


The complete chloroplast genome sequences of little millet (*Panicum sumatrense* Roth ex Roem. and Schult.) (Poaceae)

Raveendar Sebastin^a, Gi-An Lee^a , Kyung Jun Lee^a, Myoung-Jae Shin^a, Gyu-Taek Cho^a, Jung-Ro Lee^a, Kyung-Ho Ma^a and Jong-Wook Chung^b

^aNational Agrobiodiversity Center, National Institute of Agricultural Sciences, Jeonju-Si, Republic of Korea; ^bDepartment of Industrial Plant Science and Technology, Chungbuk National University, Cheongju, Republic of Korea

ABSTRACT

Little millet, *Panicum sumatrense* Roth ex Roem. & Schult., is an important cultivated species under the tribe Paniceae, sub-family Panicoideae and family Poaceae. In this study, for the first time we sequenced the complete chloroplast (cp) genome of *P. sumatrense* to investigate their phylogenetic relationship in the family Poaceae. The complete cp genome sequence of *P. sumatrense* is 139,384 bp in length with 38.6% overall GC content and exhibits a typical quadripartite structure comprising one pair of inverted repeats (22,723 bp) separated by a small single-copy region (12,583 bp) and a large single-copy region (81,355 bp). The *P. sumatrense* cp genome encodes 125 unique genes, which include 91 protein-coding genes, 4 rRNA genes, 30 tRNA genes, and 20 genes were duplicated in the inverted repeat region. This newly determined cp genome (*P. sumatrense*) could be valuable information for the breeding programs of this cereal crops in the family Poaceae.

ARTICLE HISTORY

Received 17 May 2018
Accepted 29 May 2018

KEYWORDS

Chloroplast; little millet; Illumina sequencing; *P. sumatrense*

Introduction

The grass genus *Panicum* L., comprises of about 500 species distributed worldwide in tropical and subtropical regions, is one of the largest genera of the family Poaceae (Aliscioni et al. 2003). The great significance of this genus is almost certainly polyphyletic (Zuloaga 1987) in which the species with C₃ and C₄ photosynthetic systems (Brown 1977), and C₃/C₄ intermediate pathways (Zuloaga et al. 1998) have been reported. Though many dominating grasses found in the genus *Panicum*, common millet (*Panicum miliaceum*) and little millet (*Panicum sumatrense*) are economically important species within the genus (Baltensperger 1996). Understanding the relationships between the grasses is vital for germplasm conservation and their utilization. The cp genomes have been used in studies on plant phylogeography, genetic diversity, and evolution (Liu et al. 2016; Park 2016, 2017; Tsuruta et al. 2017). The cp genome of common millet was reported (Cao et al. 2017) and many other economically important *Panicum* cp genome were also available on public database. In this study, we report the chloroplast genome of little millet to determine its phylogenetic relationships within the family Poaceae.

The little millet seeds (Accession No. IT261894) were obtained from the Genebank division of National Institute of Agricultural Sciences, Republic of Korea. Seeds were

germinated and fresh leaves were collected from 40-day-old seedlings. Total genomic DNA was extracted to build up genomic library and sequenced with pair-end (2 × 300 bp) by MiSeq instrument at LabGenomics (<http://www.Labgenomics.com/kor/>). A total of 4,915,776 cleaned reads were obtained after quality trimming of raw reads and mapped with the reference cp genome, *A. sativa* L. (GenBank accession KM974733). The reference mapping produced 206,243 aligned reads with about an average 325× coverage. The complete circular cp genome was obtained from contig alignment and scaffolding of mapped reads. DOGMA (<http://dogma.cccb.utexas.edu/>) software was used for annotation of protein-coding genes in the cp genome and manually inspected to predict transfer RNA (tRNA) and ribosomal RNA (rRNA) genes.

We determined the cp genome of *P. sumatrense* was 139,384 bp in length (NCBI accession number KX756177). The GC content was 38.6% which is similar to values previously reported with *Panicum* cp genomes. The LSC and SSC regions contained 81,355 bp and 12,583 bp, respectively, whereas the IR was 22,723 bp in length. The cp genome contained 125 known genes, including 91 protein-coding genes, 30 tRNA genes, and four rRNA genes. There were eight protein-coding genes, eight tRNA, and all four rRNA genes duplicated in the IR regions. Sixteen genes contained one or two introns, including the protein-coding genes,

CONTACT Kyung-Ho Ma  khma@korea.kr  National Agrobiodiversity Center, National Institute of Agricultural Sciences, RDA, 370 Nonsaengmyeong-Ro, Wansan-Gu, Jeonju-Si, Jeollabuk-Do 54874, Republic of Korea; Jong-Wook Chung  jwchung73@chungbuk.ac.kr  Department of Industrial Plant Science and Technology, Chungbuk National University, Cheongju, Chungbuk 28644, Republic of Korea

© 2018 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.

