

Characterization of the complete chloroplast genome of Sanming wild banana (*Musa itinerans*) and phylogenetic relationships

Yong-Yan Zhang, Fan Liu, Na Tian, Jing-Ru Che, Xue-Li Sun, Zhong-Xiong Lai and Chun-Zhen Cheng

College of Horticulture, Fujian Agriculture and Forestry University, Fuzhou, P. R. China

ABSTRACT

Sanming wild banana (*Musa itinerans*) is a cold resistant wild banana found in Fujian province of China. In this study, we characterized its complete chloroplast genome using BGISEQ-500 sequencing. The chloroplast genome is 171,815 bp in size, containing a pair of IR regions (35,142 bp), a large single copy region (89,995 bp), and a small single copy region (11,464 bp). The whole chloroplast genome contains 111 unique genes, including 78 protein-coding genes, 29 tRNAs, and 4 rRNAs. Phylogenetic maximum likelihood analysis revealed that Sanming wild banana showed the closest relationship with *Musa itinerans* that collected from Yunnan Province of China.

ARTICLE HISTORY

Received 20 June 2019
Accepted 22 June 2019

KEYWORDS

Musa itinerans; wild banana; chloroplast genome; phylogenetic relationship analysis

China is rich of wild banana species (Li et al. 2017), many of which showed high resistance to both abiotic and biotic stresses, such as cold and banana wilt disease (Li et al. 2015). In our previous study, we identified that Sanming wild banana showed the highest cold resistance (Liu, Cheng, Chen, et al. 2018; Liu, Cheng, Lin, et al. 2018), suggesting that it has great potential to be used in banana cold resistance breeding. Based on its phenotype characteristics, it was categorized into *Musa itinerans*, which is native to Southeast Asia and subtropical China (Häkkinen et al. 2008). However, molecular evidence that can clarify its evolution position and relationship with other banana species is limited. In this study, we assembled and characterized its complete chloroplast genome using BGISEQ-500 (BGI, Shenzhen, China) sequencing data and explored the phylogenetic relationship with other plant species, which can be used for the revealing of its evolution position and can also be subsequently used for the cold resistance genetic breeding and the evolutionary study of banana.

The specimen of Sanming wild banana was isolated from the banana germplasm resources nursery of Institute of Horticultural Biotechnology, Fujian Agriculture and Forestry University in Fuzhou (Fujian, China, 26°04'37.00"N;119°14'33.47"E) and the DNA stored at the Fujian Agriculture and Forestry University (No. FAFUYSJ01). High-quality genomic DNA was used for BGISEQ-500 sequencing (BGI, Shenzhen, China). Approximately 1.2 Gbp high-quality reads were obtained, and clean reads were aligned to chloroplast genomes of *Musa acuminata* (Genbank accession No. HF677508.1), *Musa balnisi*

(Genbank accession No. NC_028439.1) and *Musa itinerans* (Genbank accession No. NC_035723.1) and then further assembled into contigs using CLC Genomics Workbench v8.0 (CLC Bio, Aarhus, Denmark). The assembled chloroplast genome was assembled annotated using DOGMA (Wyman et al. 2004) and annotation correction was performed using Geneious (Kearse et al. 2012). The annotated chloroplast genome has been deposited in Genbank with the accession number MN087226.

The complete chloroplast genome of Sanming wild banana is 171,815 bp in size, containing a pair of inverted repeat (IR) regions of 35,142 bp, a large single copy region of 89,995 bp, and a small single copy region of 11,464 bp. The chloroplast contains 111 unique genes, including 78 protein-coding genes, 29 tRNA genes, and 4 rRNA genes. Most of them (89) occur as a single copy, but 10 protein-coding genes (i.e. *ycf2*, *ycf1*, *rps19*, *rps15*, *rps12*, *rps7*, *rpl23*, *rpl2*, *ndhH*, and *ndhB*), 8 tRNA genes (i.e. *trnV-GAC*, *trnR-ACG*, *trnN-GUU*, *trnL-CAA*, *trnI-GAU*, *trnI-CAU*, *trnH-GUG*, and *trnA-UGC*) and all the 4 rRNA genes (*rrn4.5*, *rrn5*, *rrn16* and *rrn23rRNA*) occur in double copies. The overall nucleotide composition of the chloroplast genome is: 31.5% A, 32.0% T, 18.5% C, and 18.0% G, with the total GC content of 36.5%.

For phylogenetic maximum-likelihood analysis, we downloaded the chloroplast genomes of 16 plant species belonging to Scitamineae from GenBank to access the relationship of Sanming wild banana with them. Sequence alignment was then performed using HomBlocks pipeline (Bi et al. 2018). RAxML-HPC2 on XSEDE version 8.2.10 (Stamatakis 2014) was used to construct a maximum-likelihood tree, for which 1000

