Contents lists available at ScienceDirect

Genomics Data

Data in Brief

Draft genome sequence of *Citrobacter freundii* strain ST2, a γ-proteobacterium that produces N-acylhomoserine lactones

Yan Wang [a], Jin Zhou [b],*

[a] *Shenzhen Polytechnic, Shenzhen, Guangdong Province, PR China*
[b] *Ocean Science and Technology Division, Graduate School at Shenzhen, Tsinghua University, Shenzhen, PR China*

ARTICLE INFO

ABSTRACT

Citrobacter freundii strain ST2, isolated from the algae bloom sample, possesses an N-acylhomoserine lactone (AHL) production activity that secretes short-chain AHL molecules. In this study, we sequenced the complete genome of *C. freundii* strain ST2 to understand the molecular regulation of the AHL system and to search for the AHL gene in this bacterium. The results show that the genome size is 4.89 Mb with a G + C content of 51.96%. 4626 function proteins were predicted and 3647 proteins were assigned to COG functional categories. A predicted AHL-coding gene *LuxR* was found at contig 4 and the length was 1541 bp. The strain temporary deposited at Shenzhen Public Platform of Screening & Application of Marine Microbial Resources (Shenzhen, China), and the genome sequence can be accessed at GenBank under the accession no. LJSQ00000000.

Specification	
Organism	*Citrobacter freundii*
Strain	ST2
Sequencer	Illumina Hiseq 2000
Data format	Processed
Experimental factor	Microbial strain
Experimental features	Whole genome sequence of *Citrobacter freundii* ST2, assembly and annotation
Consent	N/A
Sample source location	Marine spring algae bloom, Shenzhen coast, China.

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/gquery/?term=SAMN04088747.

2. Experimental design, materials and methods

Quorum sensing (QS, e.g., AHL molecules), a common form of cell signaling, is gaining significant interest in marine ecology. Quorum sensing is a form of intercellular communication within microbial populations that occurs in a density-dependent manner. QS substance can result in downstream changes in gene regulation and modulation of bioluminescence, virulence, biofilm formation, and secondary metabolite production [1]. Chemical mediators of AHL and molecular crosstalk between bacteria and eukaryotes have been described for a wide range of symbiotic organisms [2]. In algal–bacterial microenvironments, previous works have investigated the crosstalk signaling between algae and bacteria, which can broaden understanding of the complex multi-species interactions [3–5]. In this work, we found the *Citrobacter freundii* strain ST2, which was isolated from the dinoflagellate (*Scrippsiella trochoidea*) and possessed AHL activities to produce N-acylhomoserine lactones. The genome sequence of *C. freundii* strain ST2 may provide insights on the regulation of quorum sensing mechanisms, and will help us understand the ecological behavior of this specie in algae–bacteria symbiosis.

For the above-mentioned reasons, the draft genome sequence of *C. freundii* strain ST2 was determined. Genomic DNA of this bacterium was isolated using a DNA kit (Mo Bio, CA, USA) according the instructions of the manufacturer. The quality of DNA was examined using a Nanodrop spectrophotometer (Thermo Scientific, USA) and gel electrophoresis (Bio-Rad, USA). The whole-genome sequencing of *C. freundii* ST2 was performed using the Illumina Hiseq 2000 platform after the construction of the sequencing library using a Truseq DNA sample preparation kit, V2 (Illumina Inc. CA). After filtering the low-quality data, we assembled the short reads into a genome sequence using SOAP denovo v1.05, and the scaffolds were manually connected based on paired-end relationships. The gaps were filled using GapCloser with read mapping information. This resulted in 4,313,757 filtered reads and approximately 50-fold coverage. The filtered reads were de novo assembled with the CLC Genomic workstation v5.1, generating 21 contigs, containing 4,896,351 bp. The G + C

* Corresponding author at: L-308, Ocean Science and Technology Division, Graduate School at Shenzhen, Tsinghua University, Shenzhen University Town, Xili Town, Shenzhen City, 518055, PR China.
E-mail address: zhou.jin@sz.tsinghua.edu.cn (J. Zhou).

