



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

The data of genomic and phenotypic profiles of the N-acyl homoserine lactone-producing algicidal bacterium *Stenotrophomonas rhizophila* GA1



Panqing Yin^a, Qin Zhang^b, Jianming Zhu^c, Guoqiang Wu^a,
Sanjun Yin^d, Zhenhua Ma^d, Jin Zhou^{c,*}

^a The School of Life Science and Engineering, Lanzhou University of Technology, Lanzhou 730050, Gansu Province, PR China

^b Jiangxi University of Finance and Economics, 330013 Jiangxi Province, PR China

^c The Division of Ocean Science and Technology, Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, Guangdong Province, PR China

^d Hengchuang Genetic Biotech Co., Ltd, Shenzhen 518055, Guangdong Province, PR China

ARTICLE INFO

Article history:

Received 25 August 2018

Received in revised form

13 October 2018

Accepted 17 October 2018

Available online 23 October 2018

ABSTRACT

Herein, an algicidal strain, *Stenotrophomonas rhizophila* GA1, was isolated from a marine dinoflagellate and its genome was sequenced using next-generation sequencing technology. The genome size of *S. rhizophila* GA1 was determined to be 5.92 Mb with a G+C content of 62.39%, comprising eight scaffolds of 67 contigs. A total of 4579 functional proteins were assigned according to COG categories. *In silico* genome annotation protocols identified multiple putative LuxI-like genes located in the upstream position at contig 4. The thin-layer chromatography analysis showed that three kinds of acyl homoserine lactone (AHL) signals could be produced by *S. rhizophila* GA1. This work describes an algicidal bacterium capable of generating AHL molecules for its ecological adaptation. The annotated genome sequence of this strain may represent a valuable tool for studying algae-bacterium interactions and developing microbial methods to control harmful algae. The genome scaffolds generated are available

* Correspondence to: Marine-Building, Graduate School at Shenzhen, Tsinghua University, Shenzhen University Town, Xili Town, Room 902, Shenzhen 518055, Guangdong Province, PR China.

E-mail address: zhou.jin@sz.tsinghua.edu.cn (J. Zhou).

<https://doi.org/10.1016/j.dib.2018.10.051>

2352-3409/© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

in the National Center Biotechnology Information (NCBI) BioProject with accession number PRJNA485554.

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications table

Subject area	Biology
More specific subject area	Bacteriology, Genomics, Ecology
Type of data	Figures, Tables
How data was acquired	The whole genome was sequenced with an Illumina Hi-Seq. 2500
Data format	Analyzed
Experimental factors	Isolation and characterization of native strains <i>S. rhizophila</i> GA1. Genomic DNA, extraction and sequencing procedures.
Experimental features	Genome of the <i>S. rhizophila</i> GA1 was sequenced and assembled.
Data source location	The strain was isolated from the dinoflagellate (<i>Gymnodinium aeryinosum</i>) on the Shenzhen coast of China (22°59'42.19"N, 114°54'74.01"E).
Data accessibility	Data is with this article. Also, the whole-genome of <i>S. rhizophila</i> GA1 has been deposited in the GenBank database under accession numbers PRJNA485554 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA485554).
Related research article	None

Value of the data

- It is the first whole genome of algicidal bacterium *S. rhizophila* GA1.
- This data allows other researchers to extend the study about the algae-bacterium interactions.
- The data could be help us developing microbial methods to control harmful algae.

1. Data

In this study, an algicidal bacterium (*Stenotrophomonas rhizophila* GA1) was screened from the marine dinoflagellate. The morphology, optimal growth conditions and algicidal ability are shown in Fig. 1A–C. A summary of other data for the isolated strain is listed in Table 1.

The whole genome of *S. rhizophila* GA1 contained 6,598,546 bases and a G+C content of 62.39% (Table 1). The analyses of the complete genome identified 4579 open reading frames. Homologous comparison by BLAST found 3395 CDS (coding sequence) involving 25 functional COGs (clusters of orthologous groups) and a part of the CDS involving 34 KEGG (Kyoto encyclopedia of genes and genomes) metabolic pathways (Fig. 2A).

Based on the functional categories of COGs and KEGG groups, 271 genes were involved in carbohydrate metabolism and 879 genes participated in amino acid transport and energy conversion (Fig. 2B). Several genes encoding putative AHL inducers (LuxI-like proteins) were found in contig 4 of *S. rhizophila* GA1. The thin-layer chromatography (TLC) confirmed *S. rhizophila* GA1 strain has the ability to produce AHL signaling molecules (Fig. 1D). In addition, a chitinase-coding gene (*chi*) was found downstream of luxR; the protein encoded by this gene is believed to contribute to the ability of *S. rhizophila* GA1 to lyse its host (algae).

