



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/cronfa26369>

Paper:

Sun, H., Gao, T., Chen, X., Hitchings, M., 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.

<http://dx.doi.org/10.1016/j.jbiotec.2016.02.010>

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 *Arthrobacter* strain A3 (CGMCC 1.8987), a novel long-chain hydrocarbons producer

Author: Haili Sun Tianpeng Gao Ximing Chen Matthew D. Hitchings Shuyan Li Tao Chen Hua Zhang Lizhe An Paul Dyson



PII: S0168-1656(16)30055-4
DOI: <http://dx.doi.org/doi:10.1016/j.jbiotec.2016.02.010>
Reference: BIOTEC 7404

To appear in: *Journal of Biotechnology*

Received date: 2-2-2016
Accepted date: 3-2-2016

Please cite this article as: Sun, Haili, Gao, Tianpeng, Chen, Ximing, Hitchings, Matthew D., Li, Shuyan, Chen, Tao, Zhang, Hua, An, Lizhe, Dyson, Paul, Complete genome sequence of a psychotrophic *Arthrobacter* strain A3 (CGMCC 1.8987), a novel long-chain hydrocarbons producer. *Journal of Biotechnology* <http://dx.doi.org/10.1016/j.jbiotec.2016.02.010>

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 *Arthrobacter* strain A3 (CGMCC 1.8987), a novel Long-Chain Hydrocarbons producer

Haili Sun^{a,f,1}, Tianpeng Gao^{f,1}, Ximing Chen^{c,d}, Matthew D. Hitchings^b, Shuyan Li^e, Tao Chen^a, Hua Zhang^{a*}, Lizhe An^{a*}, Paul Dyson^b

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^e, Lanzhou University; Centre of Urban Ecology and Environmental Biotechnology; Lanzhou City University^f; Key Laboratory of Extreme Environmental Microbial Resources and Engineering of Gansu Province^c; Key Laboratory of Desert and Desertification, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences^d; Lanzhou 730000, China, Institute of Life Science; College of Medicine, Swansea University^b, 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

Highlights

This article is a genome announcement, which has report a complete genome sequence of a psychotrophic *Arthrobacter* strain A3. *Arthrobacter* strain A3, a psychotrophic bacterium isolated from the Tian Shan Mountain of China, can degrade the cellulose and synthesis the long-chain hydrocarbons efficiently in low temperature. The availability of this genome sequence allows us to investigate the genetic basis of adaptation to growth in a nutrient-poor permafrost environment and to evaluate of the biofuel-synthetic potential of this species. The complete genome chromosome and plasmid sequences have been deposited in GenBank database with accession number CP013745.1 and CP013746.1 respectively. This strain has been deposited at the China General Microbiological Culture Collection Center (CGMCC 1.8987).

Abstract

Arthrobacter strain A3, a psychotrophic bacterium isolated from the Tian Shan Mountain of China, can degrade the cellulose and synthesis the long-chain hydrocarbons efficiently in low temperature. Here we report the complete genome sequence of this bacterium. The complete genome sequence of *Arthrobacter* strain A3, consisting of a cycle chromosome with a size of 4.26 Mbp and a cycle plasmid with a size of 194kbp. In this genome, a hydrocarbon biosynthesis gene cluster (*oleA*, *oleB/oleC* and *oleD*) was identified. To resistant the extreme environment, this strain contains a unique mycothiol-biosynthetic pathway (*mshA-D*), which has not been found in other *Arthrobacter* species before. The availability of this genome sequence allows us to investigate the genetic basis of adaptation to growth in a nutrient-poor permafrost environment and to evaluate of the biofuel-synthetic potential of this species.

Arthrobacter are most frequently isolated in soils and play important roles in biogeochemical cycles and decontamination (Unell et al., 2008). Additionally, members of the genus *Arthrobacter*, *Shewanella* and *Micrococcus* have the ability to produce long-chain hydrocarbons as biofuel (Beller et al., 2010; Frias et al., 2009; Sukovich et al., 2010). *Arthrobacter* Strain A3 was isolated from permafrost of the Tian Shan Mountain in NW China, is a psychotrophic bacterium that can grow at temperatures between -4 °C to 25 °C, with an optimum growth temperature of 20°C (Chen et al., 2011). Gas chromatography-mass spectrometry (GC-MS) has indicated that this strain can produce hydrocarbons as biofuel at low temperatures. Here we present the complete genome sequence of *Arthrobacter* strain A3 (previous name is *Arthrobacter xinjiangense* TSBY90).

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 ~430× 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 *Arthrobacter* strain A3 was conducted with NCBI Prokaryotic Genome Annotation Pipeline (Pruitt et al., 2012). The chromosome of A3 consists of 4,264,061 bp, the G+C content is 60.78%, 3735 coding sequence were identified, and 52 tRNA genes. The plasmid of A3 consists of 194,226 bp, the G+C content is 57.11% and 175 coding sequence were identified (Table1). Over 58% (2267) of the predicted open reading frames (ORFs) present in the A3 genome could not be assigned a putative function. Previous studies identified the *oleABCD* genes involved in head-to-head olefinic hydrocarbon biosynthesis (Sukovich et al., 2010). In the A3 genome, one ole gene cluster (MB46_09735- MB46_09745) was identified. Like related hydrocarbon biosynthesis clusters from other sequenced *Arthrobacter* species, this gene cluster also contains one copie of oleA (MB46_09735), a copie of oleB/oleC fusion gene (MB46_09740) and oleD (MB46_09745) gene.