content of this genome was 51.96%. Gene prediction was performed using prodigal (v260), and 4626 open reading frames were predicted [6]. Homologous comparisons by BLAST showed 3647 CDS involving the 21 functional COG groups and a part of the CDS involving the 34 metabolic pathway KEGG groups (Fig. 1). Annotation was performed using Blast2GO [7] against NCBI-NR sequence database [8]. One complete rRNA operon and one copy each of nine 5 s-rRNA gene, seven 16 s-rRNA gene, and six 23 s-rRNA gene were identified using RNAmmer [9]. A total of 79 tRNAs were identified using tRNAscan-SE [10]. In addition, 57 minisatellite DNAs and 3 microsatellite DNAs were determined based on the Repbase transposable elements library. The genome features of *C. freundii* strain ST2 were summarized in Table 1.

Based on Blast2GO analysis, the *LuxR* gene of *C. freundii* strain ST2 was found in contig 4, and the gene length was 1541 bp. It has a highly identity of the *LuxR* gene of another Enterobacteriaceae family member (*C. freundii* UCI 31, GenBank: EJF23932.1). Furthermore, a putative AI-2 (autoinducer-2) synthase gene with a metallo-beta-lactamase domain was also found in silico contig 4.

The complete genome of *C. freundii* strain ST2 provides novel insights into the quorum-sensing and signaling confusion genomic features of strain ST2. Further work needs to be conducted to study the roles of this bacterium in the phycosphere environment under AHL molecular regulation.

3. Nucleotide sequence accession number

The genome sequences were deposited in GenBank with the accession number LJSQ0000000. The version described in this paper is the first version.

Table 1
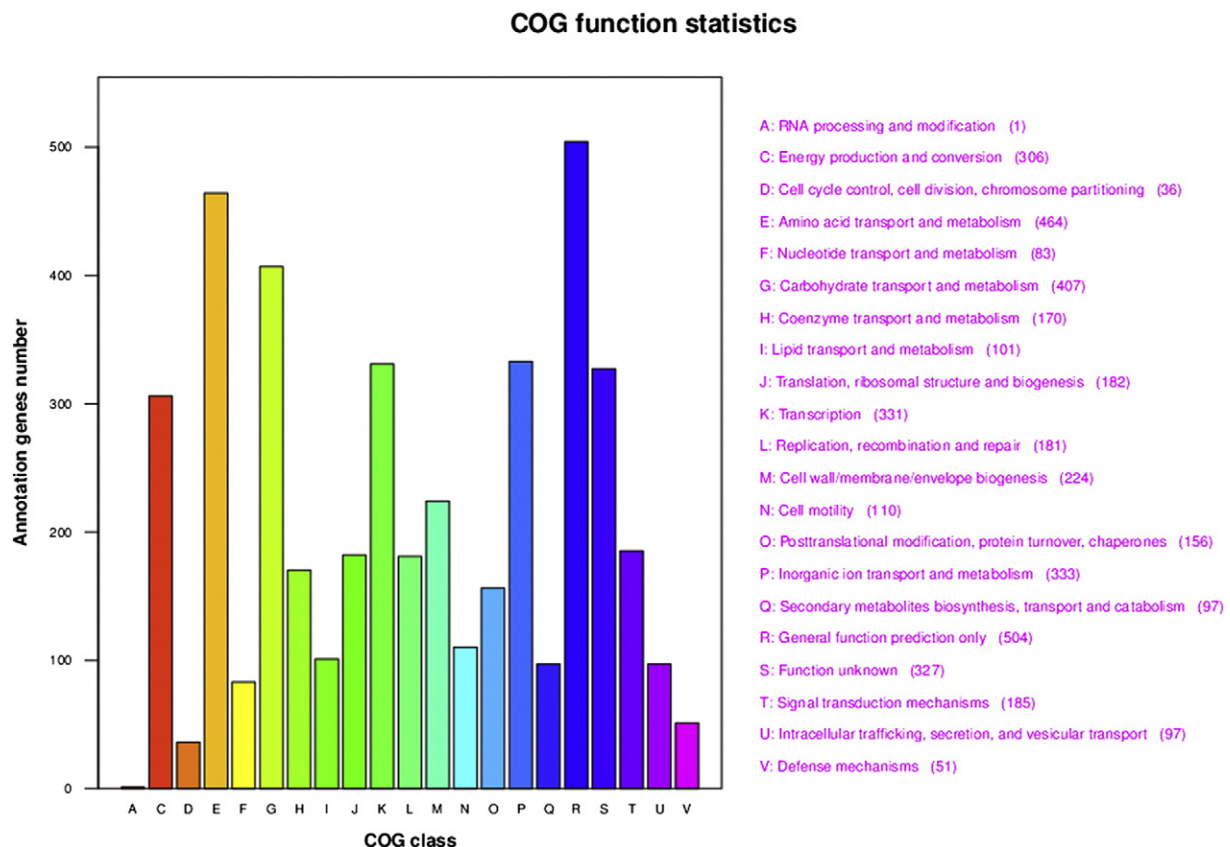Genome features of *Citrobacter freundii* strain ST2.

