



Draft Macronuclear Genome Sequence of the Ruminal Ciliate *Entodinium caudatum*

 T. Park,^a S. Wijeratne,^{b,c} T. Meulia,^{b,c} J. Firkins,^a  Z. Yu^a

^aDepartment of Animal Sciences, The Ohio State University, Columbus, Ohio, USA

^bMolecular and Cellular Imaging Center, Ohio Agricultural Research and Development Center, The Ohio State University, Wooster, Ohio, USA

^cDepartment of Plant Pathology, The Ohio State University, Wooster, Ohio, USA

ABSTRACT The transcriptionally active macronucleus of a ruminal ciliate, *Entodinium caudatum* MZG-1, was sequenced using the Illumina MiSeq and Oxford Nanopore MinION platforms. This is the first draft macronuclear genome sequence of a ruminal protozoan, and the genomic information will provide useful insight into the metabolism, physiology, and ecology of ruminal ciliates.

Entodinium caudatum, found only in the rumen or similar environments, is the most researched ruminal protozoal species because of its predominance (1) and high bacterivorous activity, which increases intraruminal nitrogen recycling and decreases nitrogen utilization efficiency (2–4). However, the inability to establish an axenic culture of any ruminal ciliate species has hindered understanding of their metabolism, physiology, and ecology and thus their actual roles in the rumen (5–8). Therefore, we sequenced the transcriptionally active macronuclear (MAC) genome of *E. caudatum* to help us understand its genomic and biological features.

E. caudatum strain MZG-1, initially isolated from the rumen of a gerenuk, was grown as a monoculture (in terms of ruminal protozoa) in SP medium (9). Cells of *E. caudatum* strain MZG-1 were separated from the associated prokaryotic cells present in the monoculture by sequential filtration and washing through four nylon filter membranes of decreasing pore size. Macronuclei were isolated following cell lysis (10, 11) and then purified using Percoll gradient centrifugation. Following confirmation by PCR amplification of the actin gene, MAC DNA was extracted using a QIAamp DNA minikit, and RNA was removed using DNase-free RNase. The MAC DNA was sequenced using the Illumina MiSeq and Oxford Nanopore MinION platforms. Approximately 40 million paired-end (2×300 -bp) reads (~6 Gb) were generated with the MiSeq platform, and 0.4 million 2D reads (~0.78 Gb) were generated with the MinION platform (12). Trimmomatic version 3.2.2 was used to trim the Illumina adapters and filter out the reads with a Phred quality score of <20 for a 4-bp window (13). The quality-checked MiSeq reads (73,209,090 in total) were then assembled *de novo* using SPAdes (14). The quality-filtered MinION 2D reads ($Q > 9$) and the contigs assembled from the MiSeq reads were assembled again using SSPACE-LongRead version 2.0 (15). The final assembled draft MAC genome sequence (107,579,855 bp) had a coverage reaching $79\times$ from a total of 20,400 contigs, with the N_{50} value being 9,873 bp and the longest contig being 147,117 bp. Putative telomeric repeats ($5'$ -CCCCAAT)_n were searched from the *E. caudatum* MAC scaffolds using SCAMPI (16), and 1,307 scaffolds were capped with the putative repeats at both ends, while 5,911 scaffolds were capped at one end. Approximately 100 bp of subtelomeric regions with 15-bp AT periodicity was detected.

Genes were predicted *de novo* using AUGUSTUS (17) trained with the *Tetrahymena thermophila* MAC genome sequences, followed by annotation of the protein-coding sequences using Blast2GO and InterProScan version 5 (18–20). In total, 15,544 different

Received 5 June 2018 Accepted 11 June 2018 Published 12 July 2018

Citation Park T, Wijeratne S, Meulia T, Firkins J, Yu Z. 2018. Draft macronuclear genome sequence of the ruminal ciliate *Entodinium caudatum*. Microbiol Resour Announc 7:e00826-18. <https://doi.org/10.1128/MRA.00826-18>.

Editor Jason Stajich, University of California, Riverside

Copyright © 2018 Park 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. Yu, yu.226@osu.edu.

protein-coding genes were annotated. The *E. caudatum* MAC genome sequence shared the most orthologs (698 in total) with that of *Oxytricha trifallax*, which has 16,000 chromosomes (21). Further analyses of the draft MAC genome and comparison with that of other ciliates will help us better understand the metabolism, physiology, and ecology of *E. caudatum* and its importance in rumen function and the interaction with other members of the rumen microbiome.

Data availability. This draft MAC genome of *E. caudatum* MZG-1 has been deposited in DDBJ/ENA/GenBank under the accession number [NBjL00000000](https://doi.org/10.1101/2020000000). The version described here is the second version, NBjL02000000.

