Draft Genome Sequence of *Holospora undulata* Strain HU1, a Micronucleus-Specific Symbiont of the Ciliate *Paramecium caudatum*

Hideo Dohra,[a,b] Haruo Suzuki,[c] Tomohiro Suzuki,[a] Kenya Tanaka,[c,d] Masahiro Fujishima[c,d]

Instrumental Research Support Office, Research Institute of Green Science and Technology, Shizuoka University, Suruga-ku, Shizuoka, Japan[a]; Department of Biological Science, Graduate School of Science, Shizuoka University, Suruga-ku, Shizuoka, Japan[b]; Department of Environmental Science and Engineering, Graduate School of Science and Engineering, Yamaguchi University, Yoshida, Yamaguchi, Japan[c]; National BioResource Project of the Ministry of Education, Culture, Sports, Science and Technology, Chiyoda-ku, Tokyo, Japan[d]

Holospora undulata is a micronucleus-specific symbiont of the ciliate *Paramecium caudatum*. We report here the draft genome sequence of *H. undulata* strain HU1. This genome information will contribute to the study of symbiosis between *H. undulata* and the host *P. caudatum*.

Bacteria of the genus *Holospora* are intranuclear symbionts of *Paramecium* spp. and are Gram-negative bacteria belonging to the order *Rickettsiales* of the *Alphaproteobacteria* (1–5). *Holospora* currently consists of nine named species, and all species show host specificity and nucleus specificity in their habitats (6). *Holospora undulata* is a micronucleus-specific symbiont of the ciliate *Paramecium caudatum* (7). Here, we present the draft genome sequence of *H. undulata* strain HU1, which provides insight into the symbiotic strategy of this organism.

The infectious form of *H. undulata* cells was isolated from homogenates of the host cells at the early stationary phase of growth, using Percoll density gradient centrifugation (8), and kept until use at -80°C. The draft genome sequence of *H. undulata* HU1 was generated at the Instrumental Research Support Office, Research Institute of Green Science and Technology, Shizuoka University, Japan, using Illumina GAIIx technology. The genomic DNA of *H. undulata* was isolated using a DNeasy blood and tissue kit (Qiagen), fragmented using a Covaris Acoustic solubilizer, and an Illumina GAIIx paired-end (101-bp) library generated 77,451,934 reads totaling 7,822 Mb. The raw sequences were filtered by the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html) to get high-quality reads (cutoff quality score, 20; cutoff read length, <80 bp), resulting in 60,663,426 reads totaling 6,076 Mb. The high-quality reads were then assembled using ABySS version 1.3.5 (9), with a k-mer size of 72 bp, and contigs <200 bp in length were eliminated. The draft genome of *H. undulata* strain HU1 contains 452 contigs consisting of 1,512,931 bp, with a G+C content of 36.2% and an average 4,016× coverage of the total length of contigs.

The draft genome sequence was annotated using the National Center for Biotechnology Information (NCBI) Prokaryotic Genomes Automatic Annotation Pipeline. The annotated genome contains 1,420 protein-coding sequences, of which 650 (46%) are hypothetical proteins without any known functions. Among the 1,420 proteins, 857 proteins (60%) were assigned to different functional categories of NCBI Clusters of Ortholo-gous Groups (COG) (10). Eighty-six proteins were assigned to multiple COG categories. The most abundant COG category was "Replication, recombination, and repair" (186 proteins) followed by "Translation, ribosomal structure, and biogenesis" (125 proteins), and "Cell wall/membrane/envelope biogenesis" (86 proteins). This genome represents a valuable resource for future genomic studies.

Nucleotide sequence accession numbers. The sequencing results are archived in the DDBJ Sequence Read Archive (DRA) database with accession no. DRA001008. The *H. undulata* strain HU1 draft genome sequence has been deposited in GenBank with the accession no. ARPM00000000. The version described in this paper is the first version, ARPM01000000.

ACKNOWLEDGMENTS

REFERENCES

1. **Vakkerov-Kouzova ND, Rautian MS.** 2011. Obtaining and characterization of "*Holospora curviuscula*" and *Holospora obtusa*, bacterial symbionts of the macronuclei of *Paramecium bursaria* and *Paramecium caudatum*. Mikrobiologiya **80**:714–718. (In Russian.)
2. **Rautian MS, Wackerow-Kouzova ND.** 2013. Phylogenetic placement of two previously described intranuclear bacteria from the ciliate *Paramecium bursaria* (Protozoa, Ciliophora): '*Holospora acuminata*' and '*Holospora curviuscula*.' Int. J. Syst. Evol. Microbiol. **63**:1930–1933.
3. **Boscaro V, Fokin SI, Schrallhammer M, Schweikert M, Petroni G.** 2013.

Revised systematics of *Holospora*-like bacteria and characterization of "*Candidatus* Gortzia infectiva," a novel macronuclear symbiont of *Paramecium jenningsi*. Microb. Ecol. **65**:255–267.

4. **Amann R, Springer N, Ludwig W, Görtz HD, Schleifer KH.** 1991. Identification *in situ* and phylogeny of uncultured bacterial endosymbionts. Nature **351**:161–164.

5. **Fokin SI, Görtz HD.** 2009. Diversity of *Holospora* bacteria in *Paramecium* and their characterization, p 161–199. *In* Fujishima M (ed), Microbiology monographs, vol 12. Endosymbionts in *Paramecium*. Springer-Verlag, Berlin, Germany.

6. **Fokin SI, Brigge T, Brenner J, Görtz HD.** 1996. *Holospora* species infecting the nuclei of *Paramecium* appear to belong into two groups of bacteria. Eur. J. Protistol. **32**:19–24.

7. **Gromov BV, Ossipov DV.** 1981. *Holospora* (ex Hafkine 1890) nom. rev., a genus of bacteria inhabiting the nuclei of paramecia. Int. J. Syst. Evol. Microbiol. **31**:348–352.

8. **Fujishima M, Nagahara K, Kojima Y.** 1990. Changes in morphology, buoyant density and protein composition in differentiation from the reproductive short form to the infectious long form of *Holospora obtusa*, a macronucleus-specific symbiont of the ciliate *Paramecium caudatum*. Zool. Sci. **7**:849–860.

9. **Birol I, Jackman SD, Nielsen CB, Qian JQ, Varhol R, Stazyk G, Morin RD, Zhao Y, Hirst M, Schein JE, Horsman DE, Connors JM, Gascoyne RD, Marra MA, Jones SJ.** 2009. *De novo* transcriptome assembly with ABySS. Bioinformatics **25**:2872–2877.

10. **Tatusov RL, Galperin MY, Natale DA, Koonin EV.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28**:33–36.