

MITOGENOME ANNOUNCEMENT



The complete mitochondrial genome of the oleaginous microalgae *Ankistrodesmus falcatus* strain UCP001 from the Peruvian Amazon

Marianela Cobos^a, Gad E. Grandez^{a,b} , J. Dylan Maddox^{a,c,d}, Carlos G. Castro^{a,b}, Hicler N. Rodríguez^{a,b}, Segundo L. Estela^a, Miguel A. Grández^a, Jae D. Paredes^a, Rodil Tello-Espinoza^e, Pedro M. Adrianzen^{b,f}, Jorge L. Marapara^{b,f} and Juan C. Castro^{b,f} 

^aLaboratorio de Biotecnología y Bioenergética (LBB), Universidad Científica del Perú (UCP), Loreto, Perú; ^bUnidad Especializada de Biotecnología, Centro de Investigaciones de Recursos Naturales de la Amazonía (CIRNA), Universidad Nacional de la Amazonia Peruana (UNAP), Loreto, Perú; ^cPritzker Laboratory for Molecular Systematics and Evolution, Field Museum of Natural History, Chicago, IL, USA; ^dEnvironmental Sciences, American Public University System, Charles Town, WV, USA; ^eDepartamento Académico de Manejo Forestal y Medio Ambiente, Facultad de Ciencias Forestales, Universidad Nacional de la Amazonia Peruana (UNAP), Ciudad Universitaria - Zungarococha, Loreto, Perú; ^fDepartamento Académico de Ciencias Biomédicas y Biotecnología, Facultad de Ciencias Biológicas, Universidad Nacional de la Amazonia Peruana (UNAP), Ciudad Universitaria - Zungarococha, Loreto, Perú

ABSTRACT

Ankistrodesmus falcatus strain UCP001 is a native oleaginous microalgae isolated from the Peruvian Amazon basin. In this study we sequenced, *de novo* assembled, and functionally annotated the complete mitochondrial genome of the native oleaginous microalgae *Ankistrodesmus falcatus* strain UCP001 (Accession number MT701044). This mitogenome is a typical circular double stranded DNA molecule of 41,048 bp in total length with G + C content of 37.4%. The mitogenome contains 49 genes, including 18 protein coding genes, 5 ribosomal (rRNA) genes and 26 transfer RNA (tRNA) genes. A phylogenetic analysis of 18 microalgae species indicated that *Ankistrodesmus falcatus* strain UCP001 was closely related to *Ourococcus multisporus* and *Raphidocelis subcapitata*. The complete mitochondrial genome sequence of *Ankistrodesmus falcatus* strain UCP001 enriches genomic resources of oleaginous native microalgae from the Peruvian Amazon for further basic and applied research.

ARTICLE HISTORY

Received 8 August 2020
Accepted 31 October 2020

KEYWORDS

Biofuel; chlorophyta; green algae; mitochondrial genome; molecular sequence annotation

Ankistrodesmus falcatus (Corda) Ralfs, 1848 strain UCP001 is a unicellular green microalgae of the Selenastraceae family isolated from the Itaya River (Cobos et al. 2017), which is a blackwater tributary river of the Amazon River in the northern region of Peru. This microalgae strain has shown a great capability to biosynthesize and store lipids (~264 mg/g of dry biomass) under nutritive stress (Ruiz et al. 2016). To date, however, genomic resources for this microalgae strain are scarce and the elucidation of their complete mitogenome will enhance our current understanding of the structure and organization of mitogenomes in green microalgae. In addition, the genomic sequence obtained would help to define, more precisely, the phylogenetic relationships among green microalgae. Finally, this genetic information could prove useful in the development of mitochondrial vector systems for biotechnology applications. Consequently, we sequenced, *de novo* assembled, and functionally annotated the complete mitochondrial genome of the oleaginous native microalgae *Ankistrodesmus falcatus* (Corda) Ralfs, 1848 strain UCP001.

