Genomics Data 6 (2015) 267–268

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

Draft genome sequence of extremely acidophilic bacterium *Acidithiobacillus ferrooxidans* DLC-5 isolated from acid mine drainage in Northeast China

Peng Chen [a,*], Lei Yan [b], Zhengrong Wu [a], Ruixiang Xu [a], Suyue Li [c], Ningbo Wang [c], Ning Liang [c], Hongyu Li [a,*]

[a] School of Pharmacy, Lanzhou University, Donggang West Road No. 199, Lanzhou 730020, PR China
[b] College of Life Science and Technology, Heilongjiang Bayi Agricultural University, Daqing 163319, PR China
[c] Gansu Institute of Business and Technology, Yannan Road No. 449, Lanzhou 730010, PR China

ARTICLE INFO

Article history:
Received 21 October 2015
Accepted 21 October 2015
Available online 23 October 2015

Keywords:
Acidophilic bacteria
Acidithiobacillus ferrooxidans
Extremophiles
Genome

ABSTRACT

Acidithiobacillus ferrooxidans type strain DLC-5, isolated from Wudalianchi in Heihe of Heilongjiang Province, China. Here, we present the draft genome of strain DLC-5 which contains 4,232,149 bp in 2745 contigs with 57.628% GC content and includes 32,719 protein-coding genes and 64 tRNA-encoding genes. The genome sequence can be accessed at DDBJ/EMBL/GenBank under the accession no. JNNH00000000.1.

Specifications	
Organism/cell line/tissue	*Acidithiobacillus ferrooxidans*
Strain (s)	DLC-5
Sequencer or array type	Illumina Hiseq2000
Data format	Processed
Experimental factors	Microbial strains
Experimental features	Draft genome sequence of *Acidithiobacillus ferrooxidans* DLC-5 assembly and annotation
Consent	N/A
Sample source location	Acid mine drainage in the Wudalianchi in Heihe of Heilongjiang Province, China

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/bioproject/?term=JNNH00000000.

2. Experimental design, materials and methods

Acidithiobacillus ferrooxidans (*A. ferrooxidans*) is a Gram-negative, extremely acidophilic, mesophilic, chemolithotrophic bacterium and the most well-studied acidophilic organism which is usually found in acid environments such as acid mine drainage [1,2]. Due to its bioleaching capabilities, it is an important member of microbial consortia involved in the industrial recovery of metal under mesophilic conditions (bioleaching or biomining). Recently, *A. ferrooxidans* has played important roles in bioleaching and harnesse environmental contamination [3,4]. Like in other acidophilic iron-oxidizing bacterium, it grows optimally at about 35 °C in 9K inorganic medium at extremely low pH (pH 1.0–2.0) and fixes both carbon and nitrogen from the atmosphere [5]. *A. ferrooxidans* derives energy from oxidizing reduced sulfur compounds and Fe^{2+} ions to form sulfate and Fe^{3+}, respectively [3].

A. ferrooxidans strain, DLC-5 was grown in 9K medium at 35 °C. DNA was isolated from 1.0–1.5 g of cell paste using Qiagen Genomic 500 DNA Kit (Qiagen, Hil-den, Germany) with a modified protocol, st/FT, for cell lysis, as described in Valdes et al. [6]. Draft genome sequence of *A. ferrooxidans* type strain DLC-5 was obtained in Illumina Hiseq2000 sequencing technology by Shanghai Majorbio Bio-pharm Technology Co., Ltd. (Shanghai, China), using the Short Oligonucleotides Alignment Program (SOAP) denovo alignment tool (http://soap.genomics.org.cn/) processes reads assemble. A library containing 300-bp inserts was constructed. Altogether, 6,372,268 paired reads; 398,580 single reads; total 1,079,535,272 bp bases with average coverage of 221.1×. Reads were filtered to remove adapter sequences, low-quality bases (Phred score, <20), removing the 5′ end that contains the bases of it is not A, G, C, T before shearing, remove reads with the containing 10% of N, giving up adapter and small fragments of length less than 25 bp after qualitative pruning. The reads were assembled into 881 contigs (>1000 bp; Contig

* Corresponding authors.
E-mail addresses: chenpeng@lzu.edu.cn (P. Chen), lihy@lzu.edu.cn (H. Li).

http://dx.doi.org/10.1016/j.gdata.2015.10.018

N50, 102 bp; Contig N90, 569 bp) and 573 scaffolds (>1000 bp; Scaffold N50, 71 bp; Scaffold N90, 333 bp).

Until now, two genome sequences of *A. ferrooxidans* strains ATCC 23270 and ATCC 53993 are available in the public databases [7,8]. These genomic data are useful for the experimental identification of unique proteins or estimation of the phylogenetic relationship among the related strains. Strain DLC-5 (CCTCC-M 2014362) is the type strain of *A. ferrooxidans*, isolated from Wudalianchi in Heihe of Heilongjiang Province, and the type species of the genus *Acidithiobacillus*, which currently contains five species. The draft genome sequences of strain DLC-5 have great importance to provide more explicit information for the physiology and metabolic potential of *A. ferrooxidans*. Further analysis of the genome sequence via gene engineering might improve the oxidation of Fe^{2+} efficiency by strain DLC-5. The genome includes two plasmids, for a total size of 3,142,890 bp, with one circular chromosome of 1,832,305 bp (58.3% GC content). For the main chromosome, 4299 bp genes were predicted, 4131 bp of which are protein-coding genes. 3250 bp of protein coding genes were assigned to a putative function with the remaining annotated as hypothetical proteins. The properties and the statistics of the genome are summarized in Table 1. The distribution of genes into COGs functional categories is presented in Table 2.