The potential of *Arthrobacter* strain A3 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 NRPS, 1 typeI PKS, 1 Siderophore, 1 Terpene and 1 unspecified cluster) are located in the chromosome and 1 bacteriocin putative gene cluster is located in the plasmid.

Other features of the genome indicate that *Arthrobacter* strain A3 has evolved a metabolism adapted to this nutrient-poor environment, specifically by recruiting the potential to utilize a variety of carbon sources from plants. In contrast to other *Arthrobacter* species, A3 has a unique predicted beta-glucoside metabolism with potential to utilise plant-derived carbohydrates that include salicin, arbutin, cellobiose and aesculin. In addition, A3 has the potential to utilize fructooligosaccharides and raffinose. The genome analysis of *Arthrobacter* strain A3 indicates considerable investment, greater than in other sequenced *Arthrobacter* species, in iron acquisition capability. 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. A3 contains a gene cluster to direct synthesis of Desferrioxamine E, a pathway it shares with other *Arthrobacter* species.

The ability to cope with oxidative, osmotic and low temperature stresses is likely a key factor for life in the permafrost. Previous studies have indicated that many Actinobacteria can synthesis a mycothiol (1D-myo-inosityl 2-(N-acetylcysteiny) amido-2-deoxy-a-D-glucopyranoside) to provide protection against oxidative stress (Fahey, 2001; Newton et al., 2008). However, until now, the mycothiol-biosynthetic pathway has not been found in representatives of the *Arthrobacter*, *Agromyces*, or *Actinomyces* genera (Newton et al., 1996). Here we report the first example of a potential mycothiol-biosynthetic pathway in an *Arthrobacter*; the A3 genome contains genes for myo-Inositol-1-phosphate synthase (MB46_13210), mshA (MB46_13160), mshB (MB46_12525), mshC (MB46_05370) and mshD (MB46_13645). The A3 genome also contains two pathways for synthesis of the osmotic stress protector trehalose: the OtsA/B (MB46_00090, MB46_00085) and TreS (MB46_19105) pathways (Chen et al., 2011). Additionally, in contrast to other sequenced *Arthrobacter* species, A3 contains a predicted hyperosmotic potassium uptake system, consisting of trkA (MB46_07900, MB46_07905, MB46_17505) and trkH (MB46_17510). In terms of cold adaption, the genome of this psychrotrophic bacterium has two aquaporin Z genes (MB46_15885, MB46_10660). 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). In addition, A3 encodes a type 1 antifreeze protein (AFP, MB46_12945). AFPs bind to small ice crystals to inhibit growth and recrystallization of ice that would otherwise be fatal (Daley et al., 2002).

ACCEPTED MANUSCRIPT

The complete genome chromosome and plasmid sequences have been deposited in GenBank database with accession number CP013745.1 and CP013746.1 respectively. This strain has been deposited at the China General Microbiological Culture Collection Center (CGMCC 1.8987).

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-2015-bt05 and lzujbky-2015-ct01), the National Natural Science Foundation of China (grant 31400437) and UK BBSRC China Partnering Grant (BB/J020419/1).

References

- Beller, H.R., Goh, E.B., Keasling, J.D., (2010) Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Applied and environmental microbiology* 76, 1212-1223.
- 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.
- Daley, M.E., Spyropoulos, L., Jia, Z., Davies, P.L., Sykes, B.D., (2002) Structure and dynamics of a beta-helical antifreeze protein. *Biochemistry* 41, 5515-5525.
- Fahey, R.C., (2001) Novel thiols of prokaryotes. *Annual review of microbiology* 55, 333-356.
- Frias, J.A., Richman, J.E., Wackett, L.P., (2009) C29 olefinic hydrocarbons biosynthesized by *Arthrobacter* species. *Applied and environmental microbiology* 75, 1774-1777.
- Newton, G.L., Arnold, K., Price, M.S., Sherrill, C., Delcardayre, S.B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R.C., Davis, C., (1996) Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes. *Journal of bacteriology* 178, 1990-1995.
- Newton, G.L., Buchmeier, N., Fahey, R.C., (2008) Biosynthesis and functions of mycothiol, the unique protective thiol of Actinobacteria. *Microbiology and molecular biology reviews : MMBR* 72, 471-494.
- 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.
- Sukovich, D.J., Seffernick, J.L., Richman, J.E., Gralnick, J.A., Wackett, L.P., (2010) Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. *Applied and environmental microbiology* 76, 3850-3862.
- 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.

Table 1
General genome features of *Arthrobacter* strain A3

Feature	Chromosome	Plasmid
Length (bp)	4,264,061	194,226
G + C content (%)	60.78	57.11
Protein coding genes	3735	175
rRNA number	20	0
tRNA number	52	0
GenBank accession	CP013745.1	CP013746.1