



# Draft Genome Sequence of *Chlorobium* sp. Strain N1, a Marine Fe(II)-Oxidizing Green Sulfur Bacterium

 Casey Bryce,<sup>a</sup>  Nia Blackwell,<sup>a,b</sup>  Daniel Straub,<sup>b,c</sup>  Sara Kleindienst,<sup>b</sup>  Andreas Kappler<sup>a</sup>

<sup>a</sup>Geomicrobiology Group, Center for Applied Geoscience, University of Tübingen, Tübingen, Germany

<sup>b</sup>Microbial Ecology Group, Center for Applied Geoscience, University of Tübingen, Tübingen, Germany

<sup>c</sup>Quantitative Biology Center (QBIC), University of Tübingen, Tübingen, Germany

**ABSTRACT** Here, we present the draft genome sequence of the halotolerant photoferrotroph *Chlorobium* sp. strain N1. This draft genome provides insights into the genomic potential of the only marine Fe(II)-oxidizing green sulfur bacterium (GSB) available in culture and expands our views on the metabolic capabilities of Fe(II)-oxidizing GSB more generally.

Aerobic photoautotrophic Fe(II)-oxidizing bacteria, so-called “photoferrotrophs,” were key drivers in the marine iron cycle prior to the rise of oxygen (1). Numerous photoferrotrophs have been isolated (reviewed in reference 2) that are either purple sulfur/nonsulfur bacteria or green sulfur bacteria (GSB), but most are from freshwater environments. Here, we present the genome sequence of the halotolerant GSB *Chlorobium* sp. strain N1, which is the only Fe(II)-oxidizing GSB obtained from a marine environment (3) and thus provides a unique opportunity to study Fe(II)-oxidizing GSB at salinities comparable to those of the ancient oceans.

*Chlorobium* sp. strain N1 was obtained from sediments of Norsminde Fjord in Aarhus Bay, Denmark, and was maintained in artificial seawater medium with 10 mM FeCl<sub>2</sub> (the physiology of *Chlorobium* sp. N1 is described in reference 3). Genomic DNA was extracted using the PowerSoil DNA extraction kit, according to the manufacturer’s instructions, and stored at –80°C. Three hundred nanograms of high-molecular-weight DNA was sheared to a target fragment size of 550 bp using a Covaris S2 focused ultrasonicator, and genomic sequencing libraries were prepared using the NEBNext Ultra II DNA library preparation for Illumina sequencing kit (New England BioLabs), according to the manufacturer’s protocol. Sequencing was conducted by IMGM Laboratories GmbH (Munich, Germany) with the Illumina MiSeq kit version 2. A total of 3,003,794 read pairs with a 250-bp read length were produced. Reads were trimmed and adapters removed using Trimmomatic version 0.36 (4), with the following parameters: ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:1:true, LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, and MINLEN:50. The quality was checked with FastQC version 0.11.5 before and after trimming, using the default parameters (5).

PhiX contamination was tested with Bowtie version 2.3.3 by aligning trimmed paired-end reads to the PhiX genome (Illumina iGenome). Forty-five percent of the paired-end reads were successfully merged with FLASH version 1.2.11 (6). Merged and unmerged paired-end reads were subsequently assembled with SPAdes version 3.11.0 (7). Genome contamination of 66.20% was estimated by CheckM version 1.0.11 (8) with Prodigal version 2.6.2 (9), HMMER version 3.1b2 (<http://hmmerr.org/>), and pplacer version 1.1.alpha19 (10) based on the reduced reference genome tree using the current CheckM database. The GC content peaked at around 35% and 65%, with similar peak heights, making the presence of one or more other genomes likely. *In silico* genome purification with MaxBin version 2.2.4 (11) extracted 36 contigs with an  $N_{50}$  value of

**Citation** Bryce C, Blackwell N, Straub D, Kleindienst S, Kappler A. 2019. Draft genome sequence of *Chlorobium* sp. strain N1, a marine Fe(II)-oxidizing green sulfur bacterium. *Microbiol Resour Announc* 8:e00080-19. <https://doi.org/10.1128/MRA.00080-19>.

