



Swansea University
Prifysgol Abertawe



Cronfa - Swansea University Open Access Repository

This is an author produced version of a paper published in :
Journal of Biotechnology

Cronfa URL for this paper:

<http://cronfa.swan.ac.uk/Record/cronfa29536>

Paper:

Zhang, H., Sun, H., Yang, R., Li, S., Zhou, M., Gao, T., An, L., Chen, X. & Dyson, P. (2016). Complete genome sequence of a psychotrophic *Pseudarthrobacter sulfonivorans* strain Ar51 (CGMCC 4.7316), a novel crude oil and multi benzene compounds degradation strain. *Journal of Biotechnology*, 231, 81-82.

<http://dx.doi.org/10.1016/j.jbiotec.2016.04.010>

This article is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Authors are personally responsible for adhering to publisher restrictions or conditions. When uploading content they are required to comply with their publisher agreement and the SHERPA RoMEO database to judge whether or not it is copyright safe to add this version of the paper to this repository.

<http://www.swansea.ac.uk/iss/researchsupport/cronfa-support/>

Accepted Manuscript

Title: Complete genome sequence of a psychotrophic *Pseudarthrobacter sulfonivorans* strain Ar51 (CGMCC 4.7316), a novel crude oil and multi benzene compounds degradation strain

Author: Hua Zhang Haili Sun Ruiqi Yang Shuyan Li Meng
Zhou Tianpeng Gao Lizhe An Ximing Chen Paul Dyson

PII: S0168-1656(16)30189-4
DOI: <http://dx.doi.org/doi:10.1016/j.jbiotec.2016.04.010>
Reference: BIOTEC 7497

To appear in: *Journal of Biotechnology*

Received date: 29-3-2016
Accepted date: 6-4-2016

Please cite this article as: Zhang, Hua, Sun, Haili, Yang, Ruiqi, Li, Shuyan, Zhou, Meng, Gao, Tianpeng, An, Lizhe, Chen, Ximing, Dyson, Paul, Complete genome sequence of a psychotrophic *Pseudarthrobacter sulfonivorans* strain Ar51 (CGMCC 4.7316), a novel crude oil and multi benzene compounds degradation strain. *Journal of Biotechnology* <http://dx.doi.org/10.1016/j.jbiotec.2016.04.010>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Complete genome sequence of a psychotrophic *Pseudarthrobacter sulfonivorans* strain Ar51 (CGMCC 4.7316), a novel crude oil and multi benzene compounds degradation strain

Hua Zhang^{a,1}, Haili Sun^{a,f,1}, Ruiqi Yang^{c,g}, Shuyan Li^c, Meng Zhou^a, Tianpeng Gao^f, Lizhe An^{a*}, Ximing Chen^{b,c*}, Paul Dyson^d

Ministry of Education Key Laboratory of Cell Activities and Stress Adaptations, School of Life Sciences^a, State Key Laboratory of Applied Organic Chemistry, College of Chemistry & Chemical Engineering^c, Lanzhou University; Centre of Urban Ecology and Environmental Biotechnology; Lanzhou City University^f; Chemical and Biological Engineering, Lanzhou jiaotong University^g; Key Laboratory of Extreme Environmental Microbial Resources and Engineering of Gansu Province^e; Key Laboratory of Desert and Desertification, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences^b; Lanzhou 730000, China; Institute of Life Science, College of Medicine, Swansea University^d, Singleton Park, Swansea SA2 8PP, UK

* Corresponding authors

1 These authors contributed equally to this work.

To whom correspondence should be addressed.

Email: lizhean@lzu.edu.cn or lzuchen@163.com

Highlights:

Pseudarthrobacter sulfonivorans strain Ar51, a psychotrophic bacterium isolated from the Tibet permafrost of China, can degrade crude oil and multi benzene compounds efficiently in low temperature. Here we report the complete genome sequence of this bacterium. The complete genome sequence of *Pseudarthrobacter sulfonivorans* strain Ar51, consisting of a cycle chromosome with a size of 5.04 Mbp and a cycle plasmid with a size of 12.39kbp. The availability of this genome sequence allows us to investigate the genetic basis of crude oil degradation and adaptation to growth in a nutrient-poor permafrost environment.

Abstract

Pseudarthrobacter sulfonivorans strain Ar51, a psychotrophic bacterium isolated from the Tibet permafrost of China, can degrade crude oil and multi benzene compounds efficiently in low temperature. Here we report the complete genome sequence of this bacterium. The complete genome sequence of *Pseudarthrobacter sulfonivorans* strain Ar51, consisting of a cycle chromosome with a size of 5.04 Mbp and a cycle plasmid with a size of 12.39kbp. The availability of this genome sequence allows us to investigate the genetic basis of crude oil degradation and adaptation to growth in a nutrient-poor permafrost environment.

