10-16-2014

Draft Genome Sequences of 10 Strains of the Genus Exiguobacterium

Tatiana A. Vishnivetskaya
University of Tennessee, Knoxville

Archana Chauhan
University of Tennessee, Knoxville

Alice C. Layton
University of Tennessee, Knoxville

Susan M. Pfiffner
University of Tennessee, Knoxville

Marcel Huntemann
DOE Joint Genome Institute

See next page for additional authors

Recommended Citation

Vishnivetskaya TA, Chauhan A, Layton AC, Pfiffner SM, Huntemann M, Copeland A, Chen A, Kyrpides NC, Markowitz VM, Palaniappan K, Ivanova N, Mikhailova N, Ovchinnikova G, Andersen EW, Pati A, Stamatis D, Reddy TBK, Shapiro N, Nordberg HP, Cantor MN, Hua XS, Woyke T. 2014. Draft genome sequences of 10 strains of the genus Exiguobacterium. Genome Announc. 2(5):e01058-14. doi:10.1128/genomeA.01058-14.

genomeA Journals.ASM.org

Draft Genome Sequences of 10 Strains of the Genus *Exiguobacterium*

Tatiana A. Vishnivetskaya,[a] Archana Chauhan,[a,b] Alice C. Layton,[a] Susan M. Pfiffner,[a] Marcel Huntemann,[c] Alex Copeland,[c] Amy Chen,[c] Nikos C. Kyrpides,[c] Victor M. Markowitz,[c] Krishna Palaniappan,[c] Natalia Ivanova,[c] Natalia Mikhailova,[c] Galina Ovchinnikova,[c] Evan W. Andersen,[c] Amrita Pati,[c] Dimitrios Stamatis,[c] T. B. K. Reddy,[c] Nicole Shapiro,[c] Henrik P. Nordberg,[c] Michael N. Cantor,[c] X. Susan Hua,[c] Tanja Woyke[c]

Center for Environmental Biotechnology, University of Tennessee, Knoxville, Tennessee, USA[a]; UT-ORNL Joint Institute for Biological Sciences, Oak Ridge, Tennessee, USA[b]; DOE Joint Genome Institute, Walnut Creek, California, USA[c]

High-quality draft genome sequences were determined for 10 *Exiguobacterium* strains in order to provide insight into their evolutionary strategies for speciation and environmental adaptation. The selected genomes include psychrotrophic and thermophilic species from a range of habitats, which will allow for a comparison of metabolic pathways and stress response genes.

Received 8 September 2014 **Accepted** 10 September 2014 **Published** 16 October 2014

Citation Vishnivetskaya TA, Chauhan A, Layton AC, Pfiffner SM, Huntemann M, Copeland A, Chen A, Kyrpides NC, Markowitz VM, Palaniappan K, Ivanova N, Mikhailova N, Ovchinnikova G, Andersen EW, Pati A, Stamatis D, Reddy TBK, Shapiro N, Nordberg HP, Cantor MN, Hua XS, Woyke T. 2014. Draft genome sequences of 10 strains of the genus *Exiguobacterium*. Genome Announc. 2(5):e01058-14. doi:10.1128/genomeA.01058-14.

Address correspondence to Tatiana A. Vishnivetskaya, tvishniv@utk.edu.

Exiguobacterium, belonging to the order *Bacillales* of the phylum *Firmicutes*, was proposed as a new genus in 1983 by Collins et al. (1) and includes 16 species. All these species are Gram-positive, rod-shaped, facultative anaerobes, motile via peritrichous flagella and have been isolated from a wide range of habitats, with temperatures ranging from −12° to 55°C (2). The ability of individual strains isolated under psychrotrophic or thermophilic conditions to grow in the mesophilic temperature range of 15° to 37°C suggests that *Exiguobacterium* species have unique and conserved genetic pathways allowing these organisms to exploit a diversity of temperature-related habitats. In addition, species with close affiliation to modern strains have been isolated from permafrost and ice estimated to be >100,000 years old.

