complete chloroplast genome of *D. esculenta* (GenBank accession number: MT818507) is 153,437 bp in length with 37.07% GC contents and exhibits a typical quadripartite structure, consisting of a pair of inverted repeat regions (IRa and IRb) with same length (25,458 bp) separated by the large single copy (LSC, 83,628 bp) and small single copy (SSC, 18,893 bp) regions. A total of 131 genes were identified, including 86 protein-coding genes, eight rRNA genes, and 37 tRNA genes; six of the protein-coding genes, six of the tRNA genes, and four rRNA genes are duplicated within the IRs. Noticeably, the start code of rpl2 is by RNA editing, and one of ycf1 is a pseudogene that has only partial fragment.

To explore the phylogenetic position and evolutionary relationship of *D. esculenta*, a phylogenetic analysis was conducted between *D. esculenta* and other 24 complete chloroplast genomes of Dioscoreaceae downloaded from the organelle genome database in NCBI using Batch Entrez, and

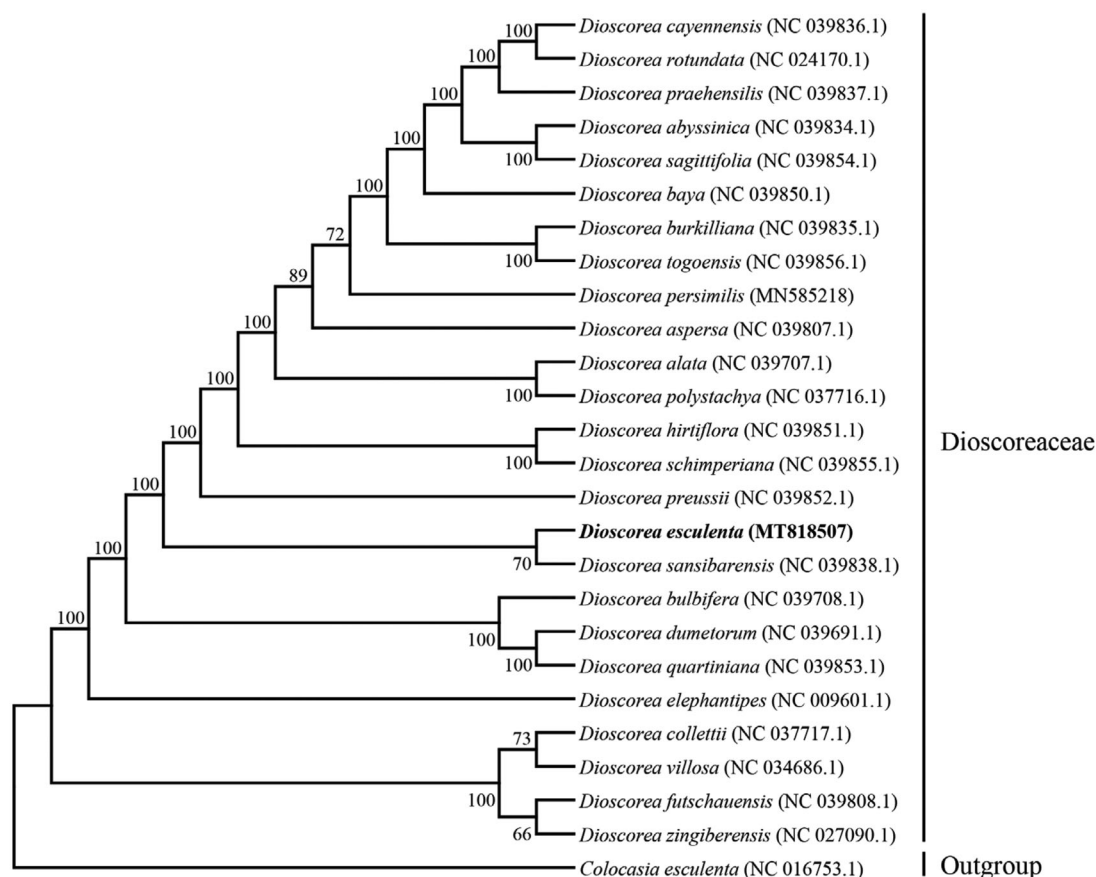


Figure 1. Phylogenetic tree showing the relationship between *D. esculenta* and 24 species belonging to Dioscoreaceae family, *C. esculenta* was taken as the outgroup. Phylogenetic tree was constructed based on the complete chloroplast genomes using maximum likelihood (ML) with 1000 bootstrap replicates. Numbers in each the node indicated the bootstrap support values.

Colocasia esculenta chloroplast genome as the outgroup. These sequences above were aligned using MAFFT v7.407 (Nakamura et al. 2018) and the phylogenetic tree was constructed using MEGA v10.0.4 (Kumar et al. 2018) with maximum likelihood (ML) method. The result was well-resolved and revealed that *D. esculenta* was belonged to Dioscoreaceae and closer to *D. sansibarensis* (Figure 1), these findings further enriched the phylogenetic relationship of the family Dioscoreaceae (Cao et al. 2018; Zhao et al. 2018; Magwé-Tindo et al. 2019).

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 funded by the Youth Science Foundation of Jiangxi Provincial Department of Education [GJJ180218], and the Construction Project of Root and Tuber Crops Industry and Technology System [JXARS-19-01].

ORCID

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

Data availability statement

The data that support the findings of this study are openly available in GenBank at <https://www.ncbi.nlm.nih.gov/nuccore/MT818507.1/>, reference number MT818507.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Cao J, Jiang D, Zhao Z, Yuan S, Zhang Y, Zhang T, Zhong W, Yuan Q, Huang L. 2018. Development of chloroplast genomic resources in Chinese yam (*Dioscorea polystachya*). *Biomed Res Int.* 2018: 6293811–6293847.
- Darkwa K, Olasanmi B, Asiedu R, Asfaw AJPB. 2020. Review of empirical and emerging breeding methods and tools for yam (*Dioscorea* spp.) improvement: status and prospects. *Plant Breed.* 139(3):424–474.
- Govindarajulu R, Parks M, Tennessen JA, Liston A, Ashman TL. 2015. Comparison of nuclear, plastid, and mitochondrial phylogenies and the origin of wild octoploid strawberry species. *Am J Bot.* 102(4): 544–554.
- 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.
- Lee HJ, Watanabe B, Nakayasu M, Onjo M, Sugimoto Y, Mizutani M. 2017. Novel steroidal saponins from *Dioscorea esculenta* (Togedokoro). *Biosci Biotechnol Biochem.* 81(12):2253–2260.

- Magwé-Tindo J, Wieringa JJ, Sonké B, Zapfack L, Vigouroux Y, Couvreur TLP, Scarcelli N. 2019. Complete plastome sequences of 14 African yam species (*Dioscorea* spp.). *Mitochondrial DNA B*. 4(1):74–76.
- Murugan M, Mohan VR. 2012. In vitro antioxidant studies of *Dioscorea esculenta* (Lour). *Burkill. Asian Pac J Trop Bio*. 2(3):S1620–S1624.
- Nakamura T, Yamada KD, Tomii K, Katoh K. 2018. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics*. 34(14):2490–2492.
- Olayemi JO, Ajaiyeoba EO. 2007. Anti-inflammatory studies of yam (*Dioscorea esculenta*) extract on wistar rats. *Afr J Biotechnol*. 6(16):1913–1915.
- 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.
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinformatics*. 14(2):178–192.
- 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.