The complete mitochondrial genome and phylogenetic analysis of sugarcane (Saccharum spp. hybrids) line 15a-53

Xiaolan Liu\textsuperscript{a,b,c}, Liping Xu\textsuperscript{b}, Youxiong Que\textsuperscript{b}, Hui Ling\textsuperscript{b,d}, Ze Yin\textsuperscript{a}, Ying Liu\textsuperscript{a} and Dinggang Zhou\textsuperscript{a,b,c} \textsuperscript{a}Hunan Key Laboratory of Economic Crops Genetic Improvement and Integrated Utilization, School of Life Science, Hunan University of Science and Technology, Xiangtan, PR China; \textsuperscript{b}Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture/National Engineering Research Center for Sugarcane, Ministry of Science and Technology, Fujian Agriculture and Forestry University, Fuzhou, PR China; \textsuperscript{c}Key Laboratory of Ecological Remediation and Safe Utilization of Heavy Metal-Polluted Soils, College of Hunan Province, Xiangtan, PR China; \textsuperscript{d}Yulin Normal University, Agricultural College, Yulin, PR China

\textbf{ABSTRACT} 

The complete mitogenome of Saccharum spp. hybrid 15a-53 was determined in this study, which contains two distinct circular chromosomes, Chromosome 1 and 2. The length of Chromosome 1 is 300,848 bp with the GC content of 43.93\%, while Chromosome 2 is 144,713 bp in length with the GC content of 43.57\%. In Chromosome 1, 7.14\% of the genome (21,468 nucleotides) is coding DNA and 92.86\% (279,380 nucleotides) are intergenic region, while in Chromosome 2, 8.20\% of genome (11,865 nucleotides) are coding DNA and 91.80\% (132,848 nucleotides) are intergenic region. Chromosome 1 contains 20 protein-coding genes (three \textit{atp} genes, three \textit{ccm} genes, two \textit{cox} genes, one \textit{mat} gene, one \textit{mtt} gene, six \textit{nad} genes, and four \textit{rps} genes), and 21 non-coding genes (15 tRNA and six rRNAs), while in Chromosome 2, there are 13 protein-coding genes (four \textit{nad} genes, three \textit{rps} genes, two \textit{atp} genes, one \textit{ccm} gene, one \textit{cob} gene, one \textit{cox} gene, and one \textit{rpl} gene) and five tRNA genes. Maximum Likelihood phylogenetic analysis indicated that 15a-53 is close to \textit{S.} spp. hybrid ROC22, \textit{S.} spp. hybrid FN15 and \textit{S.} officinarum Khon Kaen 3. The complete mitogenome herein will provide useful sequence information for phylogenetic and evolutionary studies for Saccharum and Poaceae.

Sugarcane (\textit{Saccharum} spp. hybrids complex) is the most important commercial crop for the production of sucrose and ethanol in the worldwide (Garsmeur et al. 2018; Liu et al. 2020; Zhou et al. 2020). Modern commercial sugarcane varieties are all complex interspecies hybrids at least from the three species \textit{S.} spontaneum, \textit{S.} robustum, and \textit{S.} officinarum, so the modern sugarcane cultivars exhibit an exceedingly complex interspecific aneupolyploid genome, so ‘\textit{Saccharum} spp. hybrids’ is a general term for modern sugarcane (Garsmeur et al. 2018). Insect attack is a major issue in modern sugarcane cultivation. However, traditional cross-breeding is almost impossible in improving the insect resistance of sugarcane, due to lack of insect-resistant sugarcane germplasms and the characterization of highly complex polyploid–aneuploids with huge genome and large chromosome numbers of modern commercial sugarcane cultivars (Zhou et al. 2018). Genetic engineering provides an alternative (Zhou et al. 2018). \textit{Saccharum} spp. hybrid 15a-53 is an insect-resistant transgenic \textit{cry1Ac} line from receptor variety ROC22 (a \textit{Saccharum} spp. hybrid from ROC5 × 69-463, the most widely cultivated sugarcane commercial variety during the past two decades in China) (Gao et al. 2016; Zhou et al. 2018), which could facilitate the development of insect-resistant sugarcane, and serves as germplasm for use in cross-breeding. The characterization of the complete mitogenome of sugarcane line 15a-53 and its phylogenetic relationship within Poaceae were described in this study.

The complete mitochondrial genome of the sugarcane \textit{S.} spp. hybrid 15a-53 was sequenced by Illumina HiSeq XTen and PacBio Sequel platform, assembled into the complete mitochondrial genome by SPAdes version 3.10.1 (Antipov et al. 2016), annotated by GeSeq (Tillich et al. 2017), and submitted to GenBank with the accession numbers of MT821853 (Chromosome 1) and MT821854 (Chromosome 2). The mitochondrial DNA (mtDNA) was extracted and purified from fresh yellowing seedlings of a single individual sugarcane line \textit{S.} spp. hybrid 15a-53 (Chen et al. 2011), which from Fujian Agriculture and Forestry University, Fuzhou, Fujian Province (geographic coordinates: 26°9'8"N, 119°24'24"E), China. The specimen of 15a-53 was stored in the Key Laboratory of Sugarcane Biology and Genetic Breeding, Fujian Agriculture and Forestry University with store number 15a-53-FJ2016003.

