



## Data in Brief

## Draft genome of iron-oxidizing bacterium *Leptospirillum* sp. YQP-1 isolated from a volcanic lake in the Wudalianchi volcano, China

Lei Yan <sup>a,\*</sup>, Shuang Zhang <sup>a</sup>, Gaobo Yu <sup>a</sup>, Yongqing Ni <sup>b</sup>, Weidong Wang <sup>a</sup>, Huixin Hu <sup>a</sup>, Peng Chen <sup>c</sup>

<sup>a</sup> College of Life Science and Technology, Heilongjiang Bayi Agricultural University, Xinfeng Road No.5, Daqing 163319, PR China

<sup>b</sup> School of Food Sciences, Shihezi University, North 4th Road No. 280, Shihezi 832003, PR China

<sup>c</sup> School of Pharmacy, Lanzhou University, Tianshui Road No. 222, Lanzhou 730000, PR China

## ARTICLE INFO

## Article history:

Received 29 August 2015

Accepted 2 September 2015

Available online 9 September 2015

## Keywords:

*Leptospirillum*

Genome

Iron-oxidizing bacteria

Volcanic lake

## ABSTRACT

*Leptospirillum* sp. YQP-1, a member of iron-oxidizing bacteria was isolated from volcanic lake in northeast China. Here, we report the draft genome sequence of the strain YQP-1 with a total genome size of 3,103,789 bp from 85 scaffolds (104 contigs) with 58.64% G + C content. The genome sequence can be accessed at DDBJ/EMBL/GenBank under the accession no. LIEB00000000.

© 2015 Elsevier Inc. All rights reserved. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Specifications

Organism/cell line/tissue	<i>Leptospirillum</i> sp.
Strain (s)	YQP-1
Sequencer or array type	Sequencer; Illumina MiSeq
Data format	Processed
Experimental factors	Microbial strains
Experimental features	Draft genome sequence of <i>Leptospirillum</i> sp. YQP-1 assembly and annotation
Consent	N/A
Sample source location	Volcanic lake in the Wudalianchi volcano, China

## 1. Direct link to deposited data

<http://www.ncbi.nlm.nih.gov/bioproject/?term=LIEB00000000>.

\* Corresponding author.

E-mail address: [hekouyanlei@gmail.com](mailto:hekouyanlei@gmail.com) (L. Yan).

## 2. Experimental design, materials and methods

The *Leptospirillum* genus was validly described by Hippe [1], though the first bacterium of this genus was identified by Markosyan [2]. Phylogenetically, it belongs to the class Nitrospira. *Leptospirillum* spp. are found in volcano habitat and acid mine drainage environment [3, 4]. Cells of the *Leptospirillum* genus are Gram-negative, motile vibroid to spirilla-shaped, but cocci or pseudococci also can be formed [3]. This genus comprises four species, viz., *Leptospirillum ferrooxidans*, *Leptospirillum ferriphilum*, *Leptospirillum ferrodiazotrophum*, and *Leptospirillum thermoferrooxidans* [3,5]. Until recently, only four genomes are available in the GenBank, including complete genomes of *L. ferriphilum* ML-04 [6], *L. ferriphilum* YSK (CP007243), and *L. ferrooxidans* C2-3 [4], as well as draft genome of *L. ferriphilum* DSM 14647 [7]. As an important iron-oxidizing genus, more investigations at genomic level are required to improve understanding of its physiology, ecology, genetics and potential in biotechnological applications.

*Leptospirillum* sp. YQP-1 was isolated from a volcanic lake of the Wudalianchi volcano, northeast China in July 2013. This strain has been deposited in the China Center for Type Culture Collection (accession number: CCTCC M 2014196) (<http://www.cctcc.org/index.php>). The 16S rRNA gene sequence of YQP-1 available in GenBank database (accession number: KJ573504) showed more than 99% identity with several *Leptospirillum* strains, such as *Leptospirillum* sp. 3.2 (accession number: EU372651), *L. ferrooxidans* P3a (accession number: AF356837). This indicated that this strain belongs to *Leptospirillum* genus.

