

Draft Genome Sequence of a Single Cell of SAR86 Clade Subgroup IIIa

D. B. Rusch,^{a,b} M.-J. Lombardo,^c J. Yee-Greenbaum,^c M. Novotny,^c L. M. Brinkac,^a R. S. Lasken,^c C. L. Dupont^c

Informatics Group, J. Craig Venter Institute, Rockville, Maryland, USA^a; The Center for Genomics and Bioinformatics, Indiana University, Bloomington, Indiana, USA^b; Department of Microbial and Environmental Genomics, J. Craig Venter Institute, San Diego, California, USA^c

SAR86 denotes a 16S clade of gammaproteobacteria that are ubiquitous in ocean surface waters. So far, SAR86 is resistant to cultivation; thus, little is known about the genome contents or physiology of this clade. Recently, four partial genome sequences for SAR86 subclades I and II were published. Here, we present the draft genome sequence of a single cell from SAR86 subgroup IIIa isolated from coastal waters in San Diego, CA.

Received 10 October 2012 Accepted 26 October 2012 Published 15 January 2013

Citation: Rusch DB, Lombardo M-J, Yee-Greenbaum J, Novotny M, Brinkac LM, Lasken RS, Dupont CL. 2013. Draft genome sequence of a single cell of SAR86 clade subgroup IIIa. *Genome Announc.* 1(1):e00030-12. doi:10.1128/genomeA.00030-12.

Copyright © 2013 Rusch et al. This is an open-access article distributed under the terms of the [Attribution 3.0 Unported Creative Commons License](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to C. L. Dupont, cdupont@jvci.org.

The 16S clade SAR86 has been observed in the microbial communities of each of the major oceans (1–5). Multiple subgroups are present within the SAR86 16S phylogeny, including the subgroups I, II, IIIa, and IIIb, though it is not known whether these are related to differences in ecophysiology (6). The sequencing of an ~40-kbp bacterial artificial chromosome (BAC) and fosmid clones generated from a variety of locations revealed that some SAR86 genomes contain a gene that encodes a rhodopsin capable of generating a proton motive force (PMF) using light energy (proteorhodopsin [7–9]), with divergences between the subclades. Recently, two near-complete genome sequences from subclades I and II were assembled from metagenomic data and provided insight into the metabolic potential of SAR86 (10). The two subclades displayed distinct environmental distributions related to temperature, suggesting an ecophysiological link to a phylogenetic divergence, though representatives from other clades are necessary to make this conclusion.

Seawater was collected at the Scripps Institution of Oceanography (SIO) pier on 9 April 2009 at 8:30 a.m. The water temperature was 13.5°C, with a salinity of 33.5 practical salinity units (PSUs) and a chlorophyll fluorescence of 2.7 $\mu\text{g liter}^{-1}$ (SIO Automated Shore station equipment, <http://www.sccoos.org>). Single cells were sorted and amplified using multiple displacement amplification (MDA) technology as described by Dupont et al. (10).

The MDA reaction mixture was sequenced on one lane of an Illumina GAIIx. The reads were trimmed with Trimmomatic, and iterative *de novo* assemblies were performed using the CLC Genomics workbench. The read set was normalized to reduce coverage biases caused by MDA. To aid with gap closure, high-identity (>95% nucleotide identity) paired-end metagenomic reads from the Global Ocean Sampling Project (11) were aligned to the contigs to help orient and order them. Gap closure was performed using PCR and Sanger sequencing. Annotation was performed using the Prokaryotic Annotation Pipeline (12) with visualization and manual curation in Manatee. We have assem-

bled a genome of 7 scaffolds of 13 contigs totaling 1.40 Mbp, 1,442 putative coding sequences, 33 tRNAs, and 1 rRNA operon.

A phylogenetic analysis of the full-length 16S gene placed the single cell in SAR86 subclade IIIa, along with two BACs (accession no. AY033328 and AF268217); thus, we have named the genome SAR86_IIIb_sincell1. The genome contains many of the same features observed by Dupont et al. (10), including an overabundance of *tonB* receptors, a complete Emden-Meyerhof-Parnas glycolytic pathway, a green-light-absorbing proteorhodopsin, and the molecular machinery for dimethylsulfoniopropionate (DMSP) and glutathione assimilation. Like the other genomes, it also lacks several synthetic pathways, including those for vitamins and the carotenoids needed for light harvesting. Uniquely, SAR86_IIIb_sincell1 contains a hypervariable region with poor coverage in ocean metagenomes. This hypervariable region contains the genes involved in outer membrane modification, including those for putative glycosyltransferases, fucose acetylases, GDP-mannose dehydratases, and acetyltransferases. A similar scenario has been observed in the genomic islands of marine *Synechococcus* strains WH8102 and CC9311, where it has been suggested to alter the interactions with predators and phages (13, 14).