2. Experimental design, materials, and methods

2.1. Isolation

The seawater sample was collected from the dinoflagellate (*G. aerucyinosum*) bloom on the Shenzhen coast of China. An isolated strain GA1 caused significant algal lyse. The 16S rRNA gene sequence analysis revealed that it shared 99.7% similarity to the type strain of *Stenotrophomonas rhizophila*. We provisionally named this strain *S. rhizophila* GA1.

2.2. DNA extraction, sequencing, and assembly

Genomic DNA of *S. rhizophila* GA1 was extracted using the genomic DNA extraction kit (MoBio, CA, USA) following the protocols of the manufacturer. Whole-genome sequencing of the normalized DNA was performed using IlluminaHiSeq. 2500 (San Diego, USA) instrument, as described in Glushchenko et al. [1]. De novo assembly was performed using CLC Genomics Workbench version 5.1 (CLC Bio, Denmark) and trimmed using a minimum Phred quality score of 20, a minimum length of 50 bp, allowing no ambiguous nucleotides and trimming off some low-quality nucleotides [2]. The reads were assembled with SOAPdenovo (V.2.04) [3], and the sequence was annotated using the RAST

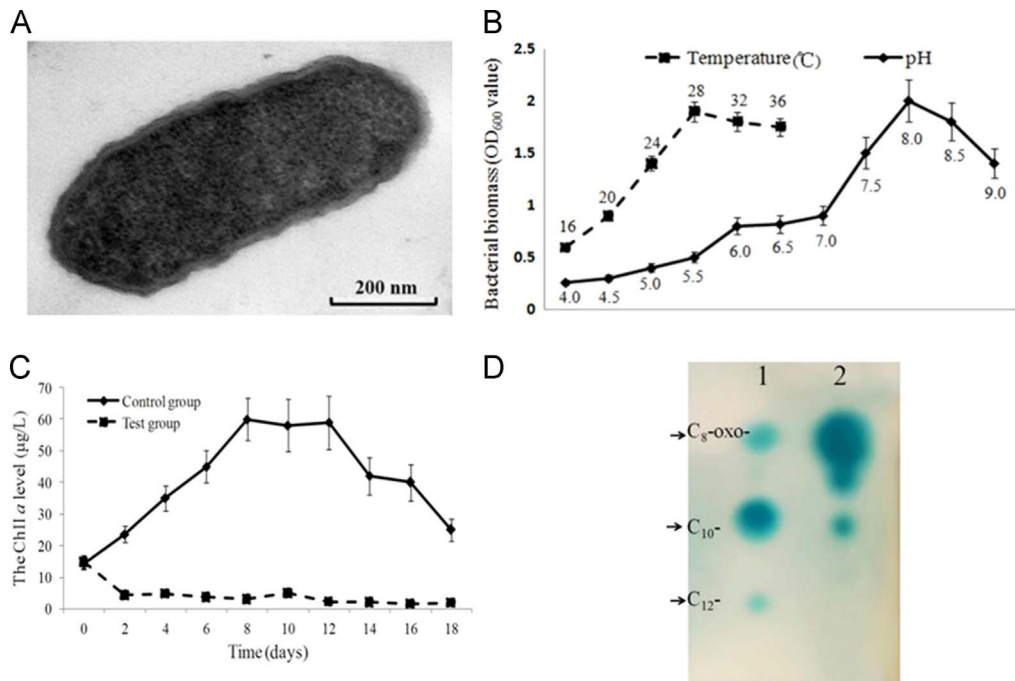


Fig. 1. (A) Electron micrographs of cells of *S. rhizophila* GA1. Preparation and electron microscopy conditions were as described by Hahnke et al. [10]. Magnification: 50,000 ×. (B) Optimal temperature and pH conditions for the growth of *S. rhizophila* GA1. Error bars indicate the mean value ± standard deviation of three optical density measurements at 600 nm (OD₆₀₀). (C) Algicidal activity of *S. rhizophila* GA1. The control group was sole-cultured *G. aerucyinosum*, and the initial concentration was 2 × 10⁴ cells/L. The test group was *G. aerucyinosum* co-cultured with *S. rhizophila* GA1 (final concentration was 1 × 10⁵ cells/mL). The total experimental cycle was 18 days. Error bars indicate the mean value ± standard deviation of three measurements of Chl *a*. (D) Analysis of AHLs from supernatant extracts of GA1 strain. AHLs extracted from cell-free culture supernatants were separated by thin-layer chromatography and detected using an overlay of agar seeded with *Agrobacterium tumefaciens* 136. Lane 1 contains AHL standards (arrows point to C₈-oxo-, C₁₀-, and C₁₂-AHL, respectively); lane 2 contains the *S. rhizophila* GA1 extracts.

