

University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Faculty Publications in the Biological Sciences

Papers in the Biological Sciences

2015

Complete Genome Sequence of an Evolved *Thermotoga maritima* Isolate

Raghuveer Singh

University of Nebraska-Lincoln

Julien Gradnigo

University of Nebraska-Lincoln

Derrick White

University of Nebraska-Lincoln, derrick.white@yahoo.com

Anna Lipzen

U.S. Department of Energy-Joint Genome Institute

Joel Martin

U.S. Department of Energy-Joint Genome Institute

See next page for additional authors

Follow this and additional works at: <http://digitalcommons.unl.edu/bioscifacpub>

 Part of the [Biology Commons](http://digitalcommons.unl.edu/bioscifacpub)

Singh, Raghuveer; Gradnigo, Julien; White, Derrick; Lipzen, Anna; Martin, Joel; Schackwitz, Wendy; Moriyama, Estuko N.; and Blum, Paul H., "Complete Genome Sequence of an Evolved *Thermotoga maritima* Isolate" (2015). *Faculty Publications in the Biological Sciences*. 511.

<http://digitalcommons.unl.edu/bioscifacpub/511>

This Article is brought to you for free and open access by the Papers in the Biological Sciences at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications in the Biological Sciences by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

Raghuveer Singh, Julien Gradnigo, Derrick White, Anna Lipzen, Joel Martin, Wendy Schackwitz, Estuko N. Moriyama, and Paul H. Blum

Complete Genome Sequence of an Evolved *Thermotoga maritima* Isolate

Raghuvver Singh,^a Julien Gradnigo,^a Derrick White,^a Anna Lipzen,^b Joel Martin,^b Wendy Schackwitz,^b Etsuko Moriyama,^a Paul Blum^a

School of Biological Sciences, University of Nebraska–Lincoln, Lincoln, Nebraska, USA^a; U.S. Department of Energy–Joint Genome Institute, Walnut Creek, California, USA^b

***Thermotoga maritima* is a hyperthermophilic bacterium with a small genome (1.86 Mbp). Genome resequencing of Tma200, a derivative produced by experimental microbial evolution, revealed the occurrence of deletions and substitution mutations. Their identification contributes to a better understanding of genome instability in this organism.**

Received 24 April 2015 Accepted 29 April 2015 Published 28 May 2015

Citation Singh R, Gradnigo J, White D, Lipzen A, Martin J, Schackwitz W, Moriyama E, Blum P. 2015. Complete genome sequence of an evolved *Thermotoga maritima* isolate. *Genome Announc* 3(3):e00557-15. doi:10.1128/genomeA.00557-15.

Copyright © 2015 Singh et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Paul Blum, pblum1@unl.edu.

Thermotoga maritima MSB8 genomovar DSM3109 is a hyperthermophilic anaerobic bacterium (1) that grows at 80°C, producing a maximum of 4 mol of H₂ per mol of glucose (2). There are a variety of duplicated genes and direct repeats in its genome, suggesting the potential for genome instability. Genome resequencing of *T. maritima* MSB8 genomovar DSM3109 in 2011 and 2013 (3, 4) indicated that the earlier sequenced *T. maritima* MSB8 reported by Nelson et al. (NC_000853.1) (5) was an evolved laboratory variant with an approximately 8-kb deletion located between TM1847 and TM1848 (annotation according to reference 5). The 8-kb deletion may have resulted from genome instability. To assess the potential for additional instability, a cell line harboring a chromosomally integrated kanamycin-resistant suicide plasmid was allowed to segregate without drug addition but with selection for maltose catabolism as part of ongoing studies involving experimental microbial evolution (unpublished data). The initially sequenced genome of *T. maritima* by Nelson et al. (5) (NC_000853.1) was used to describe the genome changes in the resulting strains. Of 50 clonal isolates screened, 10 underwent deletion formation, including a 10-kb loss between TM1322 and TM1332. One of these 10-kb deletion isolates was named Tma200. The deleted region in Tma200 encoded five hypothetical proteins, two AstB/ChuR-related proteins, one LacI family transcriptional regulator, and three ABC transporter ATP-binding proteins. In addition, two distinct repeat sequences of 921 bp and 313 bp, respectively, were identified in TM1322 (coordinates 1340942 to 1341862 and 1342246 to 1342558) and TM1332 (1350970 to 1351890 and 1352274 to 1352586). Crossover between the 921-bp direct repeats deleted the intervening region (1341863 to 1351890). Genome resequencing of this *T. maritima* derivative expands the understanding of factors underlying genome instability in this lineage.