ACKNOWLEDGMENT

This report is based upon work that was partially supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under award number 2012-67015-19437.

REFERENCES

- Dehority BA. 2005. Effect of pH on viability of *Entodinium caudatum*, *Entodinium exiguum*, *Epidinium caudatum*, and *Ophryoscolex purkynjei* in vitro. *J Eukaryot Microbiol* 52:339–342. <https://doi.org/10.1111/j.1550-7408.2005.00041.x>.
- Coleman GS, Sandford DC. 1979. The engulfment and digestion of mixed rumen bacteria and individual bacterial species by single and mixed species of rumen ciliate protozoa grown *in vivo*. *J Agric Sci* 92:729. <https://doi.org/10.1017/S0021859600053971>.
- de la Fuente G, Fondevila M, Belanche A, Morgavi D. 2011. *In vitro* predation of pure bacterial species by rumen protozoa from monofaunated sheep, determined by qPCR. *Options Mediterraneennes* 99:91–96.
- Belanche A, de la Fuente G, Moorby JM, Newbold CJ. 2012. Bacterial protein degradation by different rumen protozoal groups. *J Anim Sci* 90:4495–4504. <https://doi.org/10.2527/jas.2012-5118>.
- Coleman GS. 1962. The preparation and survival of almost bacteria-free suspensions of *Entodinium caudatum*. *J Gen Microbiol* 28:271–281. <https://doi.org/10.1099/00221287-28-2-271>.
- Hind T, Kametaka M. 1977. Gnotobiotic and axenic cultures of a rumen protozoan, *Entodinium caudatum*. *J Gen Appl Microbiol* 23:37–48. <https://doi.org/10.2323/jgam.23.37>.
- Onodera R, Henderson C. 1980. Growth factors of bacterial origin for the culture of the rumen oligotrich protozoan, *Entodinium caudatum*. *J Appl Bacteriol* 48:125–134. <https://doi.org/10.1111/j.1365-2672.1980.tb05214.x>.
- Fondevila M, Dehority BA. 2001. Preliminary study on the requirements of *Entodinium exiguum* and *E. caudatum* for live or dead bacteria when cultured *in vitro*. *Reprod Nutr Dev* 41:41–46. <https://doi.org/10.1051/rnd:2001110>.
- Dehority BA. 1998. Generation times of *Epidinium caudatum* and *Entodinium caudatum*, determined *in vitro* by transferring at various time intervals. *J Anim Sci* 76:1189–1196. <https://doi.org/10.2527/1998.7641189x>.
- Allen SL. 2000. Isolation of micronuclear and macronuclear DNA. *Methods Cell Biol* 62:241–252.
- Young KW, Thomas NA, Duval SM, Newbold CJ, Fried J, Brown DS, Hackstein JH, Michalowski T, McEwan NR. 2015. Isolation of macronuclear DNA from the rumen ciliate *Entodinium caudatum*. *Endocytobiosis Cell Res* 26:50–55.
- Mikheyev AS, Tin MM. 2014. A first look at the Oxford Nanopore MinION sequencer. *Mol Ecol Resour* 14:1097–1102. <https://doi.org/10.1111/1755-0998.12324>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Boetzer M, Pirovano W. 2014. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* 15:211. <https://doi.org/10.1186/1471-2105-15-211>.
- Bernsel A, Viklund H, Falk J, Lindahl E, von Heijne G, Elofsson A. 2008. Prediction of membrane-protein topology from first principles. *Proc Natl Acad Sci U S A* 105:7177–7181. <https://doi.org/10.1073/pnas.0711151105>.
- Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: *ab initio* prediction of alternative transcripts. *Nucleic Acids Res* 34:W435–W439. <https://doi.org/10.1093/nar/gkl200>.
- Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676. <https://doi.org/10.1093/bioinformatics/bti610>.
- Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36:3420–3435. <https://doi.org/10.1093/nar/gkn176>.
- Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. *Nucleic Acids Res* 33:W116–W120. <https://doi.org/10.1093/nar/gki442>.
- Swart EC, Bracht JR, Magrini V, Minx P, Chen X, Zhou Y, Khurana JS, Goldman AD, Nowacki M, Schotanus K, Jung S, Fulton RS, Ly A, McGrath S, Haub K, Wiggins JL, Storton D, Matese JC, Parsons L, Chang W-J, Bowen MS, Stover NA, Jones TA, Eddy SR, Herrick GA, Doak TG, Wilson RK, Mardis ER, Landweber LF. 2013. The *Oxytricha trifallax* macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. *PLoS Biol* 11:e1001473. <https://doi.org/10.1371/journal.pbio.1001473>.