



HAL
open science

Complete Genome Sequences of Four Microbacterium Strains Isolated from Metal- and Radionuclide-Rich Soils

Philippe Ortet, Nicolas Gallois, Justine Long, Séverine Zirah, C. Berthomieu, J. Armengaud, Béatrice Alpha-Bazin, mohamed Barakat, Virginie Chapon

► To cite this version:

Philippe Ortet, Nicolas Gallois, Justine Long, Séverine Zirah, C. Berthomieu, et al.. Complete Genome Sequences of Four Microbacterium Strains Isolated from Metal- and Radionuclide-Rich Soils. Microbiology Resource Announcements, 2019, 8 (42), 10.1128/mra.00846-19 . cea-02363446

HAL Id: cea-02363446

<https://hal-cea.archives-ouvertes.fr/cea-02363446>

Submitted on 14 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Complete Genome Sequences of Four *Microbacterium* Strains Isolated from Metal- and Radionuclide-Rich Soils

 Philippe Ortet,^b Nicolas Gallois,^a Justine Long,^b Séverine Zirah,^c Catherine Berthomieu,^a Jean Armengaud,^d Béatrice Alpha-Bazin,^d Mohamed Barakat,^b Virginie Chapon^a

^aAix-Marseille Université, CEA, CNRS, BIAM, Laboratoire des Interactions Protéine Métal, Saint-Paul-lez-Durance, France

^bAix-Marseille Université, CEA, CNRS, BIAM, Laboratoire d'Écologie Microbienne de la Rhizosphère et d'Environnements Extrêmes, Saint-Paul-lez-Durance, France

^cUnité Molécules de Communication et Adaptation des Microorganismes (MCAM), Muséum National d'Histoire Naturelle, CNRS, Paris, France

^dLaboratoire Innovations Technologiques pour la Détection et le Diagnostic (Li2D), Service de Pharmacologie et Immunoanalyse (SPI), CEA, INRA, Bagnols-sur-Cèze, France

ABSTRACT Here, we present the genome sequences of four *Microbacterium* strains, which were isolated at different locations in Europe from metal- or radionuclide-rich soils. High-quality complete genome sequences were obtained with PacBio and Illumina data sets with an original two-step procedure.

The *Microbacterium* genus is composed of ubiquitous high-GC-content Gram-positive *Actinobacteria*. Despite its presence in many environments and its potential in bioremediation processes (1–7), the bacterial genus *Microbacterium* is underrepresented in the genome databases and is still poorly studied. At the time of writing, there were more than 100 *Microbacterium* species with a valid name (<http://www.bacterio.net/microbacterium.html>) and 323 genomes, including 28 complete genomes, available in the databases (<https://www.ncbi.nlm.nih.gov/assembly/?term=microbacterium>). Here, we report the complete genome sequences of four members of this genus, which were isolated from metal-rich or radionuclide-rich soil samples. The *Microbacterium oleivorans* strain A9 was isolated from radionuclide-contaminated soil in Chernobyl (8) and exhibits a high uranium tolerance (9). Its draft genome sequence was previously published (10). The strains ViU2A and ViU22^T were isolated from natural uranium-rich soil samples collected in France (the soil composition is described in reference 11; the strain ViU22^T is the type strain of the species *Microbacterium lemovicicum* [12]). The strain HG3 was cultured from metal-rich black sand from Iceland and has been established as mercury tolerant (13).

Bacteria were cultured in LB at 30°C until late exponential growth phase. High-quality genomic DNA was extracted from cells using the DNeasy blood and tissue kit (Qiagen) following the manufacturer's instructions for Gram-positive bacteria. DNA integrity was checked using agarose gel electrophoresis, and DNA purity and concentration were measured using a NanoDrop spectrophotometer (Thermo Fisher Scientific). Whole-genome shotgun sequencing was carried out with PacBio long-read technology and the Illumina short-read technology. Library preparation and single-end sequencing with 100-base read lengths and a HiSeq 2000 instrument (Illumina) were performed by GenoScreen (Lille, France). Library preparation and long-read sequencing were performed by Eurofins Genomics Europe Shared Services GmbH (formerly GATC Biotech AG, Germany). Library preparation incorporated adaptor sequences compatible with PacBio RS II sequencing technology (single-molecule real-time [SMRT] sequencing) using proprietary methods of Eurofins Genomics Europe Shared Services GmbH. Sequencing was carried out on a PacBio RS II system with SMRT technology. For both sequencing methods, the total number of reads is indicated in Table 1.

