Draft Genome Sequence of *Aeribacillus pallidus* Strain 8m3, a Thermophilic Hydrocarbon-Oxidizing Bacterium Isolated from the Dagang Oil Field (China)

Andrey B. Poltaraus,[a] Diyana S. Sokolova,[b] Denis S. Grouzdev,[b] Timophey M. Ivanov,[b] Sophia G. Malakho,[c] Alena V. Korshunova,[b] Aleksey S. Rozanov,[d] Tatiyana P. Tourova,[b] Tamara N. Nazina[b]

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russian Federation[a]; Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, Russian Federation[b]; Federal State Budgetary Institution Federal Science Center for Physical Culture and Sport, Moscow, Russian Federation[c]; Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russian Federation[d]

The draft genome sequence of *Aeribacillus pallidus* strain 8m3, a thermophilic aerobic oil-oxidizing bacterium isolated from production water from the Dagang high-temperature oil field, China, is presented here. The genome is annotated to provide insights into the genomic and phenotypic diversity of the genus *Aeribacillus*.

Address correspondence to Tamara N. Nazina, nazinatn@gmail.com.

The genus *Aeribacillus* contains a single species, *Aeribacillus pallidus* (formerly *Geobacillus pallidus* = *Bacillus pallidus*) (1). Representatives of this species have been isolated from sewage (1), production water from high-temperature oilfields (2, 3), hot springs (4, 5), oil-contaminated soil (6), and from a deep geothermal reservoir (7). The only genome of the *A. pallidus* (strain GS3372) deposited at DDBJ/EMBL/GenBank has a G+C content of 57.4% (7). This value is considerably higher than the G+C content of the DNA (39 to 41 mol%) of *A. pallidus* type strain DSM 3670. The strain 8m3 was isolated from a formation water sample from the Dagang high-temperature oil field (Hebei Province, China) (38°40′4.7″N, 117°22′38.0″E) (3). Strain 8m3 is an aerobic endospore-forming bacterium able to grow at temperatures ranging from 38 to 65°C and produces biosurfactants in the course of the degradation of crude oil. According to phylogenetic analysis of the 16S rRNA gene sequence, it was identified as an *A. pallidus* strain (3).

The genomic DNA of this strain was isolated from the biomass using the DNeasy blood and tissue kit (Qiagen, Germany), according to the manufacturer's instructions. The TruSeq DNA sample preparation kit (Illumina, USA) was used to create the libraries for genome sequencing. Genomic DNA was sequenced using the MiSeq reagent kit version 2 (Illumina). The shotgun library was constructed with a 500-bp paired-end library.

Approximately 7,344,862 paired-end reads were generated, providing 200-fold genome coverage. The resulting reads were assembled with SPAdes version 3.1.0 (8). The remaining 79 contigs were submitted to the ProDeGe website (9) for automatic decontamination. Finally, the assembled sequence was submitted to the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP) (10) for annotation.

The draft genome sequence of *A. pallidus* strain 8m3 revealed a genome size of 3,818,610 bp, with an average G+C content of 38.86%, which was close to the G+C content of *A. pallidus* type strain DSM 3670. The draft genome contained 3,717 genes, 3,520 coding DNA sequences, 15 coding rRNAs (5S, 16S, and 23S), 60 tRNAs, and 5 noncoding RNAs (ncRNAs). The 16S rRNA and *gyrB* gene sequences revealed in the genome were identical to those determined earlier for the strain (3). Other genes were associated with the degradation of aromatic compounds and fatty acids, utilization of sugars and amino acids, and sporulation. New genomic data were in agreement with the results of phenotypic studies of the strain.

This genome sequence is expected to provide deep insights into the metabolic potential of this organism and its compliance to the conditions of the deep high-temperature petroleum reservoir.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. LWBR00000000. The version described in this paper is version LWBR01000000.

REFERENCES

1. **Miñana-Galbis D, Pinzón DL, Lorén JG, Manresa A, Oliart-Ros RM.** 2010. Reclassification of *Geobacillus pallidus* (Scholz et al. 1988) Banat et al. 2004 as *Aeribacillus pallidus* gen. nov., comb. nov. Int J Syst Evol Microbiol **60:**1600–1604. http://dx.doi.org/10.1099/ijs.0.003699-0.
2. **Chamkha M, Mnif S, Sayadi S.** 2008. Isolation of a thermophilic and halophilic tyrosol-degrading *Geobacillus* from a Tunisian high-

temperature oil field. FEMS Microbiol Lett **283**:23–29. http://dx.doi.org/10.1111/j.1574-6968.2008.01136.x.

3. **Tourova TP, Korshunova AV, Mikhailova EM, Sokolova DSh, Poltaraus AB, Nazina TN**. 2010. Application of *gyrB* and *parE* sequence similarity analyses for differentiation of species within the genus *Geobacillus*. Microbiology **79**:356–369. http://dx.doi.org/10.1134/S0026261710030124.

4. **Pinzón-Martínez DL, Rodríguez-Gómez C, Miñana-Galbis D, Carrillo-Chávez JA, Valerio-Alfaro G, Oliart-Ros R**. 2010. Thermophilic bacteria from Mexican thermal environments: isolation and potential applications. Environ Technol **31**:957–966. http://dx.doi.org/10.1080/09593331003758797.

5. **Yasawong M, Areekit S, Pakpitchareon A, Santiwatanakul S, Chansiri K**. 2011. Characterization of thermophilic halotolerant *Aeribacillus pallidus* TD1 from Tao Dam Hot Spring, Thailand. Int J Mol Sci **12**:5294–5303. http://dx.doi.org/10.3390/ijms12085294.

6. **Zheng C, He J, Wang Y, Wang M, Huang Z**. 2011. Hydrocarbon degradation and bioemulsifier production by thermophilic *Geobacillus pallidus* strains. Bioresour Technol **102**:9155–9161. http://dx.doi.org/10.1016/j.biortech.2011.06.074.

7. **Filippidou S, Jaussi M, Junier T, Wunderlin T, Jeanneret N, Regenspurg S, Li P-E, Lo C-C, Johnson S, McMurry K, Gleasner CD, Vuyisich M, Chain PS, Junier P**. 2015. Genome sequence of *Aeribacillus pallidus* strain GS3372, an endospore-forming bacterium isolated in a deep geothermal reservoir. Genome Announc **3**(4):e00981-15. http://dx.doi.org/10.1128/genomeA.00981-15.

8. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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. http://dx.doi.org/10.1089/cmb.2012.0021.

9. **Tennessen K, Andersen E, Clingenpeel S, Rinke C, Lundberg DS, Han J, Dangl JL, Ivanova N, Woyke T, Kyrpides N, Pati A**. 2016. ProDeGe: a computational protocol for fully automated decontamination of genomes. ISME J **10**:269–272. http://dx.doi.org/10.1038/ismej.2015.100.

10. **Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W**. 2013. Prokaryotic genome annotation pipeline. In The NCBI handbook, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.