




Draft Nuclear Genome, Complete Chloroplast Genome, and Complete Mitochondrial Genome for the Biofuel/Bioproduction Feedstock Species *Scenedesmus obliquus* Strain DOE0152z

S. R. Starkenburg,^a J. E. W. Polle,^{b,c} B. Hovde,^a H. E. Daligault,^a K. W. Davenport,^a A. Huang,^b P. Neofotis,^{b,c}  Z. McKie-Krisberg^b

Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA^a; Department of Biology, Brooklyn College of the City University of New York, Brooklyn, New York, USA^b; The Graduate Center of the City University of New York, New York, New York, USA^c

ABSTRACT The green alga *Scenedesmus obliquus* is an emerging platform species for the industrial production of biofuels. Here, we report the draft assembly and annotation for the nuclear, plastid, and mitochondrial genomes of *S. obliquus* strain DOE0152z.

The coccoid green alga *Scenedesmus obliquus* (Turpin) Kützing, which has the synonyms *Tetrademus obliquus* (Turpin) Wynne and *Acutodesmus obliquus* (Turpin) Hegewald & Hanagata and the basionym *Achnanthes obliqua* (Turpin), is a common freshwater alga (1, 2) in the family *Scenedesmaceae* within the *Chlorophyta* (= green algae). *S. obliquus* was previously investigated as a feedstock for protein production (3), but the species is now being cultivated for biofuel applications (4–12) and has also been suggested to be a source for edible oils (13).

Purified *S. obliquus* strain DOE0152z genomic DNA was sequenced and assembled using Pacific Biosciences (PacBio) (Menlo Park, CA, USA) long-read sequencing. Briefly, genomic DNA was size selected (20-kb DNA libraries) using Blue Pippin (Sage Science, Beverly, MA, USA) in a 0.75% agarose gel cassette and converted into a 20-kb single-molecule real-time (SMRT) bell library according to the manufacturer's instructions (PacBio). These size-selected libraries were sequenced on a PacBio RS II single-molecule sequencer (14) utilizing C3-P5 chemistry and 3-h movies. The library preparation protocol was repeated twice to generate two libraries/sequencing templates from the same genomic DNA pool. In total, 46 SMRT cells were sequenced, and these cells collectively produced 556,064 subreads with a mean subread length of 5,094 kb, which provided 18,115 Mbp of data or approximately 86× coverage of the assembled genome size. All PacBio subreads greater than 5 kb in length were assembled with HGAP version 2.3.0 (15). The mitochondrial and chloroplast genomes were assembled manually from fragmented contigs pulled from this initial assembly by mapping to previously sequenced *Scenedesmus* organellar genomes (GenBank numbers AF204057 and DQ396875, respectively). Chloroplast and mitochondrial genomes were extracted with DECONseq software, using a database constructed from previously published mitochondrial sequences for *S. obliquus* strain UTEX 78, and chloroplast sequences from *A. obliquus* (GenBank accession number DQ396875). Extracted contigs were mapped back onto reference sequences to ensure full coverage to plastid and mitochondrial genomes.

The final genome assembly resulted in 2,705 contigs with an N_{50} of 155,544 bp, a minimum contig size of 11,140 bp, a maximum contig size of 2,334,183 bp, and a total

Received 16 May 2017 Accepted 20 May 2017 Published 10 August 2017

Citation Starkenburg SR, Polle JEW, Hovde B, Daligault HE, Davenport KW, Huang A, Neofotis P, McKie-Krisberg Z. 2017. Draft nuclear genome, complete chloroplast genome, and complete mitochondrial genome for the biofuel/bioproduction feedstock species *Scenedesmus obliquus* strain DOE0152z. *Genome Announc* 5:e00617-17. <https://doi.org/10.1128/genomeA.00617-17>.

Copyright © 2017 Starkenburg 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 Z. McKie-Krisberg, zaid.MK@brooklyn.cuny.edu.

assembly size of 207,967,116 bp. Recently, a draft assembly of the nuclear genome was reported for *Tetradismus obliquus* strain UTEX393 with a size of approximately 109 Mbp (16). In contrast, the draft assembly of the nuclear genome of DOE0152z is almost twice the size at >210 Mbp. Furthermore, strain DOE0152z contains a plastid chromosome of 167,272 bp and the mitochondrial genome assembled to 41,704 bp. To the best of our knowledge, this is the first report of a nuclear, plastid, and mitochondrial genome assembly from the same *Scenedesmus* strain.

Accession number(s). This whole-genome project is publicly available on the LANL Greenhouse page (<https://greenhouse.lanl.gov/greenhouse/>) and was deposited in DDBJ/ENA/GenBank under the accession number [NEDT00000000](https://doi.org/10.1093/nucleic-acids/gnab001). The version described in this paper is version NEDT01000000.

ACKNOWLEDGMENT

This work was supported by funding from the Department of Energy under award numbers DE-SC0012556 and DE-NL0029949.

