



The complete chloroplast genome of *Artocarpus altilis* (Moraceae) and phylogenetic relationships

Ueric José Borges de Souza^a , Raíssa Nunes dos Santos^a, Renisson Neponuceno de Araújo Filho^b, Gil Rodrigues dos Santos^c, Renato Almeida Sarmento^{b,d}, Fabien De Bellis^{e,f} and Fabrício Souza Campos^a 

^aLaboratory of Bioinformatics & Biotechnology, Federal University of Tocantins, Gurupi, Brazil; ^bDepartment of Forestry Engineering, Federal University of Tocantins, Gurupi, Brazil; ^cLaboratory of Phytopathology, Federal University of Tocantins, Gurupi, Brazil; ^dPost-Graduate Program in Plant Production, Federal University of Tocantins, Gurupi, Brazil; ^eCIRAD, UMR AGAP Institute, Montpellier, France; ^fUMR AGAP Institute, University of Montpellier, CIRAD, INRAE, L'Institut Agro, Montpellier, France

ABSTRACT

The chloroplast (cp) is an essential organelle in higher plants. The genes of the plastome are well suited to infer phylogenetic relationships among taxa. In this study, we report the assembly of the cp genome of *Artocarpus altilis* and its phylogeny among species from Moraceae family. The cp genome of *A. altilis* was 160,822 base pair (bp) in length, comprising one large single-copy region of 88,692 bp, one small single-copy region of 19,290 bp, and a pair of inverted repeat regions (IRs) of 26,420 bp. A total of 113 different genes were predicted, including 79 protein-coding genes, 30 tRNA genes, and four rRNA genes. The phylogenetic analysis of 19 species belonging to the Moraceae family confirmed the phylogenetic proximity of the genus *Artocarpus* and *Morus* and the genetic similarity of *A. camansi* and *A. altilis*.

ARTICLE HISTORY

Received 24 March 2021
Accepted 15 June 2021

KEYWORDS



Plastid genome; plastome; phylogenetic relationships

The genus *Artocarpus* J.R. Forster and G. Forster is the most diverse of the tribe Artocarpeae and the third largest genus in the plant family Moraceae. This family includes approximately 70 species most of them native from the warm regions of South and Southeast Asia and Oceania (Zerega et al. 2010; Williams et al. 2017). Several species of *Artocarpus* are important food sources for forest fauna, and a dozen species are important crops in the regions in which they occur, such as the breadfruit (*A. altilis*), breadnut (*A. camansi* Blanco), jackfruit (*A. heterophyllus* Lam.), cempedak (*Artocarpus integer* (Thunb.) Merr.), and terap (*Artocarpus odoratissimus* Blanco) (Zerega et al. 2010). The breadfruit (*A. altilis*) is a cultigen originally derived from the wild progenitor *A. camansi*, a native species from New Guinea, where domestication was initiated (Zerega et al. 2004, 2005). The species is a staple food in tropical regions and is widely used as functional food because of the high nutritional and energy values (Gonçalves et al. 2020).

Nowadays, with the advent of next-generation sequencing technologies (NGS), the field of chloroplast (cp) genetics and genomics has expanded dramatically (Souza et al. 2019; Souza et al. 2020; Mehmood et al. 2020). At the present time, approximately 4896 complete plant cp genomes have been deposited as RefSeq in the NCBI Organelle Genome database (December 2020), although 17 of these species belong to the

Moraceae family and only two to the *Artocarpus* genus. In this study, we report the complete cp genome of *A. altilis* from whole-genome sequence data and the phylogenetic analysis to evaluate the relationships among species from the Moraceae family.

Illumina paired-end sequencing data of *A. altilis* were obtained from the NCBI Sequence Read Archive (Accession no. SRR3194007; BioSample: SAMN04508170). The plant sampling, library preparation, and parameters used for high throughput sequencing are available in De Bellis et al. (2016). Briefly, fresh leaf sample of *A. altilis* cultivar 'Novan' was collected in Malekula Island, Vanuatu. The specimen (voucher: VUT002) was provided by Dr Roger Malapa of the Vanuatu Agricultural Research and Technical Center (VARTC). Dr Malapa passed away a few years ago. A contact could be Mrs Juliane Kaoh, in charge of the VARTC horticultural division (kaohjuliane8@gmail.com) or Jean-Pierre Labouisse (jean-pierre.labouisse@cirad.fr). The VARTC genebank is located on the Island of Santo, Vanuatu (15.453°S, 167.184°E). Total genomic DNA was extracted according to the mixed alkyl trimethylammonium bromide (MATAB) protocol described by Risterucci et al. (2000) and used to generate the library with the Nextera DNA Library Preparation Kit (Illumina, San Diego, CA) according to the manufacturer's protocol. Paired-end sequencing was carried out at the Grand Plateau Technique

CONTACT Ueric José Borges de Souza  uericjose@gmail.com  Laboratory of Bioinformatics & Biotechnology, Federal University of Tocantins, Gurupi, Tocantins, Brazil

 Supplemental data for this article is available online at <https://doi.org/10.1080/23802359.2021.1945504>.

