AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Genome Sequence of a Novel Soil Actinomycete, *Protaetiibacter* sp. Strain SSC-01

Huiquan Duan,[a]* Chamil E. Fernando,[a] Scott S. Crupper,[a] Stephen D. Fields[a]

[a]Department of Biological Sciences, Emporia State University, Emporia, Kansas, USA

ABSTRACT The family *Microbacteriaceae* represents a diverse and important group of soil bacteria in the phylum *Actinobacteria*. Here, we report the genome sequence of a soil *Microbacteriaceae* strain, *Protaetiibacter* sp. strain SSC-01, the second putative species of the genus. Iron acquisition and xylose metabolism are central pathways identified in the annotated genome.

The actinobacteria are dominant taxa in temperate soils and make significant contributions to processes such as polysaccharide digestion, antibiotic-based microbial inhibition, heavy metal chelation, and plant growth stimulation (1). The placement of many of the taxa within the phylum is in flux, and the *Microbacteriaceae* family has undergone numerous recent revisions with newly proposed genera (1–3).

The *Microbacteriaceae* strain SSC-01, which was collected from cultivated garden soil in east central Kansas, United States (38.42N, 95.85W), was isolated by passaging single colonies through multiple rounds of growth on Reasoner's 2A agar (4) at 37°C. Other media supporting SSC-01 growth were blood agar (Difco BD, Sparks, MD, USA) with defibrinated sheep's blood (Hemostat Labs, Dixon, CA, USA) and tryptic soy agar (Difco BD) supplemented with 10 μM $FeCl_3 \cdot 6H_2O$. DNA was isolated from bacterial lawns grown on blood agar using the Quick-DNA fungal/bacterial miniprep kit (Zymo Research, Irvine, CA, USA). SeqMatic LLC (Fremont, CA, USA) prepared a genomic library with the TruSeq protocol targeting a 450-bp insert. Illumina MiSeq sequencing with a 500-cycle TruSeq kit (v.2 chemistry) generated 2,107,131 paired-end 251-bp reads. Adapters were removed with Cutadapt (v.2.5) (5), resulting in 1,057,609,266 bases of Illumina sequencing. A DNA library for Nanopore MinION sequencing, prepared with the ligation sequencing kit (SQK-LSK109; Oxford Nanopore Technologies, Oxford, UK), generated 24,000 reads averaging 4,285 bp. Illumina and Nanopore reads were uploaded to the public usegalaxy.org server (6) for processing with FASTQ Groomer (v.1.1.5), fastp (v.0.19.5+galaxy1), and/or Porechop (v.0.2.3) tools (7–9) using default settings. Unicycler (v.0.4.8.0) (10) was used to assemble 1.5 million paired Illumina reads (252-fold coverage) using 14,856 high-quality MinION reads (21.5-fold coverage) as a scaffold to generate a single 2,958,807-bp contig with a GC content of 71.5%. Circularity was confirmed by identifying Nanopore reads spanning the artificial ends of the assembled contig.

Annotation using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) server (11) identified 2,690 potential protein-coding genes, 52 RNA-coding genes, and 30 pseudogenes. Approximately 33% of the putative proteins have unknown function, but over 20 proteins are predicted to be involved in iron acquisition pathways, including siderophore-based import, hemolysin III toxins, and heme transporters. Hemicellulose metabolism (including xylan and xylose), terpene synthesis, and heavy metal chelation pathways each have 10 or more predicted enzymes, suggesting important roles for this microbe in plant-soil interactions (12, 13).

Needleman-Wunsch nucleotide alignments of full 16S rRNA gene sequences using NCBI default settings revealed the greatest identity to both *Lysinimonas* KACC 19322 (GenBank accession no. MT367295.1) and *Protaetiibacter intestinalis* (originally *Lysinimonas*

Citation Duan H, Fernando CE, Crupper SS, Fields SD. 2021. Genome sequence of a novel soil actinomycete, *Protaetiibacter* sp. strain SSC-01. Microbiol Resour Announc 10:e01029-20. https://doi.org/10.1128/MRA.01029-20.

Editor Kenneth M. Stedman, Portland State University

Address correspondence to Stephen D. Fields, sfields1@emporia.edu.

* Present address: Huiquan Duan, Department of Biochemistry and Molecular Biophysics, Kansas State University, Manhattan, Kansas, USA.