**Editor** Irene L. G. Newton, Indiana University, Bloomington

**Copyright** © 2019 Bryce et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Casey Bryce, [casey.bryce@uni-tuebingen.de](mailto:casey.bryce@uni-tuebingen.de).

**Received** 22 January 2019

**Accepted** 4 April 2019

**Published** 2 May 2019

265,932 bp. Using these purified contigs, CheckM calculated a genome completeness for *Chlorobium* sp. strain N1 of 98.91% and found no contamination. The contigs were uploaded to the Integrated Microbial Genomes Expert Review tool (<https://img.jgi.doe.gov/cgi-bin/submit/main.cgi>) for gene predictions and functional annotations.

The *Chlorobium* sp. strain N1 genome is 2,373,849 Mb, with 2,287 open reading frames (ORFs), of which 2,217 are protein-coding genes. One full-length 16S rRNA gene sequence was found, as well as a 16S rRNA gene fragment. At the 16S rRNA gene level, this strain is 99% similar to *Chlorobium luteolum* DSM 273, which has not yet been shown to oxidize Fe(II), but it has only 80% similarity at the genome level. This strain is 72% similar at the genome level to freshwater *Chlorobium ferrooxidans*, the first Fe(II)-oxidizing GSB identified.

This draft genome contains genes for the type I photosynthetic reaction center, CO<sub>2</sub> fixation via the reverse tricarboxylic acid (TCA) cycle, and sulfur metabolism pathways common to the *Chlorobi*, as well as for multiple fermentation pathways. A homolog to the Cyc2 iron oxidase of *Acidithiobacillus ferrooxidans* is also present. This draft genome not only provides the first insights into the genomic potential of the only marine Fe(II)-oxidizing GSB obtained in culture, but it also expands our views on the metabolic capabilities of the Fe(II)-oxidizing GSB more generally.

**Data availability.** This genome can be accessed at <http://img.jgi.doe.gov>, and associated metadata can be found in the GOLD database at <https://gold.jgi.doe.gov/> (GOLD analysis project identifier [ID] Ga0226664). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number [PRJEB30018](https://doi.org/10.1093/bioinformatics/btr507), including the assembly (accession number [SJPA0000000](https://doi.org/10.1093/bioinformatics/btr507)). The raw data have been deposited in the Sequence Read Archive under accession number [ERS2924416](https://doi.org/10.1093/bioinformatics/btr507).

## ACKNOWLEDGMENTS

We thank the European Union's Seventh Framework Programme (FP/2007–2013) (ERC grant, agreement number 307320 MICROFOX to A.K.) for supporting this project and for funding C.B. and N.B.

We also thank Katja Laufer for providing this strain and her previous physiological characterization.

## REFERENCES

- Camacho A, Walter XA, Picazo A, Zopfi J. 2017. Photoferrotrophy: remains of an ancient photosynthesis in modern environments. *Front Microbiol* 8:323. <https://doi.org/10.3389/fmicb.2017.00323>.
- Bryce C, Blackwell N, Schmidt C, Otte J, Huang Y, Kleindienst S, Tomaszewski E, Schad M, Warter V, Peng C, Byrne JM, Kappler A. 2018. Microbial anaerobic Fe(II) oxidation—ecology, mechanisms and environmental implications. *Environ Microbiol* 20:3462–3483. <https://doi.org/10.1111/1462-2920.14328>.
- Laufer K, Niemeyer A, Nikeleit V, Halama M, Byrne JM, Kappler A. 2017. Physiological characterization of a halotolerant anoxygenic phototrophic Fe(II)-oxidizing green-sulfur bacterium isolated from a marine sediment. *FEMS Microbiol Ecol* 93:fix054. <https://doi.org/10.1093/femsec/fix054>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27:2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
- Matsen FA, Kodner RB, Armbrust EV. 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11:538. <https://doi.org/10.1186/1471-2105-11-538>.
- Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32:605–607. <https://doi.org/10.1093/bioinformatics/btv638>.