Citation Ortet P, Gallois N, Long J, Zirah S, Berthomieu C, Armengaud J, Alpha-Bazin B, Barakat M, Chapon V. 2019. Complete genome sequences of four *Microbacterium* strains isolated from metal- and radionuclide-rich soils. *Microbiol Resour Anounc* 8:e00846-19. <https://doi.org/10.1128/MRA.00846-19>.

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

Copyright © 2019 Ortet 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 Philippe Ortet, philippe.ortet@cea.fr.

Received 17 July 2019

Accepted 4 August 2019

Published 17 October 2019

TABLE 1 Summary of complete genome sequences

Characteristic	Data for strain:			
	A9	Hg3	ViU2a	ViU22
GenBank accession no.	CP031421	CP031422	CP031338	CP031423
Raw data accession no. (Illumina)	SRR8416123	SRR8416223	SRR8416225	SRR8417352
No. of reads (Illumina)	18,500,154	19,895,949	15,648,367	13,317,740
Raw data accession no. (PacBio)	SRR8416122	SRR8416222	SRR8416224	SRR8417351
No. of reads (PacBio)	144,950	154,761	172,715	164,482
Mean read length (bp) (PacBio)	8,474	8,380	7,755	8,270
Genome length (Mb)	2.99	3.9	3.8	3.6
G+C content (%)	69	68.2	68.3	70.8
No. of protein-coding genes	2,880	3,764	3,708	3,297
No. of tRNA genes	50	51	54	52
No. of rRNA genes	6	6	6	6
No. of TCSs	62	70	75	55
No. of TFs	156	274	280	191

In the first step, genome *de novo* assembly was performed on the PacBio reads using Canu software version 1.0 with default parameters (14). Trimming and circularization in a single genome assembly were done using this tool. Then, to improve the quality of the genome sequence, a correction was made using the Illumina reads with the MIRA assembler version 4.0.2 (15) on the complete genome sequence (assembled using Canu and oriented on the origin of replication [*oriC*] by an in-house pipeline) as reference. We used the default configuration for MIRA except for the option *job*, which was set to "genome," "mapping," and "accurate," and for the option parameters, which were set to SOLEXA_SETTINGS -CL:pec=yes COMMON_SETTINGS -NW:cac=no -SK:mmhr=1.

The accession numbers, assembly metrics, and genome characteristics for each genome are listed in Table 1. All chromosomal sequences were circularized and oriented with the predicted *oriC* region as the beginning of the sequences. Taxonomic assignment at the species level was provided by the NCBI using their quality control test for bacterial genomes. This test uses average nucleotide identity (ANI), which compares the submitted genome sequence against the genomes of the type strains and proxytype strains that are already in GenBank, as described in reference 16. The chromosomal sequences were annotated with the NCBI Prokaryotic Genome Annotation Pipeline version 1.2 with default parameters (17). Two-component system proteins (TCSs) and transcription factor proteins (TFs) were identified using the P2RP Web server version 2.7 with default parameters (18).

Data availability. The raw data and whole-genome sequences of the four *Microbacterium* strains have been deposited in the GenBank database under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This study was financed by the Toxicology Program of the French Alternative Energies and Atomic Energy Commission (CEA) and by the NEEDS-PF Resources Program (CEA, CNRS, ORANO). Nicolas Gallois is the recipient of a Ph.D. grant funded by the CEA.

REFERENCES

- Wang H, Xiang T, Wang Y, Song J, Zhai Y, Chen X, Li Y, Zhao B, Zhao B, Ruan Z. 2014. *Microbacterium petrolarium* sp. nov., isolated from an oil-contaminated water sample. *Int J Syst Evol Microbiol* 64:4168–4172. <https://doi.org/10.1099/ijs.0.061119-0>.
- Aniszewski E, Peixoto RS, Mota FF, Leite SGF, Rosado AS. 2010. Bioemulsifier production by *Microbacterium* sp. strains isolated from mangrove and their application to remove cadmium and zinc from hazardous industrial residue. *Braz J Microbiol* 41:235–245. <https://doi.org/10.1590/S1517-83822010000100033>.
- Kumar R, Nongkhlaw M, Acharya C, Joshi SR. 2013. Uranium (U)-tolerant bacterial diversity from U ore deposit of Domiasiat in North-East India and its prospective utilisation in bioremediation. *Microbes Environ* 28:33–41. <https://doi.org/10.1264/jsme2.me12074>.
- Sarkar A, Sar P, Islam E. 2016. Hexavalent chromium reduction by *Microbacterium oleivorans* A1: a possible mechanism of chromate -detoxification and -bioremediation. *Recent Pat Biotechnol* 9:116–129. <https://doi.org/10.2174/187220830902160308192126>.
- Avramov AP, Couger MB, Hartley EL, Land C, Wellendorf R, Hanafy RA, Budd C, French DP, Hoff WD, Youssef N. 2016. Draft genome sequence of *Microbacterium oleivorans* strain Wellendorf implicates heterotrophic versatility and bioremediation potential. *Genom Data* 10:54–60. <https://doi.org/10.1016/j.gdata.2016.09.005>.

