AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Draft Genome Sequence of the Type Strain *Sphingopyxis bauzanensis* DSM 22271

Michal A. Kaminski,[a] Ewa M. Furmanczyk,[a] Andrzej Dziembowski,[a,b] Adam Sobczak,[a,b] Leszek Lipinski[a]

Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland[a]; Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, Warsaw, Poland[b]

ABSTRACT We present here the draft genome sequence of *Sphingopyxis bauzanensis* DSM 22271. The assembly contains 4,258,005 bp in 28 scaffolds and has a GC content of 63.3%. A series of specific genes involved in the catabolism or transport of aromatic compounds was identified.

Received 14 August 2017 **Accepted** 15 August 2017 **Published** 14 September 2017

Citation Kaminski MA, Furmanczyk EM, Dziembowski A, Sobczak A, Lipinski L. 2017. Draft genome sequence of the type strain *Sphingopyxis bauzanensis* DSM 22271. Genome Announc 5:e01014-17. https://doi.org/10.1128/genomeA.01014-17.

Address correspondence to Leszek Lipinski, lechu@ibb.waw.pl.

The members of the genus *Sphingopyxis* (family *Sphingomonadaceae*) have been isolated from chemically contaminated environments, mainly oil- and petrol-polluted soil and water (1, 2). Microorganisms from the family *Sphingomonadaceae* have the ability to use polycyclic aromatic hydrocarbons as a sole carbon source (3). Here, we present the draft genome sequence of *Sphingopyxis bauzanensis* type strain DSM 22271 (=BZ30, =CGMCC 1.8959, =CIP 110136), isolated from hydrocarbon-contaminated soil (4); this is the only representative genome of this species. Because it was collected from a contaminated environment, we were interested in genes encoding proteins responsible for hydrocarbon degradation.

Genomic DNA was isolated as previously described (5). Illumina paired-end (with an average insert size of 450 bp) and Nextera mate pair libraries (with an average insert size of 8 kb) were prepared according to the manufacturer's protocols (a KAPA HTP DNA library preparation kit for Illumina sequencing and a Nextera mate pair sample prep kit, respectively). Whole-genome sequencing of *S. bauzanensis* strain DSM 22271^T was performed using the Illumina MiSeq platform (2 × 300 bp) and resulted in 487,322 paired reads for the paired-end library and 2,039,112 paired reads for the mate pair library. Reads from the paired-end library were processed as follows: adapters were removed using the Cutadapt script (6), and then the reads were filtered by length (>50 bp) and quality (*Q* value >30) (7). The mate pair reads were processed with NxTrim (8). Assembly was done using SPAdes version 3.9.1 (9). Contigs longer than 1 kb were deposited in GenBank and annotated using NCBI PGAP (10). The assembly consists of 28 scaffolds containing 4,258,005 bp with a GC content of 63.3%. The DSM 22271^T genome consists of 4,136 predicted genes, of which, 3,932 are protein-coding genes. The DSM 22271^T genome has 52 RNA genes, 46 tRNAs, 3 rRNAs, and 3 noncoding RNAs (ncRNAs), and 204 pseudogenes.

Twenty-six dioxygenases were predicted in the analyzed genome sequence, of which 13 were encoded on a single scaffold, number 6. This scaffold contains 68 open reading frames encoding proteins thought to be associated with the catabolism or active transport of aromatic compounds. Scaffold 6 is flanked with integrase-encoding genes, suggesting that it is a part of a catabolic transposon. Deeper analysis of proteins encoded on scaffold 6 showed their high similarity to proteins from a catabolic module described already on plasmid pNL1 from *Novosphingobium aromaticivorans* F199 (formerly *Sphingomonas aromaticivorans* F199) (11). The genes associated with biphenyl, xylene, and naphthalene degradation identified in *S. bauzanensis* strain DSM 22271^T

were situated in a similar orientation to the pNL1 plasmid instead of in one major rearrangement. The genes *bphD*, *bphE*, and *bphF*, together with coenzyme A-transferase, were localized upstream of the *bphB* gene, separating the *bphB* sequence from *xylA*. Such gene rearrangements result in a concentration of the catabolic enzymes in the genome compared to pNL1.

Preliminary studies of the *S. bauzanensis* DSM 22271^T genome sequence suggest that this strain is well adapted for degradation of high-molecular-weight polycyclic aromatic hydrocarbons and has potential in the bioremediation of polluted environments.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. NISK00000000. The version described in this paper is the first version NISK01000000.

ACKNOWLEDGMENTS

This work was supported by the European Union's European Regional Development Fund through the Innovative Economy Operational Program, 2007–2013 (project support agreement POIG.01.01.02-14-054/09-00). Experiments were carried out with the use of CePT infrastructure financed by the European Regional Development Fund through the Innovative Economy Operational Program, 2007–2013 (project support agreement POIG.02.02.00-14-024/08-00).

REFERENCES

1. Madueño L, Macchi M, Morelli IS, Coppotelli BM. 2016. Draft whole-genome sequence of *Sphingobium* sp. 22B, a polycyclic aromatic hydrocarbon–degrading bacterium from semiarid Patagonia, Argentina. Genome Announc 4(3):e00488-16. https://doi.org/10.1128/genomeA.00488-16.
2. Yan QX, Wang YX, Li SP, Li WJ, Hong Q. 2010. *Sphingobium qiguonii* sp. nov., a carbaryl-degrading bacterium isolated from a wastewater treatment system. Int J Syst Evol Microbiol 60:2724–2728. https://doi.org/10.1099/ijs.0.020362-0.
3. Zhong J, Luo L, Chen B, Sha S, Qing Q, Tam NFY, Zhang Y, Luan T. 2017. Degradation pathways of 1-methylphenanthrene in bacterial *Sphingobium* sp. MP9-4 isolated from petroleum-contaminated soil. Mar Pollut Bull 114:926–933. https://doi.org/10.1016/j.marpolbul.2016.11.020.
4. Zhang DC, Liu HC, Xin YH, Zhou YG, Schinner F, Margesin R. 2010. *Sphingopyxis bauzanensis* sp. nov., a psychrophilic bacterium isolated from soil. Int J Syst Evol Microbiol 60:2618–2622. https://doi.org/10.1099/ijs.0.018218-0.
5. Kaminski MA, Furmanczyk EM, Dziembowski A, Sobczak A, Lipinski L. 2017. Draft genome sequence of the type strain *Sphingopyxis witflariensis* DSM 14551. Genome Announc 5:e00924-17. https://doi.org/10.1128/genomeA.00924-17.
6. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10. https://doi.org/10.14806/ej.17.1.200.
7. Joshi N, Fass J. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files, version 1.33. https://github.com/najoshi/sickle.
8. O'Connell J, Schulz-Trieglaff O, Carlson E, Hims MM, Gormley NA, Cox AJ. 2015. Sequence analysis NxTrim: optimized trimming of Illumina mate pair reads. Bioinformatics 31:2035–2037. https://doi.org/10.1093/bioinformatics/btv057.
9. 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. https://doi.org/10.1089/cmb.2012.0021.
10. 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.
11. Romine MF, Stillwell LC, Wong K, Thurston SJ, Sisk EC, Sensen C, Gaasterland T, Fredrickson JIMK, Saffer JD, Acteriol JB. 1999. Complete sequence of a 184-kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199. J Bacteriol 181:1585–1602.