



Draft Genome Sequences of 10 Bacterial Strains Isolated from an Abandoned Coal Mine in Southeast Kansas

Brady Steinbock,^a Rachel Bechtold,^{a*} Joseph L. Sevigny,^b Devin Thomas,^b W. Kelley Thomas,^b Anuradha Ghosh^a

^aDepartment of Biology, Pittsburg State University, Pittsburg, Kansas, USA

^bHubbard Center for Genome Studies, University of New Hampshire, Durham, New Hampshire, USA

ABSTRACT Here, we report 10 bacterial strains isolated from an abandoned coal mine in southeast Kansas to determine their potential for bioremediation through comparison of the genome sizes and distribution patterns of unique metabolic genes. The selected strains belong to the genera *Arthrobacter*, *Jeotgalibacillus*, *Kocuria*, *Microbacterium*, *Pantoea*, *Rhodococcus*, *Vibrio*, *Brevibacterium*, and *Paenibacillus*.

For over half a century, sizeable areas of land throughout the Midwest were mined for various underground resources, one of the most notable being coal. Once mined, the lands were left for years. Some areas were partially reclaimed following the Surface Mining Control and Reclamation Act of 1977. The major environmental pollutant that became established in those reclaimed as well as abandoned areas was acid mine drainage (AMD) (1). Pyrite, which is usually found in anaerobic environments, is exposed to atmospheric oxygen and water after coal mines are abandoned, leading to the formation of sulfuric acid and iron. The iron precipitates at the shores and beds of mined land, and sulfuric acid runoff dissolves heavy metals into ground and surface water. Naturally occurring acidophilic bacteria in mined land can increase the oxidation rate of pyrite, while other microbes may offer a solution to the detrimental effects of AMD (2). In the present study, whole-genome sequencing (WGS) analysis provided insight into the presence of genes related to stress, resistance, and metabolism. Thus, application of these isolates is not limited to restoration of AMD sites but also could be exploited for bioremediation of heavy metals and recalcitrant chemicals in contaminated sites.

Soil samples were collected from a depth of 5 cm at five specific sites in the partially reclaimed soil as well as from an AMD site. Up to 60 bacterial isolates were characterized using physiological and biochemical tests (3). Out of these, 10 strains were selected for WGS analysis based on their metabolic diversity. Selected pure cultures were grown overnight in tryptic soy broth (BD Difco, Franklin Lakes, NJ), and whole genomic DNA was isolated with a GenElute bacterial genomic DNA kit (Sigma-Aldrich Corporation, Natick, MA) following the manufacturer's protocol. Finally, a NanoDrop Lite (Thermo Fisher Scientific, Waltham, MA) analysis was performed on each isolated DNA for quantification and to ensure sample purity.

WGS was performed at the Hubbard Center for Genome Studies (University of New Hampshire, Durham, NH). A paired-end library was constructed using a Nextera DNA library preparation kit (Illumina, San Diego, CA) and sequenced with an Illumina HiSeq 2500 instrument to produce 250-bp paired-end reads. The total numbers of reads for all 10 strains are listed in Table 1. FASTQ files were trimmed for Nextera adapters and low-quality bases using Trimmomatic version 0.32 (4). For read trimming, trailing and leading bases were removed if the quality score was below 3. In addition, the reads were scanned using a 4-base sliding window and trimmed if the average quality dropped below 15. Trimmed sequencing reads were then assembled using the SPAdes

Citation Steinbock B, Bechtold R, Sevigny JL, Thomas D, Thomas WK, Ghosh A. 2019. Draft genome sequences of 10 bacterial strains isolated from an abandoned coal mine in southeast Kansas. *Microbiol Resour Announc* 8:e01001-19. <https://doi.org/10.1128/MRA.01001-19>.

Editor Irene L. G. Newton, Indiana University, Bloomington

Copyright © 2019 Steinbock 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 Anuradha Ghosh, aghosh@pittstate.edu.

* Present address: Rachel Bechtold, Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, Arkansas, USA.

