Draft genome sequence of *Therminicola potens* strain JR

Authors: Kathryne G. Byrne-Bailey[1*], Kelly C. Wrighton[1*], Ryan A. Melnyk[1], Peter Agbo[1], Terry C. Hazen[2] and John D. Coates [1]

*The first two authors contributed equally to this work and are listed alphabetically.

Author Affiliations: [1]Department of Plant and Microbiology, University of California, Berkeley, CA 94720, USA. [2]Earth Sciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA.

Corresponding author: John D. Coates (jdcoates@berkeley.edu)

Abstract

Thermincola potens strain JR is one of the first Gram-positive dissimilatory metal reducing bacteria (DMRB) for which there is a draft genome sequence. Consistent with the physiology of this organism, preliminary annotation revealed an abundance of multi-heme *c*-type cytochromes that are putatively associated with the periplasm and cell surface in a Gram-positive bacterium. Here we report the draft genome sequence of strain JR.

Therminicola potens strain JR is a Gram-positive obligate anaerobe isolated from the anode of a thermophilic microbial fuel cell (MFC), where it constituted a dominant

member of the current-producing bacterial community (11). Strain JR coupled acetate oxidation to the reduction of external electron acceptors including MFC anodes and hydrous ferric oxide (HFO) (11). This Firmicutes member is only one of a small number of *Peptococcaceae* to be genome sequenced (6, 12), and represents both the first genome sequence of an MFC isolate and a *Therminicola* species. Genomic analysis will aid in the elucidation of external electron transfer mechanisms by strain JR, thereby contributing to the knowledge of extracellular respiration by Gram-positive bacteria. By comparing and contrasting these mechanisms in Gram-positive and Gram-negative organisms we hope to identify both the conserved and disparate aspects of this seminal metabolic function.

The draft genome consisted of a single circular chromosome of approximately 3036819 bp with an average G+C content of 45.9 %. A total of 2963 protein-encoding genes were predicted and 393 (6.9 %) had no similarity to public database sequences. Sequencing performed at the JGI used a combination of 454 and Illumina sequencing platforms to a depth of 27x coverage. All JGI library construction and sequencing techniques can be found at http://www.jgi.doe.gov/. Illumina reads were assembled, into 121 contigs, using Velvet 0.7.1.18 (14) and shredded into 1Kb pseudo-reads (with 100 bp overlap). The pseudoreads were incorporated into a hybrid 454/Illumina assembly using the parallel Phrap assembler (www.phrap.com) (1, 2). Possible mis-assemblies were corrected with Dupfinisher (3). Gene modeling was performed using Prodigal (http://prodigal.ornl.gov/) and resulting protein translations were assigned products by comparisons to databases including Pfam, KEGG, and COGs using BLASTP or HMMER.

16S rRNA gene sequence analysis identified strain JR as a Firmicutes belonging to the Peptococcaceae, in the order Clostridiales and genus *Thermincola,* sharing 99% 16S sequence identity with the two previously characterized members, *Thermincola carboxdophilia* and *Thermincola ferriacetica* (9, 13). The 23S rRNA gene identified in the draft genome did not contain the specific intervening sequences carried in the genomes of related species *Pelotomaculum thermopropionicum* and *Synthrophus wolfei* (6). As observed in *P. thermopropionicum,* strain JR also lacked a PolC DNA polymerase to synthesize its leading strands and must therefore use DnaE to replicate both strands. This would explain the bias for protein coding observed on the non-leading strand.

This organism is potentially capable of CO/CO_2 fixation using the Wood-Ljungdahl pathway (reverse acetyl-coA pathway). Energy conservation could be achieved using a carbon monoxide dehydrogenase/hydrogenase complex. No evidence for key enzymes involved in the reverse TCA cycle or the 3-hydroxypropionate pathway were identified. This finding is consistent with the observed physiology of the organism, which can grow by anaerobic carboxydotrophy utilizing CO as the sole carbon and energy source or by chemolithoautotrophy with hydrogen as an electron donor and amorphous hydrous ferric iron (HFO) as an electron acceptor. Key enzymes were missing (from the draft genome) for the energy and carbohydrate pathways: Calvin cycle, Embden-Meyerhof pathway, Enter-Doudoroff pathway, glycolysis and the pentose phosphate pathways.

In comparing Pfams from strain JR to its nearest neighbors with sequenced genomes, this bacterium had a larger number of proteins with double heme (CXXCH)

motifs, the majority containing multiple heme binding domains, up to 28 in some cases (5, 7, 8). This finding may be significant for the physiology of strain JR, as *c*-type cytochromes play an essential role in the direct reduction of extracellular electron acceptors by Gram-negative bacteria such as *Geobacter* or *Shewanella* species (4, 10). Also of note, the predicted cytochrome proteins in putative gene clusters were associated with proteins containing NHL and TPR repeats, possibly indicating involvement in biofilm attachment to electrodes or direct involvement with electron transport (7). The abundance and putative association of *c*-type cytochromes in the periplasm and cell wall surface has important implications for extracellular electron transfer by a Gram-positive bacterium.

