

1 **Genomes of *Geoalkalibacter ferrihydriticus* Z-0531^T and *Geoalkalibacter***
2 ***subterraneus* Red1^T, two haloalkaliphilic metal-reducing Deltaproteobacteria**

3

4 Running title: Genomes of two *Geoalkalibacter* spp.

5

6 Jonathan P. Badalamenti¹, Rosa Krajmalnik-Brown², César I. Torres², and

7 Daniel R. Bond^{1*}

8

9 ¹ BioTechnology Institute

10 University of Minnesota – Twin Cities

11 140 Gortner Laboratory

12 1479 Gortner Avenue

13 Saint Paul, MN 55108

14

15 ² Biodesign Institute

16 Swette Center for Environmental Biotechnology

17 Arizona State University

18 Tempe, AZ 85287

19

20 * Corresponding author

21 Email: dbond@umn.edu

22 Phone: 612-624-8619

23

24 Keywords: *Geoalkalibacter*, haloalkaliphile, metal reduction, PacBio

25

26

27

28

29 **ABSTRACT**

30 We sequenced and annotated genomes of two haloalkaliphilic Deltaproteobacteria,
31 *Geoalkalibacter ferrihydriticus* Z-0531^T (DSM 17813) and *Geoalkalibacter subterraneus*
32 Red1^T (DSM 23483). During assembly, we discovered that the DSMZ stock culture of
33 *G. subterraneus* was contaminated. We re-isolated *G. subterraneus* in axenic culture and
34 re-deposited in DSMZ and JCM.

35

36

37

38

39

40

41

42 While complete genomes exist for several freshwater sediment and subsurface
43 isolates from the *Geobacteraceae* and *Desulfuromonadaceae* families of
44 Deltaproteobacteria (1), genomes from haloalkaliphilic members capable of metal
45 reduction and electrode respiration at elevated salinity or pH (2-5) are lacking. To
46 facilitate genetic studies of extracellular electron transfer in these haloalkaliphiles, we

47 assembled and annotated draft and finished genomes of *Geoalkalibacter ferrihydriticus*
48 Z-0531 and *Geoalkalibacter subterraneus* Red1, respectively.

49 For *G. ferrihydriticus*, we used the a5 pipeline (March 26, 2013 release) (6) to
50 assemble a 3,839,416-bp draft genome in 23 contigs at $\sim 100\times$ coverage ($N_{50} = 602,663$
51 bp; 57.95% GC) using 2×100 bp paired-end Illumina reads. More than 99% of
52 assembled sequence exists in 14 contigs, with the largest contig (1,394,064 bp)
53 comprising $>36\%$ of the assembly.

54 Applying the same approach to *G. subterraneus*, multiple occurrences of single-
55 copy genes and contigs with low G+C content indicated that the stock culture (5) was
56 contaminated with a Gram-positive organism related to *Tieserella* sp.. This finding was
57 subsequently confirmed by DSMZ (S. Spring, personal communication). We re-isolated
58 *G. subterraneus* on solid medium lacking yeast extract and performed PacBio long read
59 sequencing according to protocols for a 20-kb insert size. SMRT bell templates were
60 size-selected with a 4-kb cutoff using Blue Pippin electrophoresis (Sage Science).
61 Subread filtering from 4 SMRT cells (P4-C2 chemistry, 120-minute movies) yielded 1.13
62 Gbp of sequence (average read length = 4,760 bp; $N_{50} = 6,620$ bp). Assembly was
63 performed using HGAP v.2 (7) with default parameters in SMRT Analysis v. 2.1. One
64 400-Mbp SMRT cell was sufficient to assemble the *G. subterraneus* chromosome into
65 one linear contig, but long reads from four SMRT cells were required to resolve three
66 tandem copies of the rRNA operon (total length 16,597 bp) and circularize the contig.
67 The assembly was polished to $>99.999\%$ consensus concordance (QV 50) with three
68 successive passes through Quiver (7) at $250\times$ coverage, and remaining indels were
69 removed with Pilon v. 1.8 (8) using $100\times$ coverage of 2×100 bp paired-end Illumina

70 reads. The finished assembly consisted of one chromosome (3,475,523 bp) and one
71 megaplasmid (242,122 bp) totaling 3,717,645 bp (G+C content 56.68%).

