


The complete chloroplast genome and phylogenetic analysis of *Ocimum kilimandscharicum* Gurke (Camphor Basil) from India

Samuel Yesuthason Renald, Raju Balaji , Tanuja  and Madasamy Parani 

Center for DNA Barcoding, Department of Genetic Engineering, SRM Institute of Science and Technology, Kattankulathur, India

ABSTRACT

Ocimum kilimandscharicum Gurke commonly known as Camphor Basil, is a medicinal plant species that belongs to the Lamiaceae family. Here, the sequencing and characterization of complete chloroplast genome sequence of *O. kilimandscharicum* is reported for the first time using Illumina paired-end sequencing data. The size of the chloroplast (cp) genome is 151,741 bp in length, with a large single-copy (LSC) region of 82,882 bp and a small single-copy (SSC) region of 17,587 bp, separated by a pair of 25,636 bp inverted repeat (IR) regions. There are 135 predicted genes, including 90 protein-coding genes, 37 transfer RNA (tRNA) genes, and eight ribosomal RNA (rRNA) genes in the genome, and the overall GC content of the genome is 37.9%. The phylogenetic analysis based on the chloroplast genome data indicated that *O. kilimandscharicum* is closer to *O. tenuiflorum* and clustered to other *Ocimum* species in Lamiaceae.

ARTICLE HISTORY

Received 6 May 2021
Accepted 15 June 2021

KEYWORDS

Complete chloroplast genome; *Ocimum kilimandscharicum*; Camphor Basil; Lamiaceae; Karpoora Tulsi

Ocimum kilimandscharicum Gurke commonly known as Camphor Basil or Karpoora Tulsi in Tamil, belongs to Lamiaceae family. It is a shrub native to East Africa, which was brought to India and used to treat constipation, abdominal pain, cough, measles, and diarrhea (Gill et al. 2012). The leaves are rich in essential oil (0.004–0.7%), and camphor is the major component used in perfumery, flavor, and pharmaceutical industries (Lima et al. 2014). This plant is used in traditional medicine for its various pharmacological activities, including antioxidant, antimicrobial, wound healing, anti-malarial, antifungal, anti-nociceptive, anti-diarrhoeal, and insecticidal properties (Sarin et al. 2013). To provide genomic resources for investigating the evolution of *O. kilimandscharicum*, we sequenced the complete chloroplast genome sequence of this species and studied its phylogenetic relationships within the Lamiaceae family.

Fresh leaves of *O. kilimandscharicum* were collected from the Foundation for Revitalization of Local Health Traditions (FRLHT), Yelahanka, Bengaluru, Karnataka, India (GPS: 13°07'24.5"N 77°32'52.3"E). The voucher specimen was deposited at the SRM Institute of Science and Technology Herbarium (<http://www.srmist.edu/>, Dr. P. Senthilkumar, senthilp3@srmist.edu.in) under the voucher number SRMH000145. The total genomic DNA was extracted from fresh leaves as described before (Vassou et al. 2015) and stored at –80°C until use. A whole-genome DNA sequencing library was constructed using the Nextera XT Library Prep Kit. The DNA library was sequenced on NextSeq 500 platform (Illumina Inc., San Diego, CA), and we obtained 2.5 Gb of

paired-end sequencing data. The chloroplast genome of *O. kilimandscharicum* was assembled with 984x coverage using NovoPlasty v.4.3.1 (k-mer 33) (Dierckxsens et al. 2017) with *O. gratissimum* L. (Balaji et al. 2021) as a reference seed sequence. The assembled *O. kilimandscharicum* chloroplast genome was annotated by GeSeq (Tillich et al. 2017) using the chloroplast genomes of *O. tenuiflorum* (NC_043873.1) and *O. basilicum* (NC_035143.1) as reference sequences. The predicted transfer RNAs (tRNAs) were confirmed by tRNAscan-SE 2.0 (Lowe and Chan 2016).

The complete cp genome of *O. kilimandscharicum* (GenBank Accession No. MW829603) exhibits a typical quadripartite circular molecule with 151,741 bp genome size. It is composed of a large single-copy (LSC) region of 82,882 bp, a small single-copy region (SSC) of 17,587 bp, and a pair of inverted repeat (IR) regions of 25,636 bp. The GC content of the complete chloroplast genome was 37.9%, while the LSC, SSC, and IR regions are 36.0%, 31.8%, and 43.2%, respectively. The cp genome encodes an identical set of 135 predicted functional genes, including 90 protein-coding genes, 37 transfer RNA (tRNA) genes, and eight ribosomal RNA (rRNA) genes.