© 2021 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.

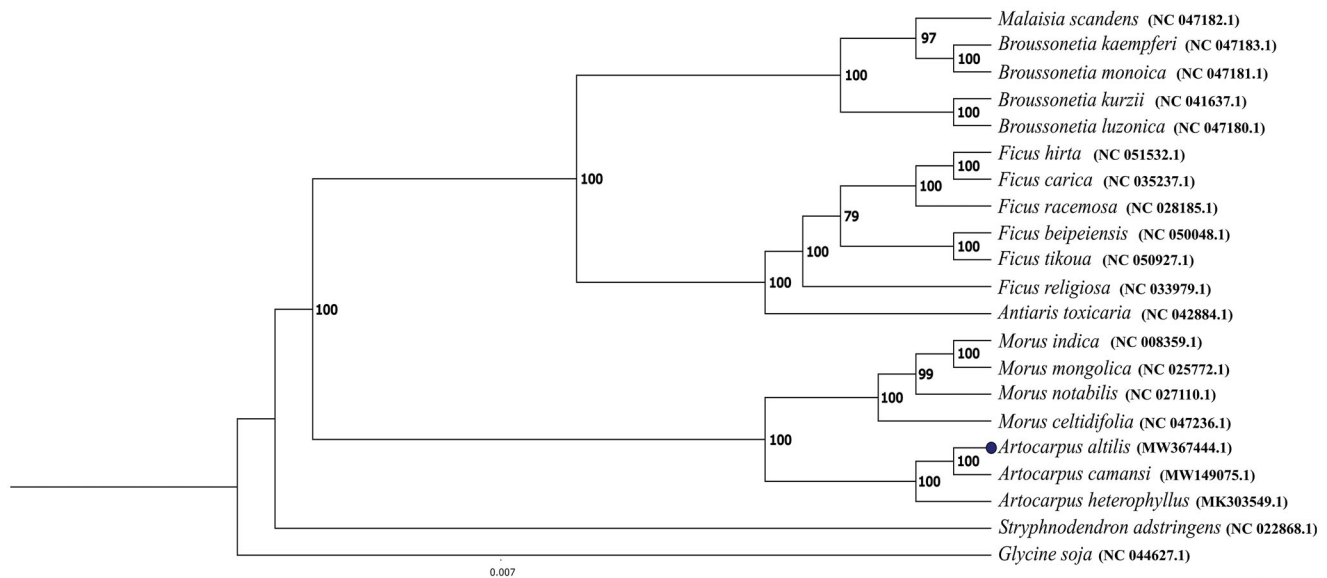


Figure 1. The phylogenetic tree of 19 species of the family Moraceae based on the sequences of 65 chloroplast genes. Bootstrap support values > 79% are given at the nodes.

Régional platform (Montpellier, France; <http://www.gptr-lr-genotypage.com>) on an Illumina MiSeq system using the MiSeq Reagent Kit version 3 (2 × 300 bp). The raw data (fastq files) was used to assemble the complete cp genome by using Fast-Plast pipeline version 1.2.8 with the Rosales order as the *bowtie_index* (McKain and Wilson 2017). Genome annotation was performed with Geseq (Tillich et al. 2017) and adjusted manually through comparisons with the annotations of *A. heterophyllus* (MK303549.1) and *A. camansi* (MW149075.1) using Geneious version 11.0.4 (Biomatters Ltd., Auckland, New Zealand).

The cp genome of *A. altilis* is 160,822 bp in length and presents the circular quadripartite structure typical of the angiosperms with a pair of inverted repeats regions (IRs) of 26,420 bp, separated by a large single-copy region (LSC) and a small single-copy region (SSC) of 88,692 and 19,290 bp, respectively (Supplementary Figure S1). The overall GC content (or guanine-cytosine content) was 36.0%. The cp genome of *A. altilis* was predicted to encode 113 different genes, with 79 protein-coding genes, 30 transfer RNA (tRNA), and four ribosomal RNA (rRNA) genes. Eighteen genes were duplicated completely in the IR region. Nine of the protein-coding genes and six tRNAs each contained one intron, while three genes each contained two introns. In general, the gene content, order, and organization of the *A. altilis* cp genome is highly similar to that of the closely related *A. camansi* (Souza et al. 2020) and *A. heterophyllus* (Liu et al. 2017), reported previously.

To conduct the phylogenetic analysis, 65 protein-coding genes were recorded in 19 species from the plant family Moraceae and two species (*Glycine soja* and *Stryphnodendron adstringens*) included as outgroups. All the genes were obtained from NCBI GenBank except those of *A. altilis* (see Supplementary Table S1 for species and accession numbers). The phylogenetic tree was constructed by maximum likelihood (ML) method using RAxML version 8 (Stamatakis 2014), with 1000 bootstrap replicates (Wang et al. 2018; Lei et al.