The genome sequences were determined for 10 *Exiguobacterium* strains, including six type strains and four environmental isolates (Table 1). High-molecular-weight genomic DNA was iso-

TABLE 1 Characteristics of 10 *Exiguobacterium* draft genomes

Organism[a]	Isolation source	Sequencing and assembly methods	Size (Mb)	G+C content (%)	No. of CDSs	No. of *rrn* operons	No. of tRNAs	No. of Transposases	No. of cold shock genes	GenBank accession no.	No. of contigs
E. acetylicum DSM 20416^T	Creamery waste, UK	PacBio, HGAP	3.28	47	3,323	9	69	40	3	JNIR00000000	3
E. oxidotolerans JCM 12280^T	Fish drain, Japan	PacBio, HGAP	3.09	47	3,053	9	69	34	6	JNIS00000000	3
E. undae DSM 14481^T	Garden pond, Germany	Illumina, AllPaths-LG	3.25	48	3,287	4	56	12	2	JHZV00000000	4
E. antarcticum DSM 14480^T	Microbial mat, Antarctica	PacBio, HGAP	3.22	47	3,250	10	69	89	7	JMKS00000000	7
E. sibiricum 7-3	Permafrost, Siberia, Russia	Illumina, AllPaths-LG	3.08	47	3,141	4	48	9	3	JHZS00000000	7
E. undae 190-11	Permafrost, Siberia, Russia	Illumina, AllPaths-LG	3.21	48	3,236	5	61	17	3	JHZU00000000	4
E. aurantiacum DSM 6208^T	Potato wash, UK	PacBio, HGAP	3.04	53	3,067	9	67	90	2	JNIQ00000000	2
E. marinum DSM 16307^T	Marine, Yellow Sea, South Korea	Illumina, AllPaths-LG	2.81	47	2,836	8	60	15	2	JHZT00000000	2
Exiguobacterium sp. GIC31	Glacier ice, Greenland	PacBio, HGAP	2.97	52	3,005	9	67	38	2	JNIP00000000	2
Exiguobacterium sp. NG55	Hot spring, Yellowstone Park, USA	PacBio, HGAP	3.14	48	3,169	11	68	27	2	JPOD00000000	5

[a] Type strains (^{T}) were obtained from the German Collection of Microorganisms and Cell Cultures (DSM) or Japan Collection of Microorganisms (JCM).

lated from strains grown overnight at 30°C in tryptic soy broth using the Joint Genome Institute (JGI) modified cetyltrimethylammonium bromide (CTAB) protocol (3). The draft genomes were generated at JGI using Illumina and Pacific Biosciences (PacBio) technologies. Illumina shotgun and long-insert mate-pair libraries were constructed and sequenced using the Illumina HiSeq 2000 platform (4). Filtered Illumina reads were assembled using AllPaths-LG (5). PacBio SMRTbell libraries were constructed and sequenced on the PacBio RS platform, and raw reads were assembled using HGAP version 2.0.1 (6). Genes were identified using Prodigal (7), followed by manual curation using GenePRIMP (8). The predicted coding sequences (CDSs) were translated and used to search the NCBI nonredundant, UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAscan-SE tool (9) was used to find tRNA genes, and rRNA genes were identified against models of the rRNA genes built from SILVA (10). Noncoding RNAs were found by searching the genomes for the corresponding Rfam profiles using Infernal (11). Gene prediction analysis and manual functional annotation were performed within the Integrated Microbial Genomes (IMG) platform (12).

