



Draft Genome Sequence of *Pseudomonas* sp. BDAL1 Reconstructed from a Bakken Shale Hydraulic Fracturing-Produced Water Storage Tank Metagenome

Daniel Lipus,^{a,c} Daniel Ross,^{c,d}  Kyle Bibby,^{a,b} Djuna Gulliver^c

Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania, USA^a;
Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA^b;
National Energy Technology Laboratory (NETL), Pittsburgh, Pennsylvania, USA^c; AECOM, Pittsburgh, Pennsylvania, USA^d

ABSTRACT We report the 5,425,832 bp draft genome of *Pseudomonas* sp. strain BDAL1, recovered from a Bakken shale hydraulic fracturing-produced water tank metagenome. Genome annotation revealed several key biofilm formation genes and osmotic stress response mechanisms necessary for survival in hydraulic fracturing-produced water.

Microbial activity associated with hydraulic fracturing-produced water and subsequently the hydraulic fracturing infrastructure is considered a significant operational concern due to the potential for corrosion, souring, and biofouling (1–4). Several studies have recently investigated the microbial community structure associated with hydraulic fracturing-produced water and hydraulic fracturing facilities (1, 5, 6). These investigations have provided interesting insights into the taxonomy of microbial populations associated with these environments. However, little is known about the functional potential and metabolic ability of these microorganisms that may result in interference with hydraulic fracturing operations.

Here, we present the draft genome of *Pseudomonas* sp. strain BDAL1, assembled from the metagenome of Bakken shale hydraulic fracturing-produced water, sampled from a produced water storage tank. Sequencing libraries were prepared using Illumina Nextera XT and sequenced using Illumina MiSeq technology (Illumina, San Diego, CA). Sequencing reads were quality trimmed (Q30) and *de novo* assembled into contigs using CLC-Genomics-Workbench version 8.5.1 (<http://www.clcbio.com/products/clc-genomics-workbench>) and SPAdes version 3.5.1 (7). Assembled contigs were grouped into genome bins with Vizbin (8) and taxonomy assessed with PhyloPythia (9). Metagenomic reads were mapped against binned contigs and reassembled using SPAdes.

The final draft genome contained 128 contigs of 5000 bp to 243,670 bp in length and an N_{50} of 63,183 bp. The total genome size was 5,425,832 bp with a mean G+C content of 58.4% and had an average of 19-fold coverage. Draft genome completeness and contamination were estimated using 833 *Pseudomonas* marker genes in CheckM (10). The final draft genome was found to be 97.4% complete and contain 0.4% contamination.

The draft genome was annotated by Rapid Annotations using Subsystems Technology (RAST) (11, 12) revealing 4,998 gene coding sequences (CDs) and 51 RNA sequences (48 tRNA, 16S, 23S, and 5S rRNA). Phylogenetic analysis of the recovered 16S rRNA gene sequence using BLASTn suggested that *Pseudomonas* sp. strain BDAL1 is closely related to *Pseudomonas syringae* pv. phaseolicola 1448A. Whole-genome alignment using an average nucleotide identity calculator confirmed the close phylogenetic relationship (99.9% nucleotide identity) (13, 14). RAST (11, 12) annotation allowed the

Received 10 January 2017 Accepted 12 January 2017 Published 16 March 2017

Citation Lipus D, Ross D, Bibby K, Gulliver D. 2017. Draft genome sequence of *Pseudomonas* sp. BDAL1 reconstructed from a Bakken shale hydraulic fracturing-produced water storage tank metagenome. Genome Announc 5: e00033-17. <https://doi.org/10.1128/genomeA.00033-17>.

Copyright © 2017 Lipus 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 Djuna Gulliver, Djuna.Gulliver@NETL.DOE.GOV.

discovery of genes participating in biofilm formation processes. These included members of the *alg* family involved in alginate metabolism and the levan production gene levansucrase (15, 16). Furthermore, multiple genes of the polysaccharide synthase locus *psl* and the glycerol transferase *glt2*, involved in exopolysaccharide (EPS) production were identified (16, 17). Annotation also allowed the identification of the osmotic stress response genes *ProX*, *ProU*, and the *Trk* complex suggesting the potential for potassium ion uptake and osmoprotectant accumulation.

