JOURNAL OF BACTERIOLOGY, Mar. 2011, p. 1485–1486 Vol. 193, No. 6
0021-9193/11/$12.00 doi:10.1128/JB.01421-10

Complete Genome Sequence of *Algoriphagus* sp. PR1, Bacterial Prey of a Colony-Forming Choanoflagellate▽

Rosanna A. Alegado,[1]* Steven Ferriera,[2,3] Chad Nusbaum,[3] Sarah K. Young,[3] Qian Zeng,[3] Alma Imamovic,[3] Stephen R. Fairclough,[1] and Nicole King[1]

Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, California 94720[1]; J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, Maryland 20850[2]; and Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts 02139[3]

Received 25 November 2010/Accepted 13 December 2010

Bacteria are the primary food source of choanoflagellates, the closest known relatives of animals. Studying signaling interactions between the Gram-negative *Bacteroidetes* bacterium *Algoriphagus* sp. PR1 and its predator, the choanoflagellate *Salpingoeca rosetta*, provides a promising avenue for testing hypotheses regarding the involvement of bacteria in animal evolution. Here we announce the complete genome sequence of *Algoriphagus* sp. PR1 and initial findings from its annotation.

The marine *Bacteroidetes* species *Algoriphagus* sp. PR1 was coisolated with the choanoflagellate *Salpingoeca rosetta* from mud core samples near Hog Island, VA (13). *Bacteroidetes* species make up 6 to 30% of the total bacteria in the oceans (4, 11). Furthermore, they play an important role in the global carbon cycle because of their ability to degrade polysaccharides and other macromolecules (6, 8, 9, 22). Of the three clades that constitute the *Bacteroidetes* phylum (*Cytophaga*, *Flavobacteria*, and *Bacteroides*), the *Cytophaga* clade, of which *Algoriphagus* is a member, has been the least studied.

The complete genome sequence of *Algoriphagus* sp. PR1 was determined using shotgun sequencing, 454 (16), and Illumina technologies (2). Initial assembly of a draft whole-genome shotgun sequence into 12 contigs was generated at the J. Craig Venter Institute (JCVI) based upon 50,413 Sanger sequencing reads from genomic libraries harboring 4-kb and 40-kb fragments. Resequencing of *Algoriphagus* sp. PR1 was performed at the Broad Institute, and a 30× assembly containing a single gap was generated using the 454 Newbler assembler for 454 data (21) and the Velvet assembler (25) for Illumina data. The remaining gap is small and appears to be contained within a single gene.

The *Algoriphagus* sp. PR1 genome was found to be a single circular 4.89-Mbp chromosome that is 38.69% GC rich, contains 3,954 predicted genes, and is similar in size to previously sequenced genomes from other marine *Bacteroidetes* (1, 18–20). *Ab initio* gene models were generated using GeneMark (3), Glimmer3 (5), and Metagene (17). Predicted genes were generated from BLAST hits to the UniRef90 database, and a synteny-based approach was used to transfer open reading frames (ORFs) from the draft PR1 genome. The final ORF set was derived by comparison of *in silico* ORFs, ORFs from BLAST hits and mapped ORFs with hits to Pfam (10), and the top BLAST hits against UniRef90. ORFs with overlap relative to noncoding RNA features were removed when appropriate. Discrepancies in the final ORFs were resolved manually. Noncoding features were identified with RNAmmer (14), tRNAScan (15), and RFAM (12). There are 39 tRNAs and 9 rRNA operons. The genome contains genes required for a complete tricarboxylic acid cycle and complete glycolysis and pentose phosphate pathways. *Algoriphagus* sp. PR1 forms pink-pigmented colonies, and the genome encodes numerous carotenoid biosynthetic enzymes.

Given the capacity of *Bacteroidetes* bacteria to degrade macromolecules, we catalogued the diversity of carbohydrate-active enzymes in *Algoriphagus* sp. PR1. We found *Algoriphagus* sp. PR1 to have 62 glycoside hydrolases, 71 glycosyltransferases, 2 polysaccharide lyases, and 10 carbohydrate esterases, constituting a high capacity for polysaccharide degradation. While the expansion of these groups of enzymes is a characteristic of the *Bacteroidetes* phylum (1, 7, 23, 24), *Algoriphagus* sp. PR1 possesses a repertoire more similar to that of gut commensal *Bacteroidetes* than marine *Bacteroidetes*, which may in part be related to its interactions with choanoflagellates. The sequencing and annotation of the *Algoriphagus* sp. PR1 genome provide a foundation for comparative studies of microbe-eukaryote interactions.

Nucleotide sequence accession numbers. The JCVI genome sequence of *Algoriphagus* sp. PR1 is available in GenBank under accession number AAXU01000000, and the accession number for the Broad genome sequence is AAXU02000000.