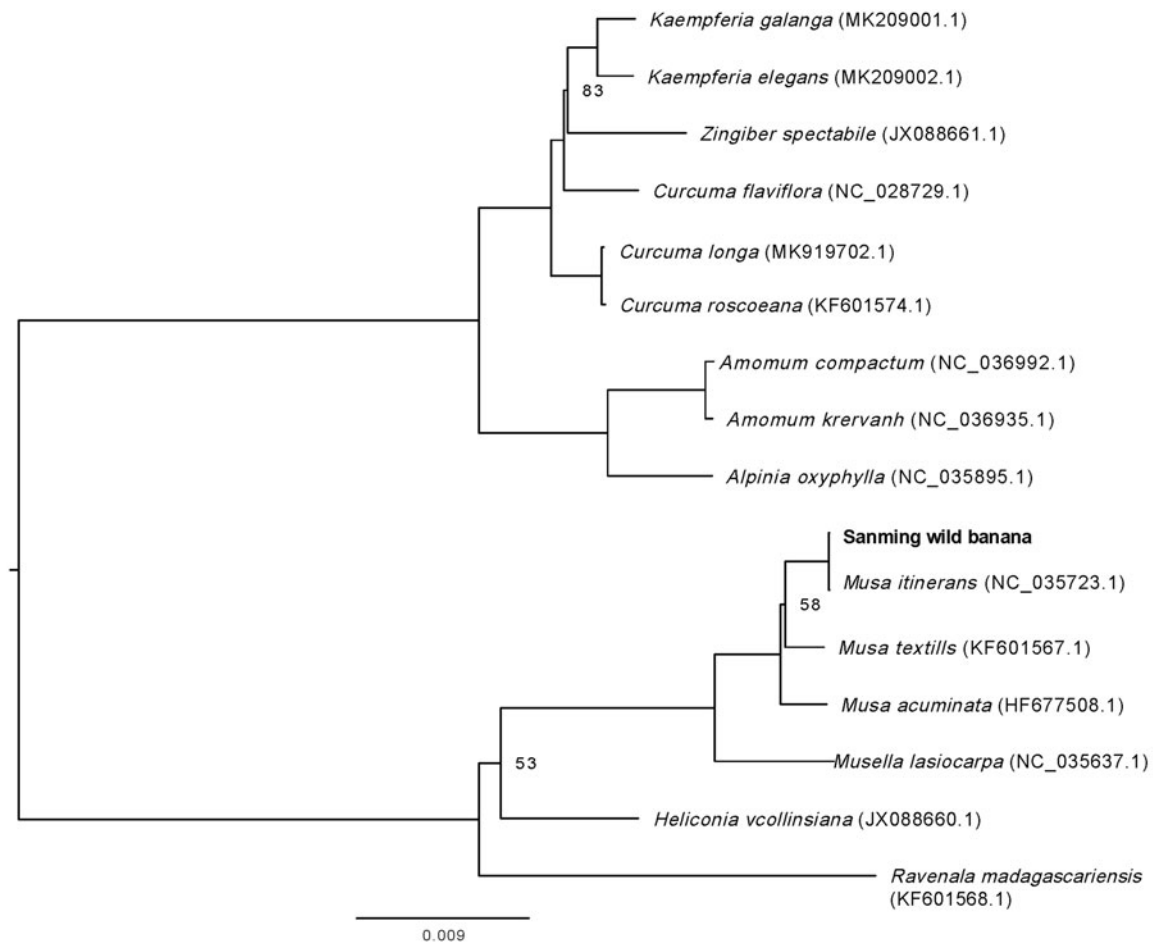


Figure 1. Maximum-likelihood phylogenetic tree based on the complete chloroplast genome sequences of 16 species from the order Scitamineae. Numbers on the nodes are bootstrap values with 1000 replicates and bootstrap values of 100 were omitted.

bootstrap replicates were used to calculate the bootstrap values. Result showed that Sanming wild banana showed a closest relationship with *Musa itinerans*, followed by *Musa balnisiensis* and *Musa textilis* (Figure 1), all from the family Musaceae, indicating that Sanming wild banana belongs to *Musa itinerans* of Musaceae. There are many wild banana species distributed in China, and the relationships among them and with other banana species need to be clarified. Our present study revealed that the chloroplast genome size of Sanming wild banana reported in this study is larger than that *Musa itinerans* collected in Yunnan Province, China (Li et al. 2017), and their differences mainly occurred in the intergenic sequences. Thus, it was suspected that the intergenic sequences of the chloroplast can be used in the phylogenetic analysis of wild banana species. Moreover, the complete chloroplast of Sanming wild banana would be subsequently used in the evolutionary studies of *Musa* species and would provide genetic information for the banana cold resistance breeding.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Natural Science Funds for Distinguished Young Scholar of the Fujian Agriculture and Forestry University [xjq201721], the Natural Science Foundation of China [31601713], the Postdoctoral Funds of the Fujian Agriculture and Forestry University [132300197], and the Construction of Plateau Discipline of Fujian Province [102/71201801101].

References

- Bi GQ, Mao YX, Xing QK, Cao M. 2018. HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching. *Genomics*. 110:18–22.
- Häkkinen M, Wang H, Ge XJ. 2008. *Musa itinerans* (Musaceae) and its intraspecific taxa in China. *Novon*. 18:50–60.
- 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.
- Li WM, Dita M, Wu W, Hu GB, Xie JH, Ge XJ. 2015. Resistance sources to *Fusarium oxysporum* f. sp. *cubense* tropical race 4 in banana wild relatives. *Plant Pathol*. 64:1061–1067.
- Li W, Liu Y, Gao LZ. 2017. The complete chloroplast genome of the endangered wild *Musa itinerans* (Zingiberales: Musaceae). *Mitochondrial DNA Part B*. 4:1562–1563.

- Liu WH, Cheng CZ, Chen FL, Ni SS, Lin YL, Lai ZX. 2018. High-throughput sequencing of small RNAs revealed the diversified cold-responsive pathways during cold stress in the wild banana (*Musa itinerans*). *BMC Plant Biology*. 18:308.
- Liu WH, Cheng CZ, Lin YL, XuHan X, Lai ZX. 2018. Genome-wide identification and characterization of mRNAs and lncRNAs involved in cold stress in the wild banana (*Musa itinerans*). *PLoS One*. 13:e0200002.
- Stamatakis A. 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30: 1312–1313.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 20: 3252–3255.