The recovery and analysis of the *Pseudomonas* sp. BDAL1 draft genome revealed a high number of biofilm formation genes, suggesting significant potential for biofilm formation and biofouling events in hydraulic fracturing-produced water storage tank and other parts of the hydraulic fracturing infrastructure. More detailed analysis of this and other bacterial genomes recovered from hydraulic fracturing environments will lead to a better understanding of microbial activity in these environments.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number [MCRW00000000](#). The version described in this paper is version MCRW01000000.

ACKNOWLEDGMENTS

This technical effort was performed under the RES contract RES1000027/183U and was supported by the Oak Ridge Institute of Science and Education (ORISE).

REFERENCES

- Cluff MA, Hartsock A, MacRae JD, Carter K, Mouser PJ. 2014. Temporal changes in microbial ecology and geochemistry in produced water from hydraulically fractured Marcellus Shale gas wells. *Environ Sci Technol* 48:6508–6517. <https://doi.org/10.1021/es501173p>.
- Kargbo DM, Wilhelm RG, Campbell DJ. 2010. Natural gas plays in the Marcellus shale: challenges and potential opportunities. *Environ Sci Technol* 44:5679–5684. <https://doi.org/10.1021/es903811p>.
- Gregory KB, Vidic RD, Dzombak DA. 2011. Water management challenges associated with the production of shale gas by hydraulic fracturing. *Elements* 7:181–186. <https://doi.org/10.2113/gselements.7.3.181>.
- Vengosh A, Jackson RB, Warner N, Darrah TH, Kondash A. 2014. A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environ Sci Technol* 48:8334–8348. <https://doi.org/10.1021/es405118y>.
- Murali Mohan A, Hartsock A, Bibby KJ, Hammack RW, Vidic RD, Gregory KB. 2013. Microbial community changes in hydraulic fracturing fluids and produced water from shale gas extraction. *Environ Sci Technol* 47:13141–13150. <https://doi.org/10.1021/es402928b>.
- Murali Mohan AM, Hartsock A, Hammack RW, Vidic RD, Gregory KB. 2013. Microbial communities in flowback water impoundments from hydraulic fracturing for recovery of shale gas. *FEMS Microbiol Ecol* 86:567–580. <https://doi.org/10.1111/1574-6941.12183>.
- 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>.
- Laczny CC, Sternal T, Plugaru V, Gawron P, Atashpendar A, Margossian HH, Coronado S, der Maaten Lv, Vlassis N, Wilmes P. 2015. VizBin—an application for reference-independent visualization and human-augmented binning of metagenomic data. *Microbiome* 3:1. <https://doi.org/10.1186/s40168-014-0066-1>.
- McHardy AC, Martín HG, Tsirigos A, Hugenholtz P, Rigoutsos I. 2007. Accurate phylogenetic classification of variable-length DNA fragments. *Nat Methods* 4:63–72. <https://doi.org/10.1038/nmeth976>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards R. 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386. <https://doi.org/10.1186/1471-2105-9-386>.
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.
- Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57:81–91. <https://doi.org/10.1099/ijs.0.64483-0>.
- Rodríguez-R LM, Konstantinidis KT. 2014. Bypassing cultivation to identify bacterial species. *Microbe* 9:111–118. <https://doi.org/10.1128/microbe.9.111.1>.
- Laue H, Schenk A, Li H, Lamberts L, Neu TR, Molin S, Ullrich MS. 2006. Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*. *Microbiology* 152:2909–2918. <https://doi.org/10.1099/mic.0.28875-0>.
- Mann EE, Wozniak DJ. 2012. *Pseudomonas* biofilm matrix composition and niche biology. *FEMS Microbiol Rev* 36:893–916. <https://doi.org/10.1111/j.1574-6976.2011.00322.x>.
- Flemming H-C, Wingender J. 2010. The biofilm matrix. *Nat Rev Microbiol* 8:623–633. <https://doi.org/10.1038/nrmicro2415>.