Arthrobacter are most frequently isolated in soils and play important roles in biogeochemical cycles and decontamination (Unell et al., 2008). *A. sulfonivorans*, which is a novel methylotrophic strain, has the ability to degrade the dimethyl sulfoxide (DMSO) efficiently (Borodina et al., 2002). Based on phylogenetic groupings, 16S rRNA gene sequence similarities and homogeneity in peptidoglycan types, quinone systems and polar lipid profiles, *A. sulfonivorans* was reclassified in the genus *Pseudarthrobacter* gen. nov. *Pseudarthrobacter sulfonivorans* strain Ar51 was isolated from permafrost of the Tibet in SW China, is a psychrotrophic bacterium that can grow at temperatures between -4 °C to 25 °C, with an optimum growth temperature of 20°C. Biodegradation experiment has indicated that this strain can utilize the crude oil, alkane and multi benzene compounds efficiently in low temperature. Here we present the complete genome sequence of *Pseudarthrobacter sulfonivorans* strain Ar51 (previous name is *Arthrobacter oxydans* Ar51).

The raw sequence was generated using a shotgun approach employing the PacBio® RS II system and whole-genome shotgun technology with the Illumina HiSeq 2000 system yielding ~410× coverage (Roberts et al., 2013). Genomic DNA was sheared to an average fragment size of approximately 20 kb, and subsequently constructed to a SMRTbell™ library with a recovery of 40%. To generate the longest reads possible for this genome, the 20 kb SMRTbell library was size-selected using the BluePippin™ size selection system to remove SMRTbells less than 15 kb. Approximately, 14% of the final SMRTbell library was recovered with an average insert size of approximately 20 kb. The size-selected SMRTbell library was bound with P6 polymerase and sequenced with C4 chemistry and loaded using the one-cell-per-well protocol in the RS II instrument. Three SMRT Cells with 1.45 Gb post-filtering data were used for assembly. De novo assembly was carried out using the Hierarchical Genome Assembly Process (HGAP) version 3 (PacBio DevNet; Pacific Biosciences) workflow, including consensus polishing with Quiver as available in the SMRT® Analysis v 2.3.

The annotation for *Pseudarthrobacter sulfonivorans* strain Ar51 was conducted with NCBI Prokaryotic Genome Annotation Pipeline (Pruitt et al., 2012). The chromosome of Ar51 consists of 5,043,757 bp, the G+C content is 64.7%, 4390 coding sequence were identified, 12 rRNA genes and 50 tRNA genes. The plasmid of Ar51 consists of 12,394 bp, the G+C content is 56.3%, 3 coding sequence and 5 rRNA genes were identified (Table 1). Over 58% (2548) of the predicted open reading frames (ORFs) present in the Ar51 genome could not be assigned a putative function.

Previous studies identified the alkane 1-monooxygenase (*alkB*) is the key factor to degrade the alkane in the bacteria (Belhaj et al., 2002). Unlike other related alkane biodegrade clusters only contain one copy of *alkB*, two copies of

alkB genes (AU252_RS14035, AU252_RS11540) were identified in the Ar51 genome. Additionally, the Ar51 also contains multi-aromatic compounds degrade genes, such as biphenyl degradation genes cluster (AU252_RS17305, AU252_RS17310, AU252_RS17315), benzoate degradation genes (*benK*, AU252_RS22695; *benD*, AU252_RS22710; AU252_RS22720; AU252_RS22725; AU252_RS22745; *benK*, AU252_RS00390).

The potential of *Pseudarthrobacter sulfonivorans* strain Ar51 to produce secondary metabolites was analyzed with the secondary metabolites search tool antiSMASH(Weber et al., 2015), which indicated that 5 putative gene clusters involved in the biosynthesis of different natural products (1 Bacteriocin, 1 typeIII PKS, 1 Siderophore, 2 unspecified cluster) are located in the chromosome and no gene cluster is located in the plasmid. Iron-uptake systems include a unique low-pH-induced iron transporter system and utilization systems. Siderophores are small, high-affinity iron chelating compounds secreted by many microorganisms as a means to acquire sufficient iron for their growth (Neilands, 1995). Ar51 contains a gene cluster to direct synthesis of Desferrioxamine E, a pathway it shares with other *Arthrobacter* species.

