

The complete mitochondrial genome of *Luffa acutangula*

Panthita Ruang-areerate^a, Jeremy Shearman^a, Wasitthee Kongkachana^a, Nukoon Jomchai^a, Thippawan Yoocha^a, Sonicha U-thoomporn^a, Nattapol Narong^a, John Robert Sheedy^b, Supat Mekiyanon^b, Wirulda Pootakham^a and Sithichoke Tangphatsornruang^a

^aNational Omics Center, National Science and Technology Development Agency (NSTDA), Pathumthani, Thailand; ^bChia Tai Company Limited, Phra Khanong District, Bangkok, Thailand

ABSTRACT

Based on PacBio *de novo* assembly, we report the first complete mitochondrial genome of *Luffa acutangula* (460,333 bp) containing nine large chloroplast-derived sequences (1.9–17.3 kb) across the mitogenome. The base composition of the mitogenome in descending order is A: 28.02%, C: 22.04%, G: 21.83% and T: 28.10%, and the G + C content is 43.87%. There are 63 mitochondrial genes including 40 protein-coding genes, 3 rRNA genes and 20 tRNA genes. Additionally, a total of 288 repeats ranging from 31 to 5,301 bp were identified, accounting for 5.7% of the mitogenome. Two large direct repeats (5,301 and 405 bp) within the mitogenome were found for the formation of four subgenomic molecules. A phylogenetic analysis showed that *L. acutangula* was closely related to other species in Cucurbitaceae. This mitogenome provides useful genetic information for evolutionary studies.

ARTICLE HISTORY

Received 31 July 2020
Accepted 7 August 2020

KEYWORDS

Luffa; Cucurbitaceae; mitochondrial genome; chloroplast-derived sequences

Luffa acutangula (ridge gourd), a member of the family Cucurbitaceae (cucurbits), is a commercially important vegetable crop, especially in Asian countries. It is used in combination with other traditional medicines to treat symptoms such as diabetes, liver diseases and skin diseases (Shendge and Belemkar 2018). The complete mitochondrial genome of *L. acutangula* is an important resource for evolutionary studies. Among angiosperm families, Cucurbitaceae are among the most variable in mitogenome sizes (Ward et al. 1981). There are four available mitogenomes in the cucurbits, which is *Cucumis melo* (melon), *Cucumis sativus* (cucumber), *Cucurbita pepo* (zucchini) and *Citrullus lanatus* (watermelon) at approximately ~2.74, 1.68, 0.98 and 0.38 Mbp, respectively (Alverson et al. 2011, 2010; Rodríguez-Moreno et al., 2011).

Based on PacBio data, the complete mitochondrial DNA sequence of *L. acutangula* (GenBank accession MT374097.1) was assembled into a master circular DNA of 460,333 bp using Canu version 1.9 (Koren et al. 2017). In comparison with the available cucurbit mitogenomes, the *L. acutangula* mitogenome (~0.46 Mb) is larger than that of watermelon (~0.38 Mb), but smaller than that of zucchini (~0.98 Mb), cucumber (~1.68 Mb) and melon (~2.74 Mb) (Alverson et al. 2011, 2010; Rodríguez-Moreno et al. 2011). The sample of *L. acutangula* young leaves was collected and acquired from Chia Tai Company Limited (Thailand). The seed of this sample was also stored in Chia Tai Company Limited, Thailand with no. AG04. Its genomic DNA was sequenced in house using the PacBio RSII platform. The base composition of the

mitogenome for A, C, G and T is 28.02, 22.04, 21.83 and 28.10%, respectively. The G + C content of the mitogenome is 43.87%, which is nearly the same as the other cucurbits (Alverson et al. 2010, 2011; Rodríguez-Moreno et al. 2011). Using BLASTn, a total of 92.28 kb of chloroplast (cp)-derived DNA distributes across nine regions of the mitogenome ranging from 1.90 to 17.26 kb, five of which are >10 kb in length. The largest cp-derived segment (17.26 kb) in *L. acutangula* is slightly smaller than the largest cp-derived segment (up to 18.53 kb) in the zucchini mitogenome (Alverson et al. 2010).

Sixty-three genes were identified on the mitogenome of *L. acutangula* using MITOFY (Alverson et al. 2010) and tRNAscan-SE (Schattner et al. 2005), including 40 protein-coding genes, 3 rRNA genes and 20 tRNA genes. Among these protein-coding genes, *atp1* and *rps19* are present in two copies. *Nad1*, *nad2*, *nad5* and *nad7* contain four introns, while *nad4* contains 3 introns, and *cox2*, *ccmFc*, *rps3*, *rps10* and *rpl2* each contain one intron. The *nad1*, *nad2* and *nad5* genes include both cis- and trans-spliced forms. The *cox1* gene has no introns. Notably, *rps14* is a pseudogene in *L. acutangula* and the other cucurbits (Alverson et al. 2010, 2011; Rodríguez-Moreno et al. 2011). Additionally, there were 17 chloroplast-like tRNA genes in the *L. acutangula* mitogenome.