Ankistrodesmus falcatus (Corda) Ralfs, 1848 strain UCP001 was provided by the Peruvian Amazon Microalgae Culture

Collection, Universidad Científica del Perú and cultured in CHU-10 medium until exponential growth phase. Microalgae cells were then harvested by centrifugation (5000 × g, 4 °C × 10 min), and total genomic DNA was purified using a standardized protocol. Libraries were constructed and sequenced with both PacBio RSII and Illumina HiSeq 2500 sequencing platforms. Low quality reads were removed with Trimmomatic v 0.38.0 (Bolger et al. 2014) and then filtered high quality reads, were *de novo* assembled with NOVOPlasty software v 4.0 (Dierckx et al. 2017). The complete circular mitogenome was annotated using a combination of the following bioinformatic tools: DOGMA (Wyman et al. 2004), GeSeq v 1.82 (Tillich et al. 2017), Mitofy v 3.0 (Alverson et al. 2010), tRNAScan-SE v 2.0 (Lowe and Chan 2016; Chan and Lowe 2019), and RNAweasel (Lang et al. 2007).

The complete circular mitogenome (Accession number MT701044) has a total length of 41,048 bp and G + C content of 37.4% that, contains 49 genes, including 18 protein coding genes, 5 ribosomal (rRNA) genes and 26 transfer RNA (tRNA) genes. Most protein coding genes are involved in the

CONTACT Marianela Cobos  mcobos@ucp.edu.pe  Laboratorio de Biotecnología y Bioenergética (LBB), Universidad Científica del Perú (UCP), Loreto, Perú; Juan C. Castro  juan.castro@unapikitos.edu.pe  Unidad Especializada de Biotecnología, Centro de Investigaciones de Recursos Naturales de la Amazonía (CIRNA), Universidad Nacional de la Amazonia Peruana (UNAP), Loreto, Perú

© 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.