Extremely acidophilic bacteria and archaea with special emphasis on bioleaching microorganisms are widely distributed in the extreme acidic environment. In this study, we analyzed the genome sequence of *A. ferrooxidans* DLC-5, which was isolated from acid mine drainage in Northeast China. Genome analysis of this strain revealed the presence of key functional characteristics. It may contribute to further studies on important process for bioleaching and acid mine drainage production, such as biofilm formation, energy resources utilization and quorum sensing that could play a role in a possible interrelationship of bioleaching heaps and other acidic environments. In addition, combining with genomes of other members in *Acidithiobacillus*, will make an important advance in understanding of the ecological roles that *Acidithiobacillus* species play in those acidic environments and their relationships with other extremely acidophilic microorganisms.

3. Nucleotide sequence accession number

The sequence of *A. ferrooxidans* DLC-5 under this Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JNNH00000000.1. The version described in this paper is version JNNH00000000.1.

Competing interests

The authors declare that there are no competing interests.

Table 1
Genome statistics.

Attribute	Value	% of total
Genome size (bp)	3,142,890	100.0
DNA coding (bp)	2,816,029	89.6
DNA G + C (bp)	1,832,305	57.63
DNA scaffolds	1333	31.0
Total genes	4299	100.0
Protein coding genes	4131	96.1
RNA genes	168	3.9
Pseudo genes	0	
Genes in internal clusters	0	
Genes with function prediction	3312	77.0
Genes assigned to COGs	3250	75.6
Genes with Pfam domains	3486	81.1
Genes with signal peptides	315	7.3
Genes with transmembrane helices	826	19.2
CRISPR repeats	0	

Table 2
Number of genes associated with general COG functional categories.

Code	Value	% age	Description
J	118	5.6	Translation, ribosomal structure and biogenesis
A	1	0.0	RNA processing and modification
K	112	5.4	Transcription
L	157	7.5	Replication, recombination and repair
B	1	0.0	Chromatin structure and dynamics
D	29	1.4	Cell cycle control, cell division, chromosome partitioning
V	51	2.4	Defense mechanisms
T	59	2.8	Signal transduction mechanisms
M	141	6.7	Cell wall/membrane biogenesis
N	30	1.4	Cell motility
U	76	3.6	Intracellular trafficking and secretion
O	96	4.6	Posttranslational modification, protein turnover, chaperones
C	159	7.6	Energy production and conversion
G	93	4.4	Carbohydrate transport and metabolism
E	146	7.0	Amino acid transport and metabolism
F	41	2.0	Nucleotide transport and metabolism
H	89	4.3	Coenzyme transport and metabolism
I	59	2.8	Lipid transport and metabolism
P	125	6.0	Inorganic ion transport and metabolism
Q	43	2.1	Secondary metabolites biosynthesis, transport and catabolism
R	212	10.1	General function prediction only
S	134	6.4	Function unknown
–	119	5.7	Not in COGs

The total is based on the total number of protein coding genes in the genome.

Acknowledgments

This work was supported by Gansu Province Science Foundation for Distinguished Young Scholars (Grant No. 1308RJDA014), Longyuan Support Project for Young Creative Talents (Grant No. GANZUTONGZI [2014] no.4), Technology Program of Lanzhou City (Grant No. 2013-4-115), and The Fundamental Research Funds for the Central Universities of China (Grant No. lzujbky-2015-57).

References

[1] D.P. Kelly, A.P. Wood, Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halothiobacillus* gen. nov. and *Thermithiobacillus* gen. nov. Int. J. Syst. Evol. Microbiol. 50 (2000) 511–516.
[2] V.B.D. Skerman, V. McGowan, P.H.A. Sneath, Approved lists of bacterial names. Int. J. Syst. Bacteriol. 30 (1980) 225–230.
[3] P. Chen, et al., Bioleaching of realgar by *Acidithiobacillus ferrooxidans* using ferrous iron and elemental sulfur as the sole and mixed energy sources. Bioresour. Technol. 102 (2011) 3260–3267.
[4] Y. Zhao, et al., The use of (5Z)-4-bromo-5-(bromomethylene)-2(5H)-furanone for controlling acid mine drainage through the inhibition of *Acidithiobacillus ferrooxidans* biofilm formation. Bioresour. Technol. 186 (2015) 52–57.
[5] R. Quatrini, E. Jedlicki, D.S. Holmes, Genomic insights into the iron uptake mechanisms of the biomining microorganism *Acidithiobacillus ferrooxidans*. J. Ind. Microbiol. Biotechnol. 32 (2005) 606–614.
[6] J. Valdes, et al., Draft genome sequence of the extremely acidophilic bacterium *Acidithiobacillus caldus* ATCC 51756 reveals metabolic versatility in the genus *Acidithiobacillus*. J. Bacteriol. 191 (2009) 5877.
[7] J. Valdes, et al., *Acidithiobacillus ferrooxidans* metabolism: from genome sequence to industrial applications. BMC Genomics 9 (2008) 597.
[8] L.H. Orellana, C.A. Jerez, A genomic island provides *Acidithiobacillus ferrooxidans* ATCC 53993 additional copper resistance: a possible competitive advantage. Appl. Microbiol. Biotechnol. 92 (2011) 761–767.