

Complete Chloroplast Genome Sequence of Phagomixotrophic Green Alga *Cymbomonas tetramitiformis*

Anchittha Satjarak,^a Amber E. Paasch,^b Linda E. Graham,^a Eunsoo Kim^b

Department of Botany, University of Wisconsin-Madison, Madison, Wisconsin, USA^a; Division of Invertebrate Zoology and Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, New York, USA^b

We report here the complete chloroplast genome sequence of *Cymbomonas tetramitiformis* strain PLY262, which is a prasinophycean green alga that retains a phagomixotrophic mode of nutrition. The genome is 84,524 bp in length, with a G+C content of 37%, and contains 3 rRNAs, 26 tRNAs, and 76 protein-coding genes.

Received 29 April 2016 Accepted 10 May 2016 Published 16 June 2016

Citation Satjarak A, Paasch AE, Graham LE, Kim E. 2016. Complete chloroplast genome sequence of phagomixotrophic green alga *Cymbomonas tetramitiformis*. *Genome Announc* 4(3):e00551-16. doi:10.1128/genomeA.00551-16.

Copyright © 2016 Satjarak et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Anchittha Satjarak, satjarak@wisc.edu.

Cymbomonas tetramitiformis is a marine prasinophycean green alga that bears a single chloroplast and lacks a cell wall (1). Within the prasinophytes, a paraphyletic assemblage of green algae that typically display features that are considered plesiomorphic for the *Chloroplastida* (green algae plus land plants), *C. tetramitiformis* falls into the order *Pyramimonadales* (prasinophyte clade I), which also includes *Halosphaera*, *Pyramimonas*, and *Pterosperma* (2). Unlike most other green algae, *C. tetramitiformis* retains the capacity to feed on bacteria while harvesting energy via photosynthesis (3).

C. tetramitiformis PLY262, acquired from the Plymouth Culture Collection of Marine Microalgae, was grown in f/2-Si at 18°C with a 12-h light/dark cycle. The total algal DNA was extracted using a Qiagen QIAamp minikit, according to the tissue-DNA protocol, and sent to the New York Genome Center (New York, NY) for whole-genome shotgun library preparation and sequencing on the Illumina MiSeq platform. A total of 19,183,050 paired sequences, each up to 300 bp in length, were generated. The raw reads were trimmed by Trimmomatic version 0.33 (4) in order to have a minimum quality score of 28 on the Phred 64 scale. The chloroplast genome was constructed by MIRA version 4.0.2 and MITObim version 1.8 (5) using the partial sequence of the *C. tetramitiformis* *rbcL* gene (GenBank accession no. L34687.1 [6]) as bait. The fold coverage of every position of the genome was calculated by mapping the trimmed reads to the newly assembled chloroplast genome using BWA nonmodel species alignment version 0.7.4 (7) and Bedtools genome coverage BAM version 2.19.1 (8) implemented in iPlant Collaborative (9). Open reading frames with length of at least 90 bp were predicted by conducting a BLAST search against the NCBI nonredundant protein database. tRNAs and rRNAs were predicted using tRNAscan-SE version 1.21 (10) and RNAmmer version 1.2 (11), respectively.

The complete chloroplast genome sequence was assembled into a circular-mapping molecule of 84,524 bp in length (1,200-fold coverage), with a G+C content of 37%. The genome was annotated with a total of 105 genes, including 76 protein-coding genes, 26 tRNAs, and 3 rRNAs. The region that contains three

rRNAs and two tRNAs appears twice in the genome in the inverted orientation, the structure that is also found in the chloroplast DNA (cpDNA) of other prasinophyte and streptophyte algae (12). No introns were detected. A notable feature of *C. tetramitiformis*, unlike the closely related *Pyramimonas parkeae* (13), is the absence of all three genes (*chlL*, *chlB*, and *chlN*) encoding subunits of light-independent protochlorophyllide oxidoreductase (LIPOR) either in the chloroplast or nuclear genome (GenBank accession no. LGRX000000000). In fact, the loss of this gene set has been reported from cpDNA of a diverse range of photosynthetic eukaryotes, including the prasinophytes *Micromonas pusilla* and *Ostreococcus tauri*, as well as members of chlorarachniophytes, euglenids, rhodophytes, cryptophytes, haptophytes, stramenopiles, dinoflagellates, and chromerids (14). The oxygen sensitivity of LIPOR (15) may at least partially explain such multiple independent losses of LIPOR genes in response to the increased level of atmospheric oxygen since the origins of eukaryotic algae.