Attributes	Values
Genome size	4.89 Mb
GC content %	51.96%
Number of Contigs	21
Total contig size	4,885,161
Scafflods	6
Total scaffold size	4,896,351
Protein encoding genes	4626
tRNAs	79
rRNAs	21
Predicted AHL gene *LuxR* site	Contig 4
Encoding-AHL gene length	1541 bp

Conflict of interest

The authors declare that there is no conflict of interests on work published in this paper.

COG function statistics

A: RNA processing and modification   (1)
C: Energy production and conversion   (306)
D: Cell cycle control, cell division, chromosome partitioning   (36)
E: Amino acid transport and metabolism   (464)
F: Nucleotide transport and metabolism   (83)
G: Carbohydrate transport and metabolism   (407)
H: Coenzyme transport and metabolism   (170)
I: Lipid transport and metabolism   (101)
J: Translation, ribosomal structure and biogenesis   (182)
K: Transcription   (331)
L: Replication, recombination and repair   (181)
M: Cell wall/membrane/envelope biogenesis   (224)
N: Cell motility   (110)
O: Posttranslational modification, protein turnover, chaperones   (156)
P: Inorganic ion transport and metabolism   (333)
Q: Secondary metabolites biosynthesis, transport and catabolism   (97)
R: General function prediction only   (504)
S: Function unknown   (327)
T: Signal transduction mechanisms   (185)
U: Intracellular trafficking, secretion, and vesicular transport   (97)
V: Defense mechanisms   (51)

Fig. 1. Function category distribution of *Citrobacter freundii* strain ST2 strain (based on COG function statistics).

Y. Wang, J. Zhou / Genomics Data 6 (2015) 234–236

References

[1] A.R. Pacheco, V. Sperandio, Inter-kingdom signaling: chemical language between bacteria and host. Curr. Opin. Microbiol. 12 (2012) 192–198.

[2] N. Mangwani, H.R. Dash, A. Chauhan, S. Das, Bacterial quorum sensing: functional features and potential applications in biotechnology. J. Mol. Microbiol. Biotechnol. 22 (2012) 215–227.

[3] S.A. Amin, M.S. Parker, E.V. Armbrust, Interactions between diatoms and bacteria. Microbiol. Mol. Biol. Rev. 76 (2012) 667–684.

[4] S.A. Amin, L.R. Hmelo, H.M. van Tol, B.P. Durham, L.T. Carlson, K.R. Heal, R.L. Morales, C.T. Berthiaume, M.S. Parker, B. Djunaedi, A.E. Ingalls, M.R. Parsek, M.A. Moran, E.V. Armbrust, Interaction and signaling between a cosmopolitan phytoplankton and associated bacteria. Nature 7554 (2015) 98–101.

[5] P.R. Jonsson, H. Pavia, G. Toth, Formation of harmful algal blooms cannot be explained by allelopathic interactions. Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 11177–11182.

[6] D. Hyatt, G.L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer, L.J. Hauser, Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. 11 (2010) 119.

[7] A. Conesa, S. Götz, J.M. García-Gómez, J. Terol, M. Talón, M. Robles, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21 (18) (2005) 3674–3676.

[8] D.A. Benson, I. Karsch-Mizrachi, K. Clark, D.J. Lipman, J. Ostell, E.W. Sayers, GenBank. Nucleic Acids Res. 40 (2012) D48–D53 (Database issue).

[9] 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.

[10] 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.