Table 1
Basic information and genome features for *S. rhizophila* GA1.

Items	Descriptions
Geographic location	The coastal of Shenzhen, China
Latitude and longitude	N22°59'42.19", E114°54'74.01"
Organism/strain	<i>Stenotrophomonas rhizophila</i> GA1
Gram strain	Negative
Cell shape	Rod
Color of colonies	Yellow
Temperature	16–36 °C
Optimal pH	8.0
Environment (biome)	Temperature, salinity, pH value, and sea biome
Environment (feature)	Water body of phycosphere (<i>G. aerucinosum</i>)
Environment (Material)	Sea water
Sequencer	Illumina HiSeq. 2500
Data format	Processed
Experimental factor	Microbial strain
Experimental features	Whole genome sequence of <i>S. rhizophila</i> GA1
Consent	N/A
Assembly and annotation	CLC Genomics Workbench Version. 5.1
Finishing strategy	Primer design, PCR and sequencing
Genome size	5.92 Mb
GC content %	62.39%
Number of Contigs	67
Total contig size	5,929,188
Scaffolds	8
Total scaffold size	6,598,546
Protein encoding genes	4579
tRNAs	64
rRNAs	26
Predicted AHL gene <i>LuxI</i> site	Contig 4
Encoding-AHL gene length	459–848 bp

annotation server [4]. tRNA and rRNA genes were predicted by tRNAscan-SE [5] and RNAmmer [6], respectively. Genes were predicted using Glimmer 3.02 [7] and annotated by searching against the NCBI-nr and KEGG databases.

2.3. Thin-layer chromatography (TLC)

To verify the AHL-synthesizing activity of *S. rhizophila* GA1, reverse-phase thin-layer chromatography (TLC) was performed as described by Shaw et al. [8] and Ma et al. [9].

2.4. Data accessibility

The genome sequence data has been deposited in the GenBank database under accession numbers PRJNA485554.