6. Hirth N, Topp E, Dörfler U, Stupperich E, Munch JC, Schroll R. 2016. An effective bioremediation approach for enhanced microbial degradation of the veterinary antibiotic sulfamethazine in an agricultural soil. *Chem Biol Technol Agric* 3:29. <https://doi.org/10.1186/s40538-016-0080-6>.
7. Bollmann A, Palumbo AV, Lewis K, Epstein SS. 2010. Isolation and physiology of bacteria from contaminated subsurface sediments. *Appl Environ Microbiol* 76:7413–7419. <https://doi.org/10.1128/AEM.00376-10>.
8. Chapon V, Piette L, Vesvres M-H, Coppin F, Marrec CL, Christen R, Theodorakopoulos N, Février L, Levchuk S, Martin-Garin A, Berthomieu C, Sergeant C. 2012. Microbial diversity in contaminated soils along the T22 trench of the Chernobyl experimental platform. *Appl Geochem* 27:1375–1383. <https://doi.org/10.1016/j.apgeochem.2011.08.011>.
9. Theodorakopoulos N, Chapon V, Coppin F, Floriani M, Vercouter T, Sergeant C, Camilleri V, Berthomieu C, Février L. 2015. Use of combined microscopic and spectroscopic techniques to reveal interactions between uranium and *Microbacterium* sp. A9, a strain isolated from the Chernobyl exclusion zone. *J Hazard Mater* 285:285–293. <https://doi.org/10.1016/j.jhazmat.2014.12.018>.
10. Ortet P, Gallois N, Long J, Barakat M, Chapon V. 2017. Draft genome sequence of *Microbacterium oleivorans* strain A9, a bacterium isolated from Chernobyl radionuclide-contaminated soil. *Genome Announc* 5:e00092-17. <https://doi.org/10.1128/genomeA.00092-17>.
11. Mondani L, Benzerara K, Carriere M, Christen R, Mamindy-Pajany Y, Février L, Marmier N, Achouak W, Nardoux P, Berthomieu C, Chapon V. 2011. Influence of uranium on bacterial communities: a comparison of natural uranium-rich soils with controls. *PLoS One* 6:e25771. <https://doi.org/10.1371/journal.pone.0025771>.
12. Mondani L, Piette L, Christen R, Bachar D, Berthomieu C, Chapon V. 2013. *Microbacterium lemovicicum* sp. nov., a bacterium isolated from a natural uranium-rich soil. *Int J Syst Evol Microbiol* 63:2600–2606. <https://doi.org/10.1099/ijs.0.048454-0>.
13. Francois F, Lombard C, Guigner JM, Soreau P, Brian-Jaisson F, Martino G, Vandervennet M, Garcia D, Molinier AL, Pignol D, Peduzzi J, Zirah S, Rebuffat S. 2012. Isolation and characterization of environmental bacteria capable of extracellular biosorption of mercury. *Appl Environ Microbiol* 78:1097–1106. <https://doi.org/10.1128/AEM.06522-11>.
14. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
15. Chevreux B WT, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. *In* Computer science and biology: proceedings of the German Conference on Bioinformatics, GCB '99 GCB, Hannover, Germany.
16. Federhen S, Rossello-Mora R, Klenk H-P, Tindall BJ, Konstantinidis KT, Whitman WB, Brown D, Labeda D, Ussery D, Garrity GM, Colwell RR, Hasan N, Graf J, Parte A, Yarza P, Goldberg B, Sichtig H, Karsch-Mizrachi I, Clark K, McVeigh R, Pruitt KD, Tatusova T, Falk R, Turner S, Madden T, Kitts P, Kimchi A, Klimke W, Agarwala R, DiCuccio M, Ostell J. 2016. Meeting report: GenBank microbial genomic taxonomy workshop (12–13 May, 2015). *Stand Genomic Sci* 11:15. <https://doi.org/10.1186/s40793-016-0134-1>.
17. 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>.
18. Barakat M, Ortet P, Whitworth DE. 2013. P2RP: a Web-based framework for the identification and analysis of regulatory proteins in prokaryotic genomes. *BMC Genomics* 14:269. <https://doi.org/10.1186/1471-2164-14-269>.