Nucleotide sequence accession number. This whole-genome shotgun project was deposited in NCBI under the accession no. [AMWX00000000](https://www.ncbi.nlm.nih.gov/bioproject/170317) (BioProject 170317).

ACKNOWLEDGMENTS

We acknowledge funding from the Gordon and Betty Moore Foundation (grant award agreement #521) and the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (DE-FC02-02ER63453). M.N., J.Y.-G., M.-J.L., and R.S.L. were supported by a grant from the Alfred P. Sloan Foundation. A NASA Astrobiology Institute Directors Discretionary Fund supported C.L.D.

REFERENCES

- Schattenhofer M, Fuchs BM, Amann R, Zubkov MV, Tarran GA, Pernthaler J. 2009. Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic ocean. *Environ. Microbiol.* 11:2078–2093.

2. Malmstrom RR, Straza TRA, Cottrell MT, Kirchmann DL. 2007. Diversity, abundance, and biomass production of bacterial groups in the western Arctic Ocean. *Aquat. Microb. Ecol.* 47:45–55.
3. González JM, Simó R, Massana R, Covert JS, Casamayor EO, Pedrós-Alíó C, Moran MA. 2000. Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl. Environ. Microbiol.* 66:4237–4246.
4. Mullins TD, Britschigi TB, Krest RL, Giovannoni SJ. 1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. *Limnol. Oceanogr.* 40:148–158.
5. Britschigi TB, Giovannoni SJ. 1991. Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. *Appl. Environ. Microbiol.* 57:1707–1713.
6. Treusch AH, Vergin KL, Finlay LA, Donatz MG, Burton RM, Carlson CA, Giovannoni SJ. 2009. Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME J.* 3:1148–1163.
7. Béjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, Jovanovich SB, Gates CM, Feldman RA, Spudich JL, Spudich EN, DeLong EF. 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* 289:1902–1906.
8. Sabehi G, Béjà O, Suzuki MT, Preston CM, DeLong EF. 2004. Different SAR86 subgroups harbour divergent proteorhodopsins. *Environ. Microbiol.* 6:903–910.
9. Sabehi G, Loy A, Jung KH, Partha R, Spudich JL, Isaacson T, Hirschberg J, Wagner M, Béjà O. 2005. New insights into metabolic properties of marine bacteria encoding proteorhodopsins. *PLoS Biol.* 3:e273.
10. Dupont CL, Rusch DB, Yooseph S, Lombardo MJ, Richter RA, Valas R, Novotny M, Yee-Greenbaum J, Selengut JD, Haft DH, Halpern AL, Lasken RS, Nealon K, Friedman R, Venter JC. 2012. Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J.* 6:1186–1199.
11. Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S, Wu D, Eisen JA, Hoffman JM, Remington K, Beeson K, Tran B, Smith H, Baden-Tillson H, Stewart C, Thorpe J, Freeman J, Andrews-Pfannkoch C, Venter JE, Li K, Kravitz S, Heidelberg JF, Utterback T, Rogers YH, Falcón LI, Souza V, Bonilla-Rosso G, Eguiarte LE, Karl DM, Sathyendranath S, Platt T, Bermingham E, Gallardo V, Tamayo-Castillo G, Ferrari MR, Strausberg RL, Nealon K, Friedman R, Frazier M, Venter JC. 2007. The *Sorcerer II* global ocean sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* 5:e77. doi:10.1371/journal.pbio.0050077.
12. Tanenbaum DM, Goll J, Murphy S, Kumar P, Zafar N, Thiagarajan M, Madupu R, Davidsen T, Kagan L, Kravitz S, Rusch DB, Yooseph S. 2010. The JCVI standard operating procedure for annotating prokaryotic metagenomic shotgun sequencing data. *Stand. Genomics Sci.* 2:229–237.
13. Palenik B, Ren Q, Dupont CL, Myers GS, Heidelberg JF, Badger JH, Madupu R, Nelson WC, Brinkac LM, Dodson RJ, Durkin AS, Daugherty SC, Sullivan SA, Khouri H, Mohamoud Y, Halpin R, Paulsen IT. 2006. Genome sequence of *Synechococcus* CC9311: insights into adaptation to a coastal environment. *Proc. Natl. Acad. Sci. USA* 103:13555–13559.
14. Palenik B, Brahamsha B, Larimer FW, Land M, Hauser L, Chain P, Lamerdin J, Regala W, Allen EE, McCarren J, Paulsen I, Dufresne A, Partensky F, Webb EA, Waterbury J. 2003. The genome of a motile marine *Synechococcus*. *Nature* 424:1037–1042.