A maximum-likelihood tree of molecular distance analysis was performed using 1000 bootstrap replicates with (GTR + G) model in MEGA version X (Kumar et al. 2018) from the alignments created using the MAFFT program (Kato and Standley 2013). *Nicotiana tabacum* (Solanaceae) and *Arabidopsis thaliana* (Brassicaceae) were designated as out-group taxa, and 16 published complete chloroplast genome

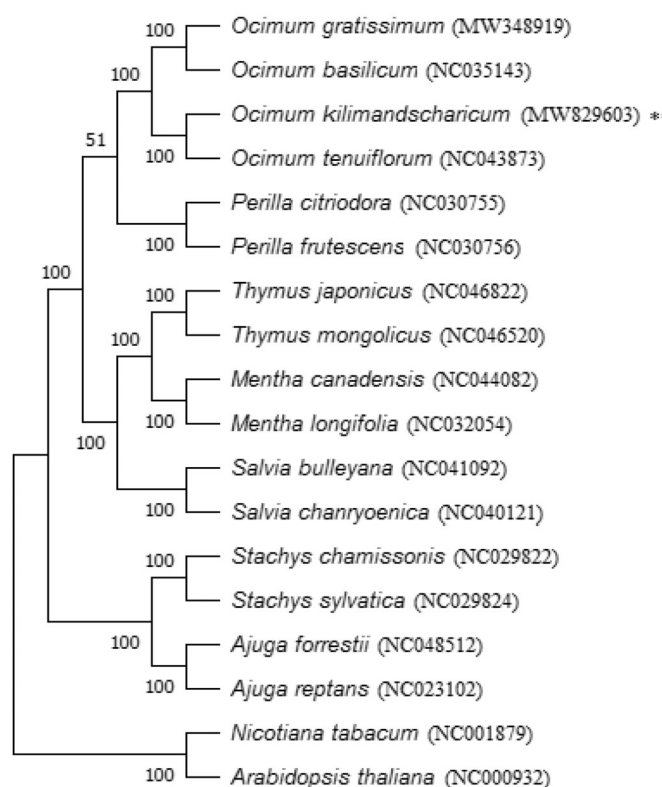


Figure 1. Phylogenetic tree reconstructed by maximum likelihood (ML) analysis based on chloroplast genome sequences, including *O. kilimandscharicum* (MW829603)* sequenced in this study. The bootstrap support values >50% from 1,000 replicates are indicated in the nodes.

sequences from the Lamiaceae family were included as in-group taxa. The phylogenetic analysis of *O. kilimandscharicum* revealed that it is closely related to *O. tenuiflorum*, *O. gratissimum*, and *O. basilicum* (Figure 1). This study enriches the genetic resources of *O. kilimandscharicum* and contributes to its species delimitation, phylogenetic analysis, DNA barcoding, and evolutionary relation of genus *Ocimum* and Lamiaceae in the future.

Acknowledgments

We would like to acknowledge the SRM Institute of Science and Technology for providing research and computation facilities. We thank Dr. K. Ravikumar and Dr. N. Dhatchnamoorthy, Foundation for Revitalisation of Local Health Tradition (FRLHT), The University of Trans-Disciplinary Health Sciences and Technology, Bengaluru-560064, Karnataka, India, for their help in the collection and authentication of plant material.

Funding

We would like to acknowledge the funding from the SRM-DBT Partnership Platform for Contemporary Research Services and Skill

Development in Advanced Life Sciences Technologies [Grant No. BT/PR12987/INF/22/205/2015] and Department of Biotechnology, Ministry of Science and Technology.

ORCID

Raju Balaji  <http://orcid.org/0000-0003-0173-1510>

Madasamy Parani  <http://orcid.org/0000-0002-2265-1715>

Data availability statement

The data that support the findings of this study are openly available in NCBI at <https://www.ncbi.nlm.nih.gov/nucleotide/MW829603.1/>, under GenBank accession number MW829603. The NGS sequencing data are available from the BioProject, SRA, and Bio-Sample ID under the accession numbers PRJNA717657, SRR14076648, and SAMN18499778 respectively.

Disclosure statement

The authors reported no potential competing interests.

References

- Balaji R, Ravichandiran K, Tanuja, Parani M. 2021. The complete chloroplast genome of *Ocimum gratissimum* from India – a medicinal plant in the Lamiaceae. *Mitochondrial DNA Part B*. 6:948–950.
- Dierckx N, Mardulyn P, Smits G. 2017. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Res*. 45(4):e18.
- Gill D, Soni N, Sagar B, Raheja S, Agarwal S. 2012. *Ocimum kilimandscharicum*: a systematic review. *J Drug Deliv Therap*. 2(3):45–52.
- 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, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 35(6):1547–1549.
- Lima VT, Vieira MC, Kassuya CAL, Cardoso CAL, Alves JM, Foglio MA, Carvalho JE, Formaggio ASN. 2014. Chemical composition and free radical-scavenging, anticancer and anti-inflammatory activities of the essential oil from *Ocimum kilimandscharicum*. *Phytomedicine*. 21(11):1298–1302.
- Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res*. 44(W1):W54–W57.
- Sarin RV, Narwal S, Bafna PA. 2013. Anti-diarrhoeal activity of aqueous extract of *Ocimum kilimandscharicum*. *J Ethnopharmacol*. 148(1):223–228.
- 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.
- Vassou SL, Kusuma G, Parani M. 2015. DNA barcoding for species identification from dried and powdered plant parts: a case study with authentication of the raw drug market. *Gene*. 15(1):86–93. 559