Complete chloroplast genome of a novel chlorophyll-deficient mutant (clm) in sweetpotato (Ipomoea batatas L.)

Hongda Zou, Xiongjian Zhang, Jingyi Chen, Zangying Wang, Lifei Huang and Boping Fang
Guangdong Provincial Key Laboratory of Crops Genetics and Improvement, Crops Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, China

ABSTRACT
The complete chloroplast genome of a novel chlorophyll-deficient mutant (clm) and its wild type (WT) in sweetpotato (Ipomoea batatas L.) was sequenced. The complete chloroplast genome of clm and WT was 161,393 bp and 161,429 bp in length, containing a large single copy (LSC) region of 87,561 bp and 87,597 bp, respectively, a small single copy (SSC) region with the same length of 30,890 bp and a pair of inverted repeat regions (IRs) with the same length of 12,052 bp. Both of them contained 132 genes including 87 protein-coding sequences, 37 tRNA, and eight rRNA. Comparing to the WT, four SNPs and three INDELs were detected and only one INDEL in the exon affecting the translation of rpoA gene. Phylogenetic analysis showed that clm and WT were closely related to Ipomoea tabescana. The complete chloroplast genome of clm and its WT will play a role in understanding the molecular mechanism of chlorophyll deficiency and developing molecular markers in sweetpotato.

The sweetpotato (Ipomoea batatas L.) is generally vegetatively propagated by means of roots or cuttings. The sports (vegetative mutations) are fairly universal in sweetpotato cultivation (Gustafsson and Gadd 2009). As one kind of common mutations in sweetpotato, although the chlorophyll-deficient mutant was found as early as 1952 (Kuwata 1950), its physiological and molecular mechanisms have not been studied in detail. The chlorophyll-deficient mutant is a valuable material for research in chlorophyll biosynthesis, chloroplast development and structure function, photosynthesis mechanism, gene function, as well as ornamental plant breeding (Frank and Chitwood 2016). Here, we first sequenced and characterized the complete chloroplast genome of a novel chlorophyll-deficient mutant (clm) and its wild type (WT) in sweetpotato (Ipomoea batatas L.), and analyzed their phylogenetic relationship in Ipomoea genus.

The fresh leaves of clm and WT were collected in Guangzhou, Guangdong Province, China (N23°39′41.36′′; E113°44′04.18′′). Voucher specimens and DNA were deposited in the Crops Research Institute, Guangdong Academy of Agricultural Sciences (voucher number: CRI20191024). The total genomic DNA of clm and WT was extracted using the Plant Genomic DNA Kit (Tiangen, Beijing, China). The whole genome sequencings were conducted by Nanjing GenePioneer Biotechnologies Inc. (Nanjing, China) using Illumina Hiseq 2500 system. The raw reads were filtered by Trimmomatic v0.32 (Bolger et al. 2014) and mapped to Xushu 18 chloroplast genome (Yan et al. 2015) as the reference by Bowtie2 v2.35 (Langmead and Salzberg 2012). The obtained high-quality chloroplast reads were assembled in chloroplast genome by SPAdes v3.10 (Bankevich et al. 2012). The final annotations were performed using GeSeq (Tillich et al. 2017) and Blast search.

The complete chloroplast genome of clm (accession number MW122506) and WT (accession number MW122507) was 161,393 bp and 161,429 bp in length, containing a large single copy (LSC) region of 87,561 bp and 87,597 bp, respectively. Both of them had a small single copy (SSC) region with the same length of 30,890 bp and a pair of inverted repeat regions (IRs) with the same length of 12,052 bp. The annotation results showed that both the chloroplast genome of clm and WT contained 132 genes, including 87 mRNA, 37 tRNA, and eight rRNA. The structure variation between clm and WT chloroplast genome was screened by MAFFT v7.450 (Katoh and Standley 2013) and perl script. Comparing to WT, four SNPs in ycf1 gene and three INDELs in rpoA gene were identified in clm. But there was only one INDEL variation affecting gene expression. The rpoA protein was shortened by the alternation in exon locus of rpoA gene, which may lead to the phenotype of chlorophyll-deficient.

To figure out the phylogenetic status of clm and its WT in Ipomoea genus, the maximum-likelihood tree was constructed based on eight complete chloroplast genomes of Ipomoea species from GenBank, and Cressa cretica and Cuscuta reflexa were used as the outgroup. These 12 whole chloroplast genomes were aligned using MAFFT v7.450.
Katoh and Standley (2013) and the maximum-likelihood tree was constructed using RAxML v8.2.10 (Stamatakis 2014) with 1000 bootstrap value (Figure 1). The result of phylogenetic analysis showed that the clm mutant and its WT were clustered together, and close to Ipomoea tabasca. It was further confirmed that I. tabasca could be a recent hybrid between I. batatas and I. trifida, and there were two I. batatas lineages due to chloroplast capture (Muñoz-Rodriguez et al. 2018). The complete chloroplast genome of clm and its WT will play an important role in understanding the molecular mechanism of chlorophyll deficiency and developing molecular markers in sweetpotato.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by the Natural Science Foundation of Guangdong Province under Grant number [2020A1515011146], the Science and Technology Program of Guangzhou under Grant number [201804010161], the Special Foundation of President of the Guangdong Agricultural Academy under Grant number [201833], and the Sweetpotato of China Agricultural Research System under Grant numbers [CARS-10-B4 and CARS-10-C17].

**Data availability statement**

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at [https://www.ncbi.nlm.nih.gov/] under the accession nos. MW122506 and MW122507. The associated BioProject and BioSample numbers are PRJNA684785, and SAMN17069323 and SAMN17069324, respectively.

**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.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina Sequence Data. Bioinformatics. 30(15):2114–2120.

Gustafsson Å, Gadd I. 2009. Mutations and crop improvement. III. Ipomoea batatas (L.) Poir. (Convolvulaceae). Hereditas. 53(1–2):77–89.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Kuwata A. 1950. Notes on chimeras in sweet potatoes. Japan J Crop Sci. 19(1–2):143–146.

Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9(4):357–359.

Frank MH, Chitwood DH. 2016. Plant chimeras: the good, the bad, and the ‘Bizzaria’. Dev Biol. 419:41–53.

Muñoz-Rodriguez P, Carruthers T, Wood JRI, Williams BRM, Weitemier K, Kronmiller B, Ellis D, Anglin NL, Longway L, Harris SA, et al. 2018. Reconciling conflicting phylogenies in the origin of sweet potato and dispersal to Polynesia. Curr Biol. 28(8):1246–1256.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313.

Tillich M, Lehwarz 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.

Yan L, Lai X, Li X, Wei C, Tan X, Zhang Y. 2015. Analyses of the complete genome and gene expression of chloroplast of sweet potato [Ipomoea batata]. PLoS One. 10(4):e0124083.