Nucleotide sequence accession number. The assembled chloroplast genome sequence of *C. tetramitiformis* has been archived at GenBank with the accession number [KX013545](https://www.ncbi.nlm.nih.gov/nuccore/KX013545).

FUNDING INFORMATION

This work, including the efforts of Amber E Paasch, was funded by Richard Gilder Graduate School. This work, including the efforts of Eunsoo Kim, was funded by an NSF CAREER grant (1453639).

REFERENCES

- Graham LE, Graham JM, Wilcox LW, Cook ME. 2016. *Algae*, 3rd ed. LJLM Press, Madison, WI.
- Lewis LA, McCourt RM. 2004. Green algae and the origin of land plants. *Am J Bot* 91:1535–1556. <http://dx.doi.org/10.3732/ajb.91.10.1535>.
- Maruyama S, Kim E. 2013. A modern descendant of early green algal phagotrophs. *Curr Biol* 23:1081–1084. <http://dx.doi.org/10.1016/j.cub.2013.04.063>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <http://dx.doi.org/10.1093/bioinformatics/btu170>.
- Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a

- baiting and iterative mapping approach. *Nucleic Acids Res* 41:e129. <http://dx.doi.org/10.1093/nar/gkt371>.
6. Daugbjerg N, Moestrup Ø, Arcander P. 1994. Phylogeny of the genus *Pyramimonas* (Prasinophyceae, Chlorophyta) inferred from the *rbcL* gene. *J Phycol* 30:991–999. <http://dx.doi.org/10.1111/j.0022-3646.1994.00991.x>.
 7. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760. <http://dx.doi.org/10.1093/bioinformatics/btp324>.
 8. Quinlan AR. 2014. BEDTools: the Swiss-army tool for genome feature analysis. *Curr Protoc Bioinformatics* 47:11.12.1–11.12.34. <http://dx.doi.org/10.1002/0471250953.bil112s47>.
 9. Goff SA, Vaughn M, McKay S, Lyons E, Stapleton AE, Gessler D, Matasci N, Wang L, Hanlon M, Lenards A, Muir A, Merchant N, Lowry S, Mock S, Helmke M, Kubach A, Narro M, Hopkins N, Micklos D, Hilgert U. 2011. The iPlant collaborative: cyberinfrastructure for plant biology. *Front Plant Sci* 2:34. <http://dx.doi.org/10.3389/fpls.2011.00034>.
 10. 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(Suppl 2):W686–W689. <http://dx.doi.org/10.1093/nar/gki366>.
 11. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <http://dx.doi.org/10.1093/nar/gkm160>.
 12. Green BR. 2011. Chloroplast genomes of photosynthetic eukaryotes. *Plant J* 66:34–44. <http://dx.doi.org/10.1111/j.1365-313X.2011.04541.x>.
 13. Turmel M, Gagnon MC, O’Kelly CJ, Otis C, Lemieux C. 2009. The chloroplast genomes of the green algae *Pyramimonas*, *Monomastix*, and *Pycnococcus* shed new light on the evolutionary history of prasinophytes and the origin of the secondary chloroplasts of euglenids. *Mol Biol Evol* 26:631–648. <http://dx.doi.org/10.1093/molbev/msn285>.
 14. Hunsperger HM, Randhawa T, Cattolico RA. 2015. Extensive horizontal gene transfer, duplication, and loss of chlorophyll synthesis genes in the algae. *BMC Evol Biol* 15:16. <http://dx.doi.org/10.1186/s12862-015-0286-4>.
 15. Yamazaki S, Nomata J, Fujita Y. 2006. Differential operation of dual protochlorophyllide reductases for chlorophyll biosynthesis in response to environmental oxygen levels in the cyanobacterium *Lep- tolyngbya boryana*. *Plant Physiol* 142:911–922. <http://dx.doi.org/10.1104/pp.106.086090>.