Received 15 August 2019

Accepted 25 September 2019

Published 17 October 2019

TABLE 1 Accession numbers, assembly metrics, and annotated features of the sequenced environmental strains isolated from an abandoned coal mine in southeast Kansas

Bacterial species	Strain	GenBank accession no.	Avg coverage (x)	No. of contigs	Total no. of reads	Genome assembly size (bp)	N_{50} value (bp)	G+C content (%)	No. of coding genes	No. of rRNAs	No. of tRNAs	No. of ncRNAs ^a
<i>Vibrio vulnificus</i>	S-C3	SMZZ00000000	447	31	11,287,962	5,792,187	849,468	46.67	5,330	24	81	6
<i>[Brevibacterium] frigoritolerans</i>	F-D2	SMZY00000000	1,408	21	12,709,518	4,568,742	1,243,018	40.87	4,411	21	77	5
<i>Microbacterium oleivorans</i>	F-B2	SMZX00000000	577	5	7,893,216	3,092,005	1,655,501	46.55	2,911	5	45	3
<i>Paenibacillus dendritiformis</i>	F-A1	SMZW00000000	340	33	10,817,884	6,777,904	735,989	53.98	5,788	21	80	4
<i>Paenibacillus amylolyticus</i>	S-B4	SMZV00000000	267	21	8,167,626	7,067,323	3,885,192	45.81	6,008	27	80	4
<i>Rhodococcus qingshengii</i>	S-E5	SMZU00000000	431	45	12,222,178	6,609,607	474,786	38.17	6,021	32	115	7
<i>Kocuria rosea</i>	S-A3	SMZT00000000	576	18	9,751,028	3,953,419	565,760	72.67	3,503	5	48	3
<i>Pantoea ananatis</i>	F-C2	SMZS00000000	239	29	6,247,448	4,805,071	405,717	53.64	4,346	13	71	20
<i>Jeotgalibacillus</i> sp.	S-D1	SMZR00000000	558	49	9,813,060	3,978,119	790,309	41.33	3,789	18	77	5
<i>Arthrobacter nitrophenolicus</i>	S-A1	SMZQ00000000	479	43	8,918,834	4,614,804	320,201	65.99	4,168	8	50	3

^a ncRNAs, noncoding RNAs.

pipeline version 3.5 (5) with default settings. QUAST version 4.6.0 (6) was used to assess the contiguity of the assemblies, and coverage statistics were calculated by mapping FASTQ reads to the assembled contigs with the Burrows-Wheeler Aligner MEM algorithm (BWA-MEM) with default settings. The assembled genomes were annotated via the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (7), while the identity of these strains was determined by performing a BLASTn (8) search on 16S rRNA and translation initiation factor-2 (IF-2) gene sequences. The assembly metrics and annotated features are given in Table 1.

Data availability. The draft genome sequences of these environmental strains, as well as the accession numbers for both the assembly and raw reads, are available at DDBJ/ENA/GenBank under the BioProject number [PRJNA523266](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA523266), and the described accession numbers are listed in Table 1.

ACKNOWLEDGMENTS

This project was partly funded by Pittsburg State University Graduate and Continuing Studies and the Kansas IDeA Network of Biomedical Research Excellence (K-INBRE) (grant P20 GM103418, National Institute of General Medical Sciences). The whole-genome sequencing was supported by the New Hampshire IDeA Network of Biomedical Research Excellence (NH-INBRE) (grant P20 GM103506, National Institute of General Medical Sciences).

We declare no conflicts of interest.

REFERENCES

- Johnson DB, Hallberg KB. 2003. The microbiology of acidic mine waters. *Res Microbiol* 154:466–473. [https://doi.org/10.1016/S0923-2508\(03\)00114-1](https://doi.org/10.1016/S0923-2508(03)00114-1).
- Kolmert Å, Johnson DB. 2001. Remediation of acidic waste waters using immobilised, acidophilic sulfate-reducing bacteria. *J Chem Technol Biotechnol* 76:836–843. <https://doi.org/10.1002/jctb.453>.
- Bechtold R. 2017. Bacterial diversity at an abandoned coal mine in southeast Kansas. Master's thesis. Pittsburg State University, Pittsburg, Kansas. <https://digitalcommons.pittstate.edu/etd/335/>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, Mclean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol* 20:714–737. <https://doi.org/10.1089/cmb.2013.0084>.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).