Compared with the psychrotrophic *Arthrobacter* strain A3, the Ar51 has the stronger ability to synthesis trehalose as the osmotic stress protector (Chen et al., 2011; Sun et al., 2016). In the genome of Ar51, it contains three pathways to synthesis trehalose: the OtsA/B (AU252_RS06465, AU252_RS06470, AU252_RS12615), TreY/TreZ (AU252_RS17210, AU252_RS17205) and TreS (AU252_RS10705, AU252_RS06525) pathways. Glutathione (GSH) is an important antioxidant in plants, animals, fungi, and some bacteria and archaea, preventing damage to important cellular components caused by reactive oxygen species such as free radicals, peroxides, lipid peroxides and heavy metals(Pompella et al., 2003). However, most of *Arthrobacter* species has no ability to synthesis GSH (Sun et al., 2016). Ar51 contains GSH synthesis pathway, consisting of gamma-glutamyltranspeptidase (AU252_RS22220) and glutathione synthetase (AU252_RS14885). In terms of cold adaption, the genome of this psychrotrophic bacterium has two aquaporin Z genes (AU252_RS16525, AU252_RS18930). Although bacterial aquaporin proteins can theoretically contribute to osmoregulation, studies indicate that they likely function to improve freeze tolerance under rapid-freezing conditions (Tanghe et al., 2006).

The complete genome chromosome and plasmid sequences have been deposited in GenBank database with accession number CP013747.1 and CP013748.1 respectively. This strain has been deposited at the China General Microbiological Culture Collection Center (CGMCC 4.7316).

ACKNOWLEDGEMENTS

This study was supported by the National Key Basic Research Program of China (973) (2013CB429904), the State Key Program of the National Natural

Science Foundation of China (31230014), the Major Project of Cultivating New Varieties of Transgenic Organisms (2014ZX08009-003-002), the Fundamental Research Funds for the Central Universities (lzujbky-2014-100 and lzujbky-2015-bt05), the National Natural Science Foundation of China (grant 31400437 and 31460162) and UK BBSRC China Partnering Grant (BB/J020419/1).

Table 1
General genome features of *Pseudarthrobacter sulfonivorans* strain Ar51

Feature	Chromosome	Plasmid
Length (bp)	5,043,757	12,394
G + C content (%)	64.7	56.3
Protein coding genes	4390	3
rRNA number	12	5
tRNA number	50	0
GenBank accession	CP013747.1	CP013748.1

References

- Belhaj, A., Desnoues, N., Elmerich, C., (2002) Alkane biodegradation in *Pseudomonas aeruginosa* strains isolated from a polluted zone: identification of *alkB* and *alkB*-related genes. *Research in microbiology* 153, 339-344.
- Borodina, E., Kelly, D.P., Schumann, P., Rainey, F.A., Ward-Rainey, N.L., Wood, A.P., (2002) Enzymes of dimethylsulfone metabolism and the phylogenetic characterization of the facultative methylotrophs *Arthrobacter sulfonivorans* sp. nov., *Arthrobacter methylotrophus* sp. nov., and *Hyphomicrobium sulfonivorans* sp. nov. *Archives of microbiology* 177, 173-183.
- Chen, X.M., Jiang, Y., Li, Y.T., Zhang, H.H., Li, J., Chen, X., Zhao, Q., Zhao, J., Si, J., Lin, Z.W., Zhang, H., Dyson, P., An, L.Z., (2011) Regulation of expression of trehalose-6-phosphate synthase during cold shock in *Arthrobacter* strain A3. *Extremophiles : life under extreme conditions* 15, 499-508.
- Neilands, J.B., (1995) Siderophores: structure and function of microbial iron transport compounds. *The Journal of biological chemistry* 270, 26723-26726.
- Pompella, A., Visvikis, A., Paolicchi, A., De Tata, V., Casini, A.F., (2003) The changing faces of glutathione, a cellular protagonist. *Biochemical pharmacology* 66, 1499-1503.
- Pruitt, K.D., Tatusova, T., Brown, G.R., Maglott, D.R., (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic acids research* 40, D130-135.
- Roberts, R.J., Carneiro, M.O., Schatz, M.C., (2013) The advantages of SMRT sequencing. *Genome biology* 14, 405.
- Sun, H., Gao, T., Chen, X., Hitchings, M.D., Li, S., Chen, T., Zhang, H., An, L., Dyson, P., (2016) Complete genome sequence of a psychotrophic *Arthrobacter* strain A3 (CGMCC 1.8987), a novel long-chain hydrocarbons producer. *Journal of biotechnology* 222, 23-24.
- Tanghe, A., Van Dijck, P., Thevelein, J.M., (2006) Why do microorganisms have aquaporins? *Trends in microbiology* 14, 78-85.
- Unell, M., Nordin, K., Jernberg, C., Stenstrom, J., Jansson, J.K., (2008) Degradation of mixtures of phenolic compounds by *Arthrobacter chlorophenolicus* A6. *Biodegradation* 19, 495-505.
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Brucoleri, R., Lee, S.Y., Fischbach, M.A., Muller, R., Wohlleben, W., Breitling, R., Takano, E., Medema, M.H., (2015) antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic acids research* 43, W237-243.