

Genome Sequence of *Youngiibacter fragilis*, the Type Strain of the Genus *Youngiibacter*

Colin B. Wawrik,^a Amy V. Callaghan,^b Blake W. Stamps,^b Boris Wawrik^b

Franziskanergymnasium Kreuzburg, Großkrotzenburg, Germany^a; Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma, USA^b

The genome of *Youngiibacter fragilis*, the type strain of the newly described genus *Youngiibacter*, was sequenced. The genome consists of 3.996 Mb, with a G+C content of 46.6 mol%. *Y. fragilis* originates from coal-bed methane-produced water and may provide insight into the microbiological basis of biogas production in coal beds.

Received 10 December 2013 Accepted 17 December 2013 Published 23 January 2014

Citation Wawrik CB, Callaghan AV, Stamps BW, Wawrik B. 2014. Genome sequence of *Youngiibacter fragilis*, the type strain of the genus *Youngiibacter*. *Genome Announc.* 2(1): e01183-13. doi:10.1128/genomeA.01183-13.

Copyright © 2014 Wawrik et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Boris Wawrik, bwawrik@ou.edu.

Youngiibacter is a newly described genus of *Clostridiaceae* (1). Its type strain, *Youngiibacter fragilis*, is a strictly anaerobic, Gram-negative, non-spore-forming rod that ferments a range of carbohydrates to ethanol, formate, acetate, and CO₂. It was isolated from produced water of a coal-bed methane (CBM) production well in the Beluga River Gas Field, AK. CBM production emerged in the 1980s and has since become an important source of natural gas in the United States, China, Japan, Australia, and New Zealand (2). Geochemical and microbiological investigations have repeatedly demonstrated that biological production of methane plays an important role in the generation of methane in coal formations (2, 3). It is thought that a diverse microbial community mediates the initial activation of the complex coal matrix, and that subsequent fermentations of degradation products are coupled to the production of methane via close interactions between bacteria and methanogenic archaea (2). The biochemical mechanisms that govern anaerobic hydrocarbon activation and degradation are an area of active research. Genomic analysis of bacterial and archaeal isolates from such environments may shed light on the role that individual microbial community members play in subsurface organic matter transformations.

Whole-genome shotgun sequencing of *Y. fragilis* was conducted via Illumina MiSeq technology by generating a library with 350-bp inserts and sequencing 250-bp paired-ends. A total of 4.68×10^6 paired-end reads were obtained, yielding 2.34×10^9 bp of raw sequences. Reads were trimmed by removing adapters, retaining sequence data with quality scores of >30 . The surviving paired-end reads with sequence lengths of >100 bp (97.8% of reads) were assembled using CLC Genomics Workbench (CLC bio, Cambridge, MA). Contigs longer than 500 bp were considered for further analysis. Scaffolding of contigs was attempted using SSPACE (4) but did not improve the assembly. Assembly produced 242 contigs at ca. $350\times$ coverage. The N_{50} of the assembly is 41,656 bp, and the longest contig is 216,900 bp. The 16S rRNA gene was reconstructed from reads using EMIRGE (5), yielding a single 16S rRNA gene that was 100% identical to the known sequence of *Y. fragilis* (1). No evidence of plasmids was observed.

Annotation was conducted using the Rapid Annotation using Subsystem Technology (RAST) pipeline (<http://rast.nmpdr.org>). The genome contains 3,722 predicted proteins in 374 subsystems. Among genes with predicted functions, 23% are dedicated to carbohydrate metabolism, consistent with the phenotypic characterization of *Y. fragilis*. The genome lacks genes for dissimilatory sulfate reduction, nitrate reduction, flagellar motion, and chemotaxis. Given the origin of *Y. fragilis*, the genome was further investigated for the presence of genes involved in anaerobic hydrocarbon activation, with special focus on those relevant to coal degradation. Genes involved in known pathways for the anaerobic activation of alkanes, toluene, xylenes, ethylbenzene, phenol, acetophenone, benzene, and naphthalene were not detected. These data indicate that *Y. fragilis* is not likely to play a role in the initial activation of coal, but rather that it plays a role in the reprocessing and secondary fermentation of carbon in the subsurface.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [AXUN00000000](https://www.ncbi.nlm.nih.gov/nuccore/AXUN00000000). The version described in this paper is version AXUN02000000.

ACKNOWLEDGMENTS

We thank Henry Neeman and Jason Alexander at the OU Supercomputing Center for Education & Research (OSCER) for computing resources and their exceptional support with software deployment.

This work is supported by NSF grants OCE 0961900 and MCB 0921265.

REFERENCES

1. Lawson PA, Wawrik B, Allen TD, Johnson CN, Marks CR, Tanner RS, Harriman BH, Strąpoc D, Callaghan AV. 18 September 2013. *Youngiibacter fragilis* gen. nov., sp. nov., isolated from natural gas production-water and reclassification of *Acetivibrio multivorans* as *Youngiibacter multivorans* comb. nov. *Int. J. Syst. Evol. Microbiol.* <http://dx.doi.org/10.1099/ijs.0.053728-0>.
2. Strąpoc D, Mastalerz M, Dawson K, Macalady J, Callaghan AV, Wawrik B, Turich C, Ashby M. 2011. Biogeochemistry of microbial coal-bed methane. *Annu. Rev. Earth Planetary Sci.* 39:617–656. <http://dx.doi.org/10.1146/annurev-earth-040610-133343>.

3. Wawrik B, Mendivelso M, Parisi VA, Suflita JM, Davidova IA, Marks CR, Van Nostrand JD, Liang Y, Zhou J, Huizinga BJ, Strapoc D, Callaghan AV. 2012. Field and laboratory studies on the bioconversion of coal to methane in the San Juan Basin. *FEMS Microbiol. Ecol.* **81**:26–42. <http://dx.doi.org/10.1111/j.1574-6941.2011.01272.x>.
4. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**:578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
5. Miller CS, Baker BJ, Thomas BC, Singer SW, Banfield JF. 2011. EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol.* **12**:R44. <http://dx.doi.org/10.1186/gb-2011-12-5-r44>.