REFERENCES

- Kützing FT. 1834. Synopsis Diatomearum oder, Versuch einer systematischen Zusammenstellung der Diatomeen. Schwetschke, Halle, Germany.
- Smith GM. 1933. Fresh-water algae of the United States. McGraw-Hill, New York, NY.
- Becker EW. 1994. Microalgae: biotechnology and microbiology. Cambridge University Press, Cambridge, United Kingdom.
- Breuer G, Lamers PP, Martens DE, Draaisma RB, Wijffels RH. 2012. The impact of nitrogen starvation on the dynamics of triacylglycerol accumulation in nine microalgae strains. *Bioresour Technol* 124:217–226. <https://doi.org/10.1016/j.biortech.2012.08.003>.
- Breuer G, Lamers PP, Janssen M, Wijffels RH, Martens DE. 2015. Opportunities to improve the areal oil productivity of microalgae. *Bioresour Technol* 186:294–302. <https://doi.org/10.1016/j.biortech.2015.03.085>.
- Ferrigo D, Galla G, Sforza E, Morosinotto T, Barcaccia G, Ceschi Berrini CC. 2015. Biochemical characterization and genetic identity of an oil-rich *Acutodesmus obliquus* isolate. *J Appl Phycol* 27:149–161. <https://doi.org/10.1007/s10811-014-0315-5>.
- Neofotis P, Huang A, Sury K, Chang W, Joseph F, Gabr A, Twary S, Qiu W, Holguin O, Polle JEW. 2016. Characterization and classification of highly productive microalgae strains discovered for biofuel and bioproduct generation. *Algal Res* 15:164–178. <https://doi.org/10.1016/j.algal.2016.01.007>.
- Silva CM, Ferreira AF, Dias AP, Costa M. 2016. A comparison between microalgae virtual biorefinery arrangements for bio-oil production based on lab-scale results. *J Cleaner Product* 130:58–67. <https://doi.org/10.1016/j.jclepro.2015.09.053>.
- Singh SP, Singh P. 2014. Effect of CO₂ concentration on algal growth: a review. *Renew Sustain Energ Rev* 38:172–179. <https://doi.org/10.1016/j.rser.2014.05.043>.
- Song M, Pei H, Hu W, Ma G. 2013. Evaluation of the potential of 10 microalgal strains for biodiesel production. *Bioresour Technol* 141: 245–251. <https://doi.org/10.1016/j.biortech.2013.02.024>.
- Unkefer CJ, Sayre RT, Magnuson JK, Anderson DB, Baxter I, Blaby IK, Brown JK, Carleton M, Cattolico RA, Dale T, Devarenne TP, Downes CM, Dutcher SK, Fox DT, Goodenough U, Jaworski J, Holladay JE, Kramer DM, Koppisch AT, Lipton MS, Marrone BL, McCormick M, Molnár I, Mott JB, Ogden KL, Panisko EA, Pellegrini M, Polle J, Richardson JW, Sabarsky M, Starkenburg SR, Stormo GD, Teshima M, Twary SN, Unkefer PJ, Yuan JS, Olivares JA. 2017. Review of the algal biology program within the National Alliance for Advanced Biofuels and Bioproducts. *Algal Res* 22:187–215. <https://doi.org/10.1016/j.algal.2016.06.002>.
- Lammers PJ, Huesemann M, Boeing W, Anderson DB, Arnold RG, Bai X, Bhole M, Brhanavan Y, Brown L, Brown J, Brown JK, Chisholm S, Meghan Downes C, Fulbright S, Ge Y, Holladay JE, Ketheesan B, Khopkar A, Koushik A, Laur P, Marrone BL, Mott JB, Nirmalakhandan N, Ogden KL, Parsons RL, Polle J, Ryan RD, Samocha T, Sayre RT, Seger M, Selvaratnam T, Sui R, Thomasson A, Unc A, Van Voorhies W, Waller P, Yao Y, Olivares JA. 2017. Review of the cultivation program within the National Alliance for Advanced Biofuels and Bioproducts. *Algal Res* 22:166–186. <https://doi.org/10.1016/j.algal.2016.11.021>.
- Huang Y, Zhang D, Xue S, Wang M, Cong W. 2016. The potential of microalgae lipids for edible oil production. *Appl Biochem Biotechnol* 180:438–451. <https://doi.org/10.1007/s12010-016-2108-6>.
- Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebrea R. 2009. Real-time DNA sequencing from single polymerase molecules. *Science* 323:133–138. <https://doi.org/10.1126/science.1162986>.
- Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- Carreres BM, de Jaeger L, Springer J, Barbosa MJ, Breuer G, van den End EJ, Kleinegris DMM, Schäffers I, Wolbert EJH, Zhang H, Lamers PP, Draaisma RB, Martins Dos Santos VA, Wijffels RH, Eggink G, Schaap PJ, Martens DE. 2017. Draft genome sequence of the oleaginous green alga *Tetradismus obliquus* UTEX 393. *Genome Announc* 5(3):e01449-16. <https://doi.org/10.1128/genomeA.01449-16>.