Received 5 October 2020
Accepted 14 January 2021
Published 4 February 2021

TABLE 1 Species that are closely related to strain SSC-01 that have available complete genomic sequences[a]

Species/strain	GenBank accession no.	No. of orthologues shared with SSC-01[b]	Mean identity with SSC-001(%)
Leifsonia xyli CTCB07	NC_006087.1	1,581	55.6
Lysinimonas sp. KACC 19322	NZ_CP043504.1	2,129	71.9
Lysinimonas sp. SLBN-160	NZ_JACBYT010000001.1	2,057	55.9
Protaetiibacter intestinalis	NZ_CP032630.1	2,317	75.7

[a] Predicted orthologous proteins and their average percent identity with their SSC-1 counterparts were determined with the sequence-based comparative software of the RAST SEED browser (14).
[b] Amino acid identity of 30% used as the threshold for orthology.

sp. strain 2DFWR-13) (GenBank accession no. MH989600.1) at 98.6%. Global comparative analysis with the RAST SEED viewer (14) showed that the average identity of all orthologous protein sequences of SSC-01 is higher for *P. intestinalis* than for other closely related species (Table 1), leading to tentative identification of strain SSC-01 as a *Protaetiibacter* sp. This second putative species of the genus represents a different ecological niche than *P. intestinalis*, which was isolated from a larval *Protaetia* moth gut (3). A Mexico City, Mexico, landfill metagenome (GenBank accession no. RDDZ01000083.1), however, suggests the existence of other soil *Protaetiibacter* species.

Data availability. This bacterial genome sequence has been deposited in DDBJ/ENA/GenBank under the accession no. CP059987. Raw sequence data used for assembly have been deposited in DDBJ/ENA/GenBank under the accession no. PRJNA649951. The assembly described in this paper is the first version (GenBank accession no. CP059987.1).

ACKNOWLEDGMENTS

Funding was provided through an Institutional Development Award from the National Institute of General Medical Sciences of the National Institutes of Health under grant P20 GM103418.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

REFERENCES

1. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Meier-Kolthoff JP, Klenk H-P, Clément C, Ouhdouch Y, van Wezel GP. 2016. Taxonomy, physiology, and natural products of *Actinobacteria*. Microbiol Mol Biol Rev 80:1–43. https://doi.org/10.1128/MMBR.00019-15.
2. Jang Y-H, Kim S-J, Tamura T, Hamada M, Weon H-Y, Suzuki K, Kwon S-W, Kim W-G. 2013. *Lysinimonas soli* gen. nov., sp. nov., isolated from soil, and reclassification of *Leifsonia kribbensis* Dastager *et al.* 2009 as *Lysinimonas kribbensis* sp. nov., comb. nov. Int J Syst Evol Microbiol 63:1403–1410. https://doi.org/10.1099/ijs.0.042945-0.
3. Heo J, Cho H, Kim MA, Hamada M, Tamura T, Saitou S, Kim S-J, Kwon S-W. 2019. *Protaetiibacter intestinalis* gen. nov., of the family *Microbacteriaceae*, isolated from gut of *Protaetia brevitarsis seulensis*, reclassification of *Lysinimonas kribbensis* Jang *et al.* 2013 as *Pseudolysinimonas kribbensis* gen. nov., comb. nov. and emended description of the genus *Lysinimonas* Jang *et al.* 2013. Int J Syst Evol Microbiol 69:2101–2107. https://doi.org/10.1099/ijsem.0.003444.
4. Reasoner DJ, Geldreich EE. 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl Environ Microbiol 49:1–7. https://doi.org/10.1128/AEM.49.1.1-7.1985.
5. Marcel M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.
6. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46:W537–W544. https://doi.org/10.1093/nar/gky379.
7. Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A, Galaxy Team. 2010. Manipulation of FASTQ data with Galaxy. Bioinformatics 26:1783–1785. https://doi.org/10.1093/bioinformatics/btq281.
8. Loman NJ, Quinlan AR. 2014. Poretools: a toolkit for analyzing Nanopore sequence data. Bioinformatics 30:3399–3401. https://doi.org/10.1093/bioinformatics/btu555.
9. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
10. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
11. 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.
12. Aznar A, Dellagi A. 2015. New insights into the role of siderophores as triggers of plant immunity: what can we learn from animals? J Exp Bot 66:3001–3010. https://doi.org/10.1093/jxb/erv155.
13. Sharma V, Salwan R, Al-Ani LKT (ed). 2020. Molecular aspects of plant beneficial microbes in agriculture. Academic Press, London, United Kingdom.
14. 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.