Genomic DNA was isolated from Tma200 as described previously (6). A DNA library was prepared from ~500-bp fragments of randomly sheared genomic DNA. This library was sequenced using an Illumina HiSeq 2000 sequencer and generated 100-bp paired-end reads. FASTQ files containing the short reads were mapped to the most recent reference genome of *T. maritima* reported by Latif et al. (4) (NC_021214.1) using Bowtie 2 version

2.1.0 and IGV version 2.3 to locate mutations and deletions that were then verified by DNA sequencing of PCR amplicons. A full consensus genome (1,859,582 bp) was generated using SAMtools version 1.1 and BCFtools alignment processing utilities version 1.1 (7) with the *T. maritima* NC_021214.1 as a reference.

The genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genome/annotation_prok). This pipeline identified 1,918 genes, 1,862 coding DNA sequence (CDS), 6 pseudogenes, 3 rRNAs, 46 tRNAs, and 7 clustered regularly interspaced short palindromic repeat (CRISPR) clusters.

Nucleotide accession number. The complete genome sequence has been deposited in GenBank under the accession number CP010967.

ACKNOWLEDGMENTS

This work was supported by the Department of Energy–Joint Genome Institute (DOE–JGI) under the Community Sequencing Program (CSP 1011924) project ID 1011924 and the UNL Cell Development Facility. The work conducted by the DOE–JGI, a DOE Office of Science User Facility, was supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

REFERENCES

- Huber R, Langworthy TA, König H, Thomm M, Woese CR, Sleytr UB, Stetter KO. 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Arch Microbiol* 144:324–333. <http://dx.doi.org/10.1007/BF00409880>.
- Schröder C, Selig M, Schönheit P. 1994. Glucose fermentation to acetate, CO₂ and H₂ in the anaerobic hyperthermophilic eubacterium *Thermotoga maritima*: involvement of the Embden–Meyerhof pathway. *Arch Microbiol* 161:460–470. <http://dx.doi.org/10.1007/BF00307766>.
- Boucher N, Noll KM. 2011. Ligands of the thermophilic ABC transporters encoded in a newly sequenced genomic region of *Thermotoga maritima* MSB8 screened by differential scanning fluorimetry. *Appl Environ Microbiol* 77:6395–6399. <http://dx.doi.org/10.1128/AEM.05418-11>.
- Latif H, Lerman JA, Portnoy VA, Tarasova Y, Nagarajan H, Schrimpe-Rutledge AC, Smith RD, Adkins JN, Lee DH, Qiu Y, Zengler K. 2013. The genome organization of *Thermotoga maritima* reflects its lifestyle. *PLoS Genet* 9:e1003485. <http://dx.doi.org/10.1371/journal.pgen.1003485>.
- Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH,

- Hickey EK, Peterson JD, Nelson WC, Ketchum KA, McDonald L, Uterback TR, Malek JA, Linher KD, Garrett MM, Stewart AM, Cotton MD, Pratt MS, Phillips CA, Richardson D, Heidelberg J, Sutton GG, Fleischmann RD, Eisen JA, White O, Salzberg SL, Smith HO, Venter JC, Fraser CM. 1999. Evidence for lateral gene transfer between archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* 399: 323–329. <http://dx.doi.org/10.1038/20601>.
6. Maezato Y, Dana K, Blum P. 2011. Engineering thermoacidophilic archaea using linear DNA recombination. *Methods Mol Biol* 765:435–445. http://dx.doi.org/10.1007/978-1-61779-197-0_26.
7. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <http://dx.doi.org/10.1093/bioinformatics/btp352>.