2020). A ML analysis yielded a tree topology with 79–100% bootstrap values at each node based on GTR + G + I model (Figure 1). The phylogeny supports the close relationship between *Artocarpus* and *Morus* genus and also reinforces the genetic similarity of the plastid genomes of *A. altilis* and *A. camansi*. Besides, more recently, a phylogenetic study showed that *A. camansi* and *A. mariannensis* form a clade with *A. altilis* (Williams et al. 2017), which is not surprising, since these two species have been successfully used in breeding with *A. altilis* (Jones et al. 2013).

Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that influenced the work reported in this article.

Funding

This research was supported by the Federal University of Tocantins and by the corresponding author's own resources. U.J.B.S and R.N.S. are CNPq DTI scholarships. This research received no external funding. We also would like to thank the Doctoral Program in Biodiversity and Biotechnology – BIONORTE Network (PPG-BIONORTE) - CAPES / AUXPE Process – PROAP 0889/2018.

ORCID

Ueric José Borges de Souza  <http://orcid.org/0000-0003-1442-4771>
 Fabrício Souza Campos  <http://orcid.org/0000-0002-5948-472X>

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 no. MW367444. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA312880, SRR3194007, and SAMN04508170, respectively.

References

- De Bellis F, Malapa R, Kagy V, Lebegin S, Billot C, Labouisse J. 2016. New development and validation of 50 SSR markers in breadfruit (*Artocarpus altilis*, Moraceae) by next-generation sequencing. *Appl Plant Sci.* 4(8):1600021.
- Gonçalves NGG, de Araújo JIF, Magalhães FEA, Mendes FRS, Lobo MDP, Moreira A, Moreira AR. 2020. Protein fraction from *Artocarpus altilis* pulp exhibits antioxidant properties and reverses anxiety behavior in adult zebrafish via the serotonergic system. *J Func Foods.* 66:103772.
- Jones AMP, Murch SJ, Wiseman J, Ragone D. 2013. Morphological diversity in breadfruit (*Artocarpus*, Moraceae): insights into domestication, conservation, and cultivar identification. *Genet Resour Crop Evol.* 60(1):175–192.
- Lei C, Jiyuan L, Zhengqi F, Hengfu Y, Xinlei L. 2020. Characterization and phylogenetic significance of the complete chloroplast genome of *Camellia Kissii*, an economic crop for producing oil. *Mitochondrial DNA Part B.* 5(1):362–363.
- Liu J, Niu YF, Ni SB, Liu ZY, Zheng C, Shi C. 2017. The complete chloroplast genome of *Artocarpus heterophyllus* (Moraceae). *Mitochondrial DNA Part B.* 3(1):13–14.
- McKain MR, Wilson M. 2017. Fast-Plast: rapid de novo assembly and finishing for whole chloroplast genomes. <https://github.com/mrmckain/Fast-Plast>.
- Mehmood F, Ubaid Z, Bao Y, Pocza P, Mirza B. 2020. Comparative plastomics of *Ashwagandha* (*Withania*, Solanaceae) and identification of mutational hotspots for barcoding medicinal plants. *Plants.* 9(6):752.
- Risterucci AM, Grivet L, N'Goran JA, Pieretti I, Flament MH, Lanaud C. 2000. A high-density linkage map of *Theobroma cacao* L. *Theor Appl Genet.* 101(5–6):948–955.
- Souza UJB, Nunes R, Targueta CP, Diniz-Filho JAF, Telles MPC. 2019. The complete chloroplast genome of *Stryphnodendron adstringens* (Leguminosae-caesalpinioideae): comparative analysis with related *Mimosoid* species. *Sci Rep.* 9(1):1–12.
- Souza UJB, Vitorino LC, Bessa LA, Silva FG. 2020. The complete plastid genome of *Artocarpus camansi*: a high degree of conservation of the plastome structure in the family Moraceae. *Forests.* 11(11):1179.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9):1312–1313.
- 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.
- Wang X, Cheng F, Rohlsen D, Bi C, Wang C, Xu Y, Wei S, Ye Q, Yin T, Ye N. 2018. Organellar genome assembly methods and comparative analysis of horticultural plants. *Hortic Res.* 5:3.
- Williams EW, Gardner EM, Harris IIR, Chaveerach A, Pereira JT, Zerega NJC. 2017. Out of Borneo: biogeography, phylogeny and divergence date estimates of *Artocarpus* (Moraceae). *Ann Bot.* 119(4):611–627.
- Zerega NJC, Ragone D, Motley TJ. 2004. Complex origins of breadfruit (*Artocarpus altilis*, Moraceae): implications for human migrations in Oceania. *Am J Bot.* 91(5):760–766.
- Zerega NJC, Ragone D, Motley TJ. 2005. Systematics and species limits of breadfruit (*Artocarpus*, Moraceae). *Syst Bot.* 30(3):603–615.
- Zerega NJC, Supardi N, Motley TJ. 2010. Phylogeny and recircumscription of Artocarpeae (Moraceae) with a focus on *Artocarpus*. *Syst Bot.* 35(4):766–782.