**Table 1**  
Genome features of *Leptospirillum* sp. YQP-1 draft genome.

Attributes	Values
Genome size (bp)	3,103,789
Number of scaffolds	85
Scaffold N50 (bp)	101,596
G + C content (%)	58.64
Number of contigs	104
Contig N50 (bp)	86,194
CDS	3038
rRNAs	8
tRNAs	42
Pseudo genes	55
Tandem repeat sequence	106

Genomic DNA was isolated from *Leptospirillum* sp. YQP-1 using a Bacterial Genomic DNA Extraction Kit (Biomed, Beijing, China) according to the manufacturer's instructions. Quality check was performed using NanoDrop spectrophotometer (Thermo Scientific, Waltham, USA). The draft genome sequence of the YQP-1 strain was obtained by paired-end sequencing on the Illumina MiSeq platform at the Biomarker Technologies Co., Ltd. (Beijing, China). De novo assembly was performed using Velvet Version 1.2.10 [8]. Protein-coding sequences were predicted by Glimmer software version 3.0 [9], while ribosomal RNA (rRNA) and transfer RNA (tRNA) genes were predicted using an RNAmmer 1.2 server [10] and tRNAscan-SE Search Server version 1.21 [11], respectively. Genes annotated using BLAST searches of nonredundant protein sequences from the NCBI, Swiss-Prot, COG [12], and KEGG database [13]. Tandem repeats were predicted using Tandem Repeats Finder Version 4.04.

The genome size of *Leptospirillum* sp. YQP-1 was estimated to be 3,103,789 bp from 85 scaffolds with a mean G + C content of 58.64%. It contains 104 contigs with N50 contig length of 86,194 bp. A total of 3038 coding sequences (CDS) were identified including 50 RNA genes (8 rRNA and 42 tRNA genes). The number of tandem repeat sequence was 106 with total length of 12,049 bp, which made up 0.3882% of genome (Table 1). Further studies of *Leptospirillum* sp. YQP-1 which might help many beneficial applications such as for bioremediation and bioleaching, are essential and required to be extended.

### 3. Nucleotide sequence accession number

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LIEB00000000. The version described in this paper is LIEB00000000.

### Conflict of interest

The authors declare that there is no conflict of interests with respect to the work published in this paper.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (41471201, 31100006), Natural Science Foundation of Heilongjiang Province of China (C201442), and Program of Science and Technology Innovation Teams Building in Heilongjiang Province (2012TD006).

### References

- [1] H. Hippe, *Leptospirillum* gen. nov. (ex Markosyan 1972), nom. rev., including *Leptospirillum ferrooxidans* sp. nov. (ex Markosyan 1972), nom. rev. and *Leptospirillum thermoferrooxidans* sp. nov. (Golovacheva et al. 1992). *Int. J. Syst. Evol. Microbiol.* 50 (2000) 501–503.
- [2] G.E. Markosyan, A new iron-oxidizing bacterium: *Leptospirillum ferrooxidans* gen. nov. sp. *Biol. J. Armenia* 25 (1972) 26–29.
- [3] E.R. Donati, W. Sand, *Microbial Processing of Metal Sulfides*. Springer, 2007.
- [4] R. Fujimura, Y. Sato, T. Nishizawa, K. Oshima, S.-W. Kim, M. Hattori, T. Kamijo, H. Ohta, Complete genome sequence of *Leptospirillum ferrooxidans* strain C2-3, isolated from a fresh volcanic ash deposit on the island of Miyake, Japan. *J. Bacteriol.* 194 (2012) 4122–4123.
- [5] D.S.A. Goltsman, M. Dasari, B.C. Thomas, M.B. Shah, N.C. VerBerkmoes, R.L. Hettich, J.F. Banfield, New group in the *Leptospirillum* clade: cultivation-independent community genomics, proteomics, and transcriptomics of the new species "*Leptospirillum* group IV UBA BS". *Appl. Environ. Microbiol.* 79 (2013) 5384–5393.
- [6] S. Mi, J. Song, J. Lin, Y. Che, H. Zheng, J. Lin, Complete genome of *Leptospirillum ferriphilum* ML-04 provides insight into its physiology and environmental adaptation. *J. Microbiol.* 49 (2011) 890–901.
- [7] J.P. Cárdenas, M. Lazzano, F.J. Ossandon, M. Corbett, D.S. Holmes, E. Watkin, Draft genome sequence of the iron-oxidizing acidophile *Leptospirillum ferriphilum* type strain DSM 14647. *Genome Announc.* 2 (2014) (e01153-01114).
- [8] D.R. Zerbino, E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18 (2008) 821–829.
- [9] 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.
- [10] K. Lagesen, P. Hallin, E.A. Rødland, H.-H. Stærfeldt, T. Rognes, D.W. Ussery, RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35 (2007) 3100–3108.
- [11] P. Schattner, A.N. Brooks, T.M. Lowe, The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33 (2005) W686–W689.
- [12] R.L. Tatusov, N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B. Kiryutin, E.V. Koonin, D.M. Krylov, R. Mazumder, S.L. Mekhedov, A.N. Nikolskaya, The COG database: an updated version includes eukaryotes. *BMC Bioinforma.* 4 (2003) 41.
- [13] M. Kanehisa, S. Goto, S. Kawashima, Y. Okuno, M. Hattori, The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32 (2004) D277–D280.