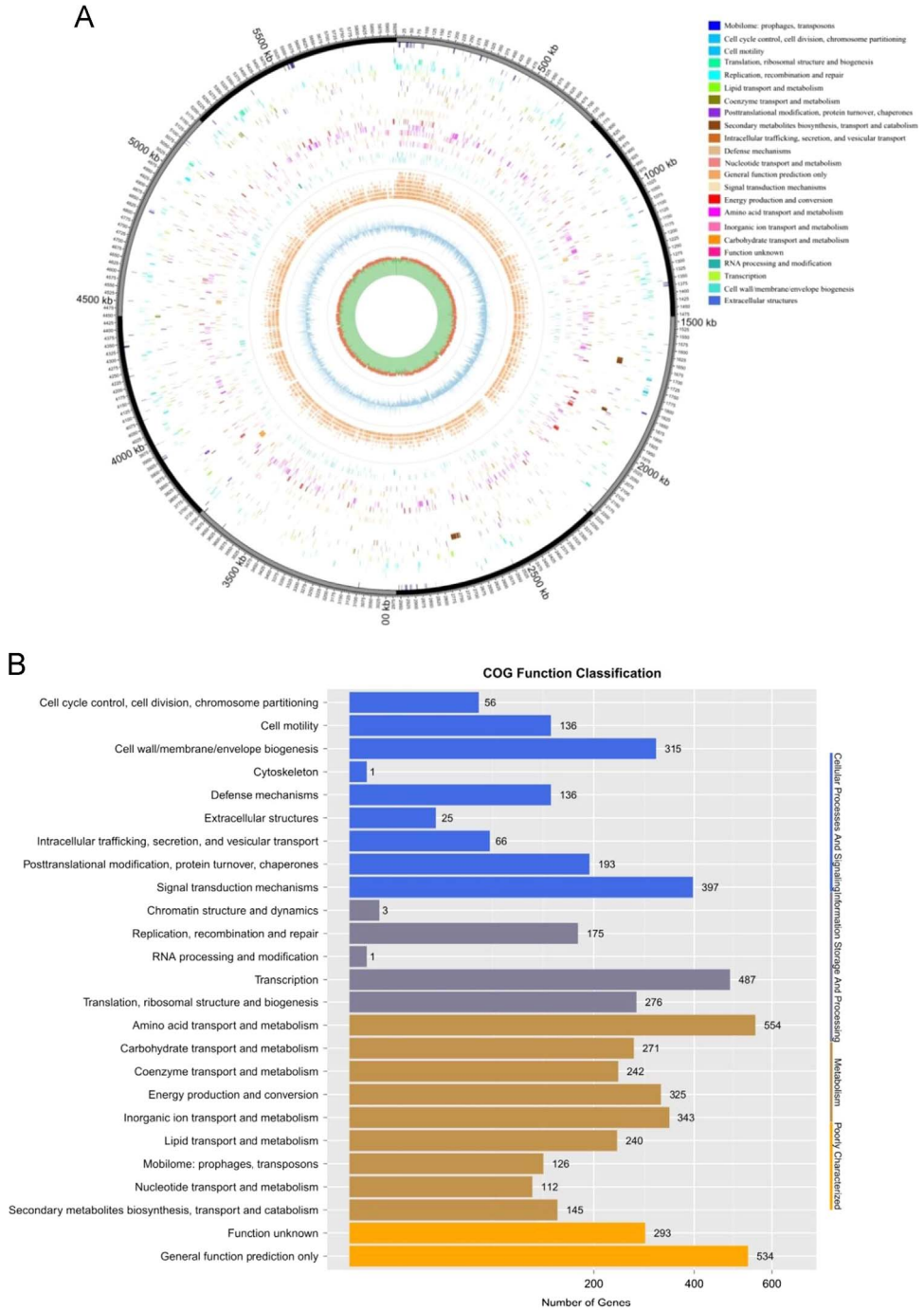


Fig. 2. (A) Circular map for the whole genome of *S. rhizophila* GA1. From the outside to the center: encoding genes, predicted CDSs transcribed in the clockwise (or counter clockwise) direction, ncRNA, GC percent (%), and GC skew (G + C/G-C) in a 1000-bp window. (B) Functional category distribution of *S. rhizophila* GA1 (based on COG function statistics).

Acknowledgments

This work was supported by Marine Fishery Science and Technology and Industry Development of Guangdong Province (A201603D05), and the S&T Projects of Shenzhen Science and Technology Innovation Committee (JCYJ20170412171959157, JCYJ20170817160708491, and CYZZ20170331112457200).

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at <https://doi.org/10.1016/j.dib.2018.10.051>.

References

- [1] O.E. Glushchenko, A.E. Samoilov, E.I. Olekhovich, B.A. Kovarsky, A.V. Tyakht, A.V. Pavlenko, et al., Data on gut meta-genomes of the patients with *Helicobacter pylori* infection before and after the antibiotic therapy, *Data Brief* 11 (2017) 68–71.
- [2] A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, *J. Comput. Biol.* 19 (2012) 455–477.
- [3] R. Luo, B. Liu, Y. Xie, Z. Li, W. Huang, J. Yuan, et al., SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler, *Gigascience* 1 (2012) 18.
- [4] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, et al., The RAST server: rapid annotations using sub-systems technology, *BMC Genom.* 9 (2008) 75.
- [5] T.M. Lowe, S.R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, *Nucleic Acids Res.* 25 (1997) 955–964.
- [6] K. Lagesen, P. Hallin, E.A. Rødland, H.H. Staerfeldt, T. Rognes, D.W. Ussery, RNAmmer: consistent and rapid annotation of ribosomal RNA genes, *Nucleic Acids Res.* 35 (2007) 3100–3108.
- [7] A.L. Delcher, K.A. Bratke, E.C. Powers, S.L. Salzberg, Identifying bacterial genes and endosymbiont DNA with Glimmer, *Bioinformatics* 23 (2007) 673–679.
- [8] P.D. Shaw, G. Ping, S.L. Daly, C. Cha, J.E. Cronan Jr., K.L. Rinehart, et al., Detecting and characterizing N-acyl-homoserine lactone signal molecules by thin-layer chromatography, *Proc. Natl. Acad. Sci. USA* 94 (1997) 6036–6041.
- [9] Z.P. Ma, Y.M. Lao, H. Jin, G.H. Lin, Z.H. Cai, J. Zhou, Diverse profiles of ai-1 type quorum sensing molecules in cultivable bacteria from the mangrove (*Kandelia obovata*) rhizosphere environment, *Front. Microbiol.* 7 (2016) 1957.
- [10] R.L. Hahnke, C.M. Bennke, B.M. Fuchs, A.J. Mann, E. Rhiel, H. Teeling, R. Amann, J. Harder, Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea, *Environ. Microbiol.* 17 (2015) 3515–3526.