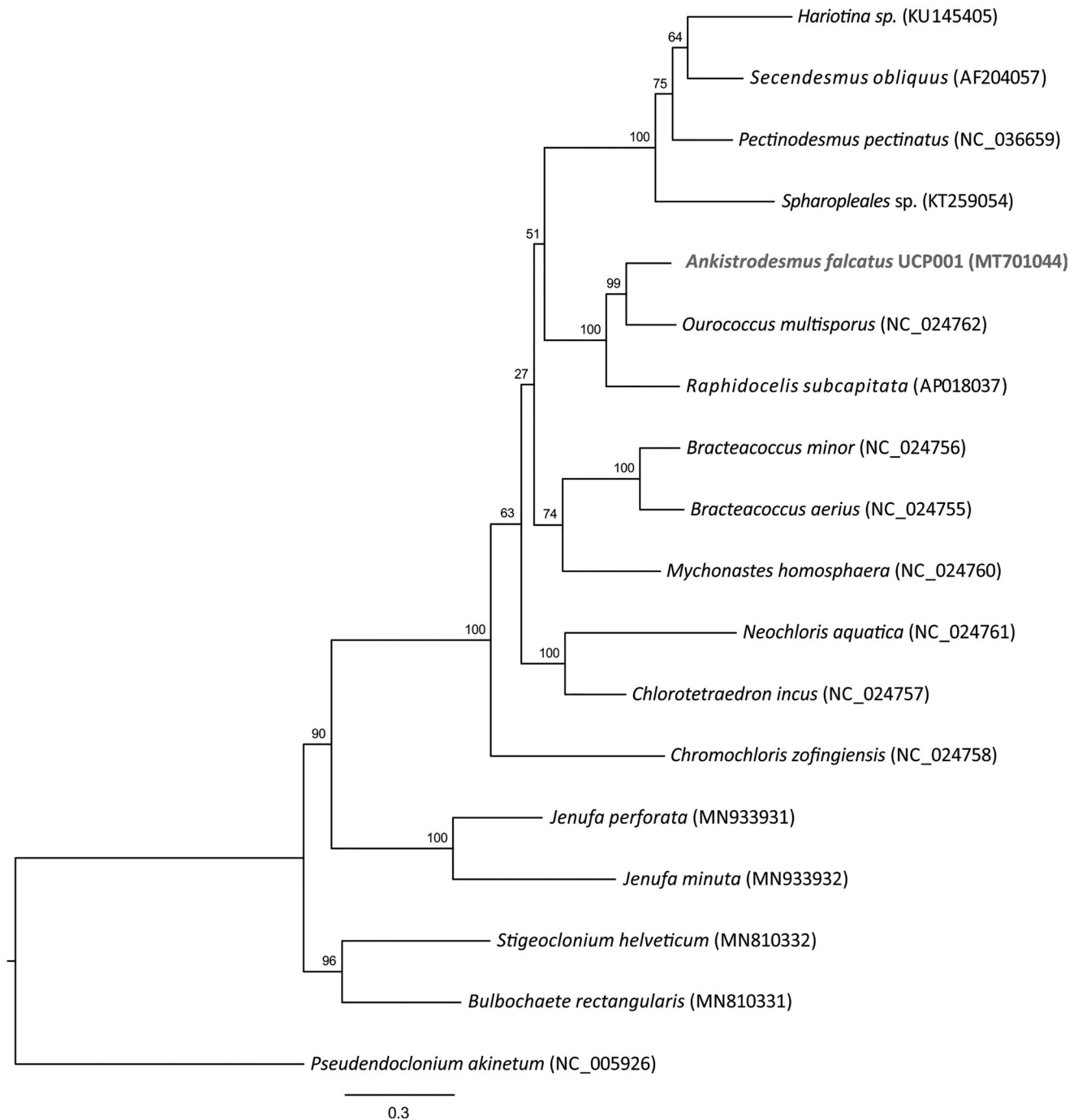


Figure 1. Maximum Likelihood (ML) phylogenetic tree of *Ankistrodesmus falcatus* strain UCP001 and 17 other green microalgae species based on 13 mitogenome proteins. Numbers on nodes indicate bootstrap support values based on 1000 replications. GenBank accession numbers are provided parenthetically.

mitochondrial electron transport and oxidative phosphorylation system. Also, we identified six intronic open-reading frames, encoding LAGLIDADG-2 Homing Endonucleases.

To infer the phylogenetic relationship of *Ankistrodesmus falcatus* UCP001 with other green microalgae, 17 additional mitogenomes were obtained from GenBank. In total, 13 genes (*atp6*, *atp9*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*) were extracted, aligned with MAFFT v 7.0 (Katoh et al. 2019) using default parameters, trimmed with trimAl (Capella-Gutiérrez et al. 2009) with automated

mode selection, and concatenated to produce a 12,499 bp alignment. PartitionFinder2 (Lanfear et al. 2017) was used to determine the best-fit partitioning scheme and models of evolution that were then used in RAXML-NG (Kozlov et al. 2019) with 10 random and 10 parsimony-based starting trees and 1000 bootstrap replications. The phylogenetic relationship indicated that *Ankistrodesmus falcatus* strain UCP001 belongs to a strongly supported monophyletic clade with *Ourococcus multisporus* and *Raphidocelis subcapitata* (Figure 1).

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding

This work was supported by Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica (FONDECYT) with grant contract No. 213-2015-FONDECYT. JDM was supported, in part, by a grant from Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica [44-2019-FONDECYT] during a portion of this research. Additionally, we thank the Laboratorio de Biotecnología y Bioenergética, Universidad Científica del Perú (UCP) and the Pritzker Laboratory for Molecular Systematics and Evolution, Field Museum of Natural History, and the Grainger Bioinformatics Center for providing necessary infrastructure and equipment.

ORCID

Gad E. Grandez  <http://orcid.org/0000-0001-5938-211X>
 Juan C. Castro  <http://orcid.org/0000-0003-0794-3178>

Data availability statement

Mitogenome data supporting this study are openly available in GenBank at: <https://www.ncbi.nlm.nih.gov/nuccore/MT701044>. Associated BioProject, SRA, and BioSample accession numbers are <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA628966>, <https://www.ncbi.nlm.nih.gov/sra/SRX9346510>, and SAMN16512711, respectively.

References

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.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30(15):2114–2120.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinforma Oxf Engl.* 25(15):1972–1973.
- Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol Biol Clifton NJ.* 1962:1–14.
- Cobos M, Paredes JD, Maddox JD, Vargas-Arana G, Flores L, Aguilar CP, Marapara JL, Castro JC. 2017. Isolation and characterization of native microalgae from the Peruvian Amazon with potential for biodiesel production. *Energies.* 10(2):224.
- Dierckx N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45(4):e18.
- Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 20(4):1160–1166.
- Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics.* 35(21):4453–4455.
- Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol.* 34(3):772–773.
- Lang BF, Laforest M-J, Burger G. 2007. Mitochondrial introns: a critical view. *Trends Genet.* 23(3):119–125.
- Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* 44(W1):W54–57.
- Ruiz MC, Rodríguez JDP, Gómez JCC. 2016. Inducción de la producción de lípidos totales en microalgas sometidas a estrés nutritivo. *Acta Biológica Colomb.* 21(1):17–26.
- 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.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organelle genomes with DOGMA. *Bioinformatics.* 20(17):3252–3255.