In addition to genes, a total of 288 repeats (121 groups) ranging from 31 to 5,301 bp (~5.7% of the mitogenome) were identified within the *L. acutangula* mitogenome using BLASTn against itself. Most of the repeats (242) are between 31 to 100 bp in length accounting for ~2.6% of the total

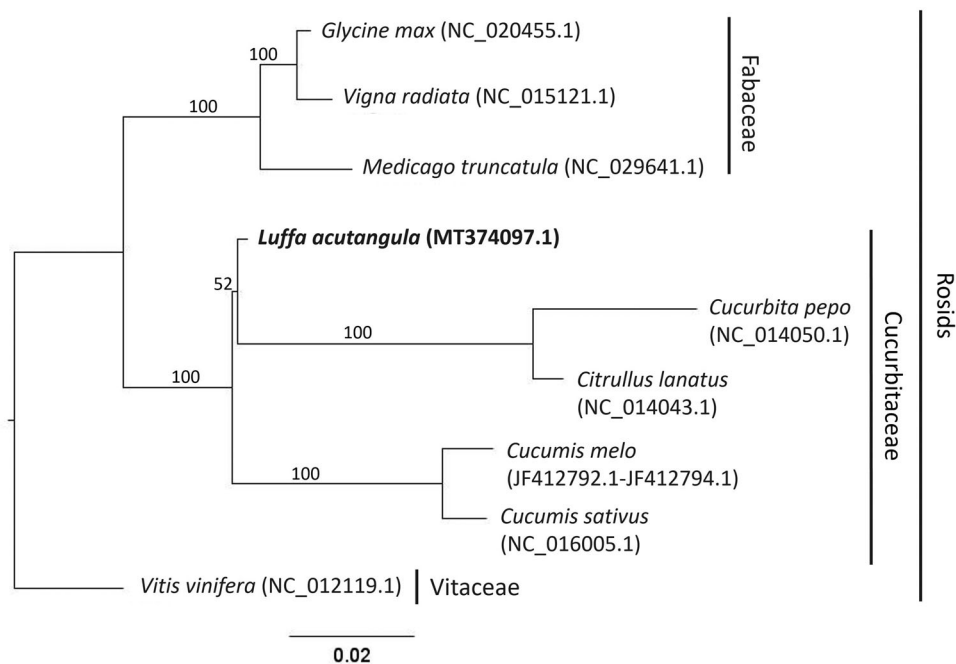


Figure 1. Phylogeny of the *Luffa acutangula* mitochondrial genome with eight plant species. Maximum likelihood phylogenetic tree was constructed based on the amino-acid sequences of 27 mitochondrial protein-coding genes with 1,000 bootstrap replicates using RAxML. Numbers in each of the node indicated the bootstrap support values.

mitogenome. Four circular-subgenomic molecules (119,030; 208,118; 252,215 and 341,303 bp) in *L. acutangula* were formed from two large direct repeats (5,301 bp (1–5,301 and 208,119–213,419) and 405 bp (284,435–284,838 and 403,465–403,869)) supported by the long read data.

Using RAxML version 8.2.10 with 1,000 bootstrap replicates (Stamatakis 2014), a maximum likelihood analysis was performed on eight plant mitogenomes based on amino acid sequences of 27 protein-coding genes including *atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFc*, *ccmFn*, *cob*, *cox1*, *cox3*, *matR*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad7*, *nad9*, *rpl5*, *rpl16*, *rps1*, *rps4*, *rps10* and *rps12*. The phylogenetic tree showed that *L. acutangula* was in the same paraphyletic clade with the four cucurbits including *C. pepo*, *C. lanatus*, *C. melo* and *C. sativus* with over 50 bootstrap values (Figure 1). These results provide insights on basic genetic information underlying sequence expansions of the *L. acutangula* mitogenome and important information on the evolution of Cucurbitaceae.

Disclosure statement

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

Funding

This study was supported by the National Omics Center under the National Science and Technology Development Agency (NSTDA), Thailand, grant number: 1000221. The authors are thankful to Prof. S. Seraphin (PAC, NSTDA) for fruitful discussion on manuscript preparation.

Data availability statement

The data that support the findings of this study are openly available in GenBank of NCBI at <https://www.ncbi.nlm.nih.gov/>, accession number: MT374097.1.

References

- Alverson AJ, Rice DW, Dickinson S, Barry K, Palmer JD. 2011. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell*. 23(7):2499–2513.
- Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. 2010. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol*. 27(6):1436–1448.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*. 27(5):722–736.
- Rodríguez-Moreno L, González VM, Benjak A, Martí MC, Puigdomènech P, Aranda MA, Garcia-Mas J. 2011. Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. *BMC Genomics*. 12:424–414.
- Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res*. 33(Web Server issue):W686–W689.
- Shendge PN, Belemkar S. 2018. Therapeutic potential of *Luffa acutangula*: a review on its traditional uses, phytochemistry, pharmacology and toxicological aspects. *Front Pharmacol*. 9:1177–1114.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9):1312–1313.
- Ward B, Anderson R, Bendich A. 1981. The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). *Cell*. 25(3):793–803.