The initial phase of sequencing, assembly, and annotation efforts was supported by a Gordon and Betty Moore Foundation Junior Investigator award (to N.K.) and the Gordon and Betty Moore Foundation Marine Microbial Sequencing Project. Resequencing and genome finishing at the Broad Institute were supported by funding from NHGRI/NIH as part of the Origins of Multicellularity Project. Subsequent data analysis was conducted at UC Berkeley and supported by an NIH National Research Service award and fellowship grant to R.A.A. (5F32GM086054). N.K. is a scholar in the Integrated Microbial Biodiversity Program of the Canadian Institute for Advanced Research.

* Corresponding author. Mailing address: Department of Molecular and Cell Biology, Life Sciences Addition, #3200, Berkeley, CA 94720-3200. Phone: (510) 643-9417. Fax: (510) 643-6791. E-mail: rosie.alegado@berkeley.edu.

▽ Published ahead of print on 23 December 2010.

REFERENCES

1. **Bauer, M., et al.** 2006. Whole genome analysis of the marine Bacteroidetes 'Gramella forsetii' reveals adaptations to degradation of polymeric organic matter. Environ. Microbiol. **8:**2201–2213.
2. **Bentley, D. R., et al.** 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature **456:**53–59.
3. **Borodovsky, M., and J. McIninch.** 1993. Recognition of genes in DNA sequence with ambiguities. Biosystems **30:**161–171.
4. **Cottrell, M. T., and D. L. Kirchman.** 2000. Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. Appl. Environ. Microbiol. **66:**1692–1697.
5. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636–4641.
6. **DeLong, E. F., D. G. Franks, and A. L. Alldredge.** 1993. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. Limnol. Oceanogr. **38:**924–934.
7. **Duchaud, E., et al.** 2007. Complete genome sequence of the fish pathogen Flavobacterium psychrophilum. Nat. Biotechnol. **25:**763–769.
8. **Fandino, L. B., L. Riemann, G. F. Steward, and F. Azam.** 2005. Populations dynamics of *Cytophaga-Flavobacteria* during marine phytoplankton blooms analyzed by real-time quantitative PCR. Aquat. Microb. Ecol. **40:**251–257.
9. **Fandino, L. B., L. Riemann, G. F. Steward, R. A. Long, and F. Azam.** 2001. Variations in bacterial community structure during a dinoflagellate bloom analyzed by DGGE and 16S rDNA sequencing. Aquat. Microb. Ecol. **23:** 119–130.
10. **Finn, R. D., et al.** 2006. Pfam: clans, web tools and services. Nucleic Acids Res. **34:**D247–D251.
11. **Glöckner, F. O., B. M. Fuchs, and R. Amann.** 1999. Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization. Appl. Environ. Microbiol. **65:**3721–3726.
12. **Griffiths-Jones, S., et al.** 2005. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res. **33:**D121–D124.
13. **King, N., C. T. Hittinger, and S. B. Carroll.** 2003. Evolution of key cell signaling and adhesion protein families predates animal origins. Science **301:**361–363.
14. **Lagesen, K., et al.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108.
15. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
16. **Margulies, M., et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.
17. **Noguchi, H., J. Park, and T. Takagi.** 2006. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. Nucleic Acids Res. **34:**5623–5630.
18. **Oh, H. M., I. Kang, S. Ferriera, S. J. Giovannoni, and J. C. Cho.** 2010. Complete genome sequence of *Croceibacter atlanticus* HTCC2559T. J. Bacteriol. **192:**4796–4797.
19. **Oh, H. M., et al.** 2011. Complete genome sequence of strain HTCC2170, a novel member of the genus *Maribacter* in the family *Flavobacteriaceae*. J. Bacteriol. **193:**303–304.
20. **Oh, H. M., et al.** 2009. Complete genome sequence of *Robiginitalea biformata* HTCC2501. J. Bacteriol. **191:**7144–7145.
21. **Quinn, N. L., et al.** 2008. Assessing the feasibility of GS FLX pyrosequencing for sequencing the Atlantic salmon genome. BMC Genomics **9:**404.
22. **Rath, J., K. Y. Wu, G. J. Herndl, and E. F. DeLong.** 1998. High phylogenetic diversity in a marine-snow-associated bacterial assemblage. Aquat. Microb. Ecol. **14:**261–269.
23. **Xie, G., et al.** 2007. Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. Appl. Environ. Microbiol. **73:**3536–3546.
24. **Xu, J., et al.** 2003. A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. Science **299:**2074–2076.
25. **Zerbino, D. R., and E. Birney.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18:**821–829.