The *Exiguobacterium* strains have low G+C contents (average, 48.4%) and vary slightly in their genome size, number of CDSs, and ribosomal RNA (*rrn*) operons (Table 1). Whole-genome sequencing identified 13 transposase families, which is consistent with those found in previous publications (2, 13). The two most abundant transposase families, transposase/inactivated derivatives and IS*605* (*orfB*), are present in all strains. The strains contain two to seven cold shock protein genes (COG1278), one molecular chaperone GrpE (heat shock protein, COG0576), one ribosome-associated heat shock protein (S4 paralog, COG1188), three chaperonin GroEL (HSP60 family, COG0459), three cochaperonin GroES (HSP10, COG0234), and four fatty acid desaturase (COG3239) genes per strain.

The presence of multiple genes encoding stress-responsive proteins may explain the broad temperature range for growth and the ability of the *Exiguobacterium* strains to colonize and thrive in diverse ecological niches.

Nucleotide sequence accession numbers. These whole-genome shotgun projects have been deposited in DDBJ/EMBL/GenBank under accession numbers JNIR00000000, JNIS00000000, JHZV00000000, JMKS00000000, JHZS00000000, JHZU00000000, JNIQ00000000, JHZT00000000, JNIP00000000, and JPOD00000000. The versions described in this paper are the first versions, JNIR01000000, JNIS01000000, JHZV01000000, JMKS01000000, JHZS01000000, JHZU01000000, JNIQ01000000, JHZT01000000, JNIP01000000, and JPOD01000000.

ACKNOWLEDGMENTS

This research was performed through the Community Science Program, CSP 2012. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231.

We thank the Yellowstone National Park Service for coordinating and allowing sampling under permit YELL-1502. We are grateful to J. M. Tiedje, S. Kathariou, R. F. Ramaley, and V. Miteva for providing strains *E. undae* 190-11, *E. sibiricum* 7-3, *Exiguobacterium* sp. NG55, and *Exiguobacterium* sp. GIC31.

REFERENCES

1. **Collins MD, Lund BM, Farrow JAE, Schleifer KH.** 1983. Chemotaxonomic study of an alkaliphilic bacterium, *Exiguobacterium aurantiacum* gen. nov., sp. nov. J. Gen. Microbiol. **129:**2037–2042.
2. **Vishnivetskaya TA, Kathariou S, Tiedje JM.** 2009. The *Exiguobacterium* genus: biodiversity and biogeography. Extremophiles **13:**541–555. http://dx.doi.org/10.1007/s00792-009-0243-5.
3. **Wilson K.** 2001. Preparation of genomic DNA from bacteria. Curr. Protoc. Mol. Biol. Chapter 2:Unit 2.4. http://dx.doi.org/10.1002/0471142727.mb0204s56.
4. **Bennett S.** 2004. Solexa Ltd. Pharmacogenomics **5:**433–438. http://dx.doi.org/10.1517/14622416.5.4.433.
5. **Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB.** 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc. Natl. Acad. Sci. U. S. A. **108:**1513–1518. http://dx.doi.org/10.1073/pnas.1017351108.
6. **Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J.** 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods **10:**563–569. http://dx.doi.org/10.1038/nmeth.2474.
7. **Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ.** 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11:**119. http://dx.doi.org/10.1186/1471-2105-11-119.
8. **Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC.** 2010. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat. Methods **7:**455–457. http://dx.doi.org/10.1038/nmeth.1457.
9. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964. http://dx.doi.org/10.1093/nar/25.5.0955.
10. **Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO.** 2007. Silva: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. **35:**7188–7196. http://dx.doi.org/10.1093/nar/gkm864.
11. **Nawrocki EP, Eddy SR.** 2013. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics **29:**2933–2935. http://dx.doi.org/10.1093/bioinformatics/btt509.
12. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278. http://dx.doi.org/10.1093/bioinformatics/btp393.
13. **Vishnivetskaya TA, Kathariou S.** 2005. Putative transposases conserved in *Exiguobacterium* isolates from ancient Siberian permafrost and from contemporary surface habitats. Appl. Environ. Microbiol. **71:**6954–6962. http://dx.doi.org/10.1128/AEM.71.11.6954-6962.2005.