The complete mitogenome of 15a-53 contains two distinct circular chromosomes, Chromosome 1 and 2. The Chromosome 1 is 300,848 bp in length with the GC content...
of 43.93%, and 7.14% of genome (21,468 nucleotides) are coding DNA while 92.86% of genome (279,380 nucleotides) are intergenic region. Chromosome 1 contains 20 PCGs (protein-coding genes, three atp genes, three ccm genes, two cox genes, one mat gene, one mtt gene, six nad genes and four rps genes), and 15 tRNA and six rRNAs non-coding genes. All these PCGs in Chromosome 1 use the initiation codon ATG except for nad1 and matR, which begin with ACG and ATA, respectively. Regarding the stop codon of the PCGs in Chromosome 1, matR, ccmC, ccmFn, mttB, nad6, nad7, and rps1 terminate with TAG; atp4, atp8, cox1, cox2, nad1, nad9, and rps7 terminate with TAA; atp1, ccmB, rps2 and rps13 terminate with TGA, while nad2 and nad5 stop with CGG and GTA, respectively. The Chromosome 2 is 144,713 bp in length with the GC content of 43.57%, and 8.20% of genome (11,865 nucleotides) are coding DNA and 91.80% of genome (132,848 nucleotides) are intergenic region. Chromosome 2 contains 13 protein-coding genes (four nad genes, three rps genes, two atp genes, one ccm gene, one cob gene, one cox gene, and one rpl gene) and five tRNA genes. All the PCGs in Chromosome 2 start with ATG except for nad2 and nad5, which begin with TTG and CCA, respectively, and atp9, rps3 and cob stop with TAG, while ccmFc, cox3, nad4 and rps12 stop with TGA, and the other genes (atp6, nad2, nad3, nad5, rps4 and rpl16) terminate with TAA.

The maximum likelihood phylogenetic tree with 1000 bootstrap replications was generated using PhyML version 3.0 (http://www.atgc-montpellier.fr/phyml/) based on the complete mitogenomes of sugarcane variety S. spp. hybrid 15a-53, eight other species from the family Poaceae, and two species from the family Cruciferae. GenBank accession numbers are as follows: S. spp. hybrid FN15 (MT411890 and MT411891), S. spp. hybrid ROC22 (SRR11358604), Sorghum bicolor (NC_008360.1), Zea mays (NC_007982.1), Oryza sativa (NC_011033.1), Triticum aestivum (NC_037304.1), Hordeum vulgare (IBSC_v2.dna Mt:1:525599:1), Arabidopsis thaliana (NC_037304.1), and Brassica napus (NC_008285.1). A. thaliana and B. napus were used as outgroups. The phylogenetic tree showed that S. spp. hybrid 15a-53 is very close to S. spp. hybrid ROC22, S. spp. hybrid FN15 and S. officinarum Khon Kaen 3 (Figure 1). The complete mitochondrial genome herein will provide useful sequences information for phylogenetic and evolutionary studies for Saccharum and Poaceae.

**Disclosure statement**

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

**Funding**

501100001809This work was supported by National Natural Science Foundation of China [Grant No. 31701491], the Natural Science Foundation of Hunan Province, China [Grant No. 2019JJ50176], the Open Funding of Hunan Key Laboratory of Economic Crops Genetic Improvement and Integrated Utilization [Grant No. E22009, E22011], the Scientific Research Funding of Hunan University of Science and Technology [Grant No. ES1761].

**ORCID**

Dinggang Zhou [http://orcid.org/0000-0001-5665-9312](http://orcid.org/0000-0001-5665-9312)

**Data availability statement**

The data that support the findings of this study are available. The mitochondrial genome sequences rawdata of S. spp. hybrid 15a-53 were deposited in SRA with the accession number: SRR11358601 at the URL (https://www.ncbi.nlm.nih.gov/sra/?term=SRR11358601). The mitochondrial genome sequences of analyzed species (S. spp. hybrid FN15 and S. spp. hybrid ROC22; S. bicolor, Z. mays, O. sativa, T. aestivum, S. officinarum Khon Kaen 3, A. thaliana, and B. napus) were from the NCBI GenBank databases (the URL https://www.ncbi.nlm.nih.gov/genbank/)
and the mitogenome of *H. vulgare* is available in Ensembl (http://asia.ensembl.org/index.html).

**References**

Antipov D, Korobeynikov A, McLean JS, Pevzner PA. 2016. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. Bioinformatics. 32(7):1009–1015.

Chen J, Guan R, Chang S, Du T, Zhang H, Xing H. 2011. Substoichiometrically different mitotypes coexist in mitochondrial genomes of *Brassica napus* L. PLoS One. 6(3):e17662.

Gao S, Yang Y, Wang C, Guo J, Zhou D, Wu Q, Su Y, Xu L, Que Y. 2016. Transgenic sugarcane with a *cry1Ac* gene exhibited better phenotypic traits and enhanced resistance against sugarcane borer. PLoS One. 11(4):e0153929.

Garsmeur O, Droc G, Antonise R, Grimwood J, Potier B, Aitken K, Jenkins J, Martin G, Charron C, Hervouet C, et al. 2018. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. Nat Commun. 9(1):2638.

Liu X, Yin Z, Liu Y, Li Z, Yan M, Que Y, Xu L, Zhou D. 2020. The complete mitochondrial genome of sugarcane (*Saccharum spp.* variety FN15. Mitochondrial DNA B. 5(3):2163–2165.

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.

Zhou D, Liu X, Gao S, Guo J, Su Y, Ling H, Wang C, Li Z, Xu L, Que Y. 2018. Foreign *cry1Ac* gene integration and endogenous borer stress-related genes synergistically improve insect resistance in sugarcane. BMC Plant Biol. 18(1):342.

Zhou D, Yin Z, Liu X, Li Z, Yan M, Que Y, Xu L. 2020. The complete mitochondrial genome sequence and phylogenetic analysis of sugarcane (*Saccharum spp.*) cultivar ROC22. Mitochondrial DNA B. 5(2):1915–1916.