Acknowledgments. We thank Lynne Goodwin for her management of the sequencing project at the JGI. Funding for this work was provided to JDC through the DOE Laboratory Directed Research and Development (LDRD) program, provided by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. and by the Sustainable Products and Solutions Program at UC Berkeley. KW was supported by a Tien Fellowship from UC Berkeley.

References

1. **Ewing, B., and P. Green.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res **8:**186-94.
2. **Ewing, B., L. Hillier, M. C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res **8:**175-85.
3. **Han, C. S., and P. Chain.** 2006. Presented at the Proceeding of the 2006 international conference on bioinformatics & computational biology.
4. **Kim, B. C., B. L. Postier, R. J. Didonato, S. K. Chaudhuri, K. P. Nevin, and D. R. Lovley.** 2008. Insights into genes involved in electricity generation in *Geobacter sulfurreducens* via whole genome microarray analysis of the OmcF-deficient mutant. Bioelectrochemistry **73:**70-5.
5. **Kolker, E., A. F. Picone, M. Y. Galperin, M. F. Romine, R. Higdon, K. S. Makarova, N. Kolker, G. A. Anderson, X. Qiu, K. J. Auberry, G. Babnigg, A. S. Beliaev, P. Edlefsen, D. A. Elias, Y. A. Gorby, T. Holzman, J. A. Klappenbach, K. T. Konstantinidis, M. L. Land, M. S. Lipton, L. A. McCue, M. Monroe, L. Pasa-Tolic, G. Pinchuk, S. Purvine, M. H. Serres, S. Tsapin, B. A. Zakrajsek, W. Zhu, J. Zhou, F. W. Larimer, C. E. Lawrence, M. Riley, F. R. Collart, J. R. Yates, 3rd, R. D. Smith, C. S. Giometti, K. H. Nealson, J. K. Fredrickson, and J. M. Tiedje.** 2005. Global profiling of *Shewanella oneidensis* MR-1: expression of hypothetical genes and improved functional annotations. Proc Natl Acad Sci U S A **102:**2099-104.
6. **Kosaka, T., S. Kato, T. Shimoyama, S. Ishii, T. Abe, and K. Watanabe.** 2008. The genome of *Pelotomaculum thermopropionicum* reveals niche-associated evolution in anaerobic microbiota. Genome Res **18:**442-8.
7. **Rollefson, J. B., C. E. Levar, and D. R. Bond.** 2009. Identification of genes involved in biofilm formation and respiration via mini-Himar transposon mutagenesis of *Geobacter sulfurreducens*. J Bacteriol **191:**4207-17.
8. **Shi, L., T. C. Squier, J. M. Zachara, and J. K. Fredrickson.** 2007. Respiration of metal (hydr)oxides by *Shewanella* and *Geobacter*: a key role for multihaem c-type cytochromes. Mol Microbiol **65:**12-20.
9. **Sokolova, T. G., N. A. Kostrikina, N. A. Chernyh, T. V. Kolganova, T. P. Tourova, and E. A. Bonch-Osmolovskaya.** 2005. *Thermincola carboxydiphila* gen. nov., sp. nov., a novel anaerobic, carboxydotrophic, hydrogenogenic bacterium from a hot spring of the Lake Baikal area. Int J Syst Evol Microbiol **55:**2069-73.
10. **Weber, K. A., L. A. Achenbach, and J. D. Coates.** 2006. Microorganisms pumping iron: anaerobic microbial iron oxidation and reduction. Nat Rev Microbiol **4:**752-64.
11. **Wrighton, K. C., P. Agbo, F. Warnecke, K. A. Weber, E. L. Brodie, T. Z. DeSantis, P. Hugenholtz, G. L. Andersen, and J. D. Coates.** 2008. A novel ecological role of the Firmicutes identified in thermophilic microbial fuel cells. ISME J **2:**1146-56.
12. **Wu, M., Q. Ren, A. S. Durkin, S. C. Daugherty, L. M. Brinkac, R. J. Dodson, R. Madupu, S. A. Sullivan, J. F. Kolonay, D. H. Haft, W. C. Nelson, L. J.**

Tallon, K. M. Jones, L. E. Ulrich, J. M. Gonzalez, I. B. Zhulin, F. T. Robb, and J. A. Eisen. 2005. Life in hot carbon monoxide: the complete genome sequence of *Carboxydothermus hydrogenoformans* Z-2901. PLoS Genet **1:**e65.

13. **Zavarzina, D. G., T. G. Sokolova, T. P. Tourova, N. A. Chernyh, N. A. Kostrikina, and E. A. Bonch-Osmolovskaya.** 2007. *Thermincola ferriacetica* sp. nov., a new anaerobic, thermophilic, facultatively chemolithoautotrophic bacterium capable of dissimilatory Fe(III) reduction. Extremophiles **11:**1-7.
14. **Zerbino, D. R., and E. Birney.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res **18:**821-9.

DISCLAIMER