72 Annotation via the NCBI Prokaryotic Genome Annotation Pipeline revealed
73 features consistent with other *Geobacteraceae* and *Desulfuromonadaceae* genomes,
74 including a complete TCA cycle with a eukaryotic-like citrate synthase (9) and an
75 abundance of putative histidine kinases (>40 in both genomes) and multiheme *c*-type
76 cytochromes for extracellular respiration (43 and 51 in *G. ferrihydriticus* and *G.*
77 *subterraneus*, respectively) (1). Notably, only *G. ferrihydriticus* contains a putative
78 *hgcAB* gene cluster indicative of mercury methylation (GFER_06575 and GFER_06580)
79 (10).

80 **Strain and nucleotide sequence accession numbers.** *G. subterraneus* has been
81 re-deposited in DSMZ and JCM. Sequences have been deposited in GenBank under
82 accession numbers JWJD00000000 (*G. ferrihydriticus*) and CP010311 and CP010312
83 (*G. subterraneus* chromosome and plasmid, respectively). Raw Illumina and PacBio
84 reads, as well as base modification data for *G. subterraneus*, have been deposited to the
85 NCBI Sequence Read Archive under accession numbers SRX808753 and SRX808316
86 for *G. ferrihydriticus* and *G. subterraneus*, respectively.

87 **Acknowledgments.** This work was supported by the Minnesota Environment
88 and Natural Resources Trust Fund. Illumina sequencing was performed at the Biodesign
89 Institute at Arizona State University. We thank Karl Oles (Mayo Clinic Bioinformatics
90 Core) for performing PacBio library preparation and sequencing, and Jason Smith and
91 Scott Binford for assistance with experimental design and data analysis.

92

93 **REFERENCES**

- 94 1. **Butler JE, Young ND, Lovley DR.** 2010. Evolution of electron transfer out of the
95 cell: comparative genomics of six *Geobacter* genomes. *BMC Genomics* **11**:40.
- 96 2. **Badalamenti JP, Krajmalnik-Brown R, Torres CI.** 2013. Generation of high
97 current densities by pure cultures of anode-respiring *Geoalkalibacter* spp. under
98 alkaline and saline conditions in microbial electrochemical cells. *mBio* **4**:e00144–
99 13.
- 100 3. **Carmona-Martinez AA, Pierra M, Trably E, Bernet N.** 2013. High current
101 density via direct electron transfer by the halophilic anode respiring bacterium
102 *Geoalkalibacter subterraneus*. *Phys. Chem. Chem. Phys.* **15**:19699–19707.
- 103 4. **Zavarzina DG, Kolganova TV, Boulygina ES, Kostrikina NA, Tourova TP,**
104 **Zavarzin GA.** 2006. *Geoalkalibacter ferrihydriticus* gen. nov. sp. nov., the first
105 alkaliphilic representative of the family *Geobacteraceae*, isolated from a soda lake.
106 *Microbiol.* **75**:673–682.
- 107 5. **Greene AC, Patel BKC, Yacob S.** 2009. *Geoalkalibacter subterraneus* sp. nov.,
108 an anaerobic Fe(III)- and Mn(IV)-reducing bacterium from a petroleum reservoir,
109 and emended descriptions of the family *Desulfuromonadaceae* and the genus
110 *Geoalkalibacter*. *Int. J. Syst. Evol. Microbiol.* **59**:781–785.
- 111 6. **Tritt A, Eisen JA, Facciotti MT, Darling AE.** 2012. An integrated pipeline for
112 de novo assembly of microbial genomes. *PLoS ONE* **7**:e42304.
- 113 7. **Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum**
114 **A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J.** 2013.
115 Nonhybrid, finished microbial genome assemblies from long-read SMRT

- 116 sequencing data. *Nat. Meth.* **10**:563–569.
- 117 8. **Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo**
118 **CA, Zeng Q, Wortman J, Young SK, Earl AM.** 2014. Pilon: an integrated tool
119 for comprehensive microbial variant detection and genome assembly
120 improvement. *PLoS ONE* **9**:e112963.
- 121 9. **Bond DR, Mester T, Nesbø CL, Izquierdo-Lopez AV, Collart FL, Lovley DR.**
122 2005. Characterization of citrate synthase from *Geobacter sulfurreducens* and
123 evidence for a family of citrate synthases similar to those of eukaryotes throughout
124 the *Geobacteraceae*. *Appl. Environ. Microbiol.* **71**:3858–3865.
- 125 10. **Parks JM, Johs A, Podar M, Bridou R, Hurt RA, Smith SD, Tomanicek SJ,**
126 **Qian Y, Brown SD, Brandt CC, Palumbo AV, Smith JC, Wall JD, Elias DA,**
127 **Liang L.** 2013. The genetic basis for bacterial mercury methylation. *Science*
128 **339**:1332–1335.
- 129