



# Draft Genome Sequence of a Novel *Thermofilum* sp. Strain from a New Zealand Hot Spring Enrichment Culture

Anna-Louise Reysenbach,<sup>a</sup> John A. Donaho,<sup>a</sup> Todd M. Hinsch,<sup>a</sup> John F. Kelley,<sup>a</sup> Kathleen Kouba,<sup>a</sup> Mircea Podar,<sup>b</sup> Matthew B. Stott<sup>c,d</sup>

<sup>a</sup>Department of Biology, Portland State University, Portland, Oregon, USA

<sup>b</sup>Oak Ridge National Labs, Oak Ridge, Tennessee, USA

<sup>c</sup>Extremophile Research Group, GNS Science, Wairakei, New Zealand

<sup>d</sup>School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

**ABSTRACT** A draft genome of a new *Thermofilum* sp. strain was obtained from an enrichment culture metagenome. Like its relatives, *Thermofilum* sp. strain NZ13 is adapted to organic-rich thermal environments and has to depend on other organisms and the environment for some key amino acids, purines, and cofactors.

Members of the crenarchaeal family *Thermofilaceae* (1) are widespread in terrestrial and marine hot springs (2, 3), with only two valid described species, *Thermofilum pendens* from Iceland (2) and *T. uzonense* from Kamchatka, Russia (3). Both have sequenced genomes (3, 4). Additional genomes are available for *Thermofilum* spp. from Kamchatka, "*T. adornatus*" 1910b (5) and "*T. carboxyditrophus*" 1505 (GenBank assembly accession number GCA\_000813245). With the exception of "*T. librum*" (6), all *Thermofilum* spp. require *Thermoproteales* cell extracts or spent culture broth for growth. The extreme reduction in biosynthetic pathways for purines, amino acids, and cofactors in all *Thermofilum* genomes analyzed thus far suggests that these organotrophic thermophiles rely on other organisms and the environment for these small molecules. Here, we report the nearly complete draft genome of a new *Thermofilum* sp. strain, NZ13, obtained from a hot spring at Hell's Gate (Tikitere) (38°03'47"S, 176°21'39"E; pH 6.0 and 74°C) in New Zealand.

Enrichments were maintained stably for 4 days at 80°C in a modified DSMZ medium (number 88) containing yeast extract (0.5 g/liter), tryptone (0.5 g/liter), and a headspace of hydrogen (100% vol/vol). DNA was extracted from the enrichment using the Qiagen DNAeasy blood and tissue DNA kit. Nextera DNA libraries were sequenced using an Illumina MiSeq instrument. Adaptors and low-quality reads were removed using Trimmomatic (7) and assembled using IDBA-UD (8). Contigs of less than 1 kb were removed. The contigs were binned using MaxBin (9), resulting in the nearly complete genome of *Thermofilum* sp. strain NZ13. Three additional bins for a *Fervidicoccus* sp., a *Desulfurococcaceae* sp., and a *Pyrobaculum* sp. were also generated. Optimization of the *Thermofilum* sp. NZ13 assembly was achieved by mapping the reads back using Bowtie 2 (10) and SAMtools (11, 12), reassembled using IDBA-UD, and further validated with Emergent Self-Organizing Maps (ESOM) (13). The draft genome (33 contigs) is 1.88 Mb and has 56.7% G+C content, which is higher than that of some *Thermofilum* genomes but similar to that of *T. pendens*. The genome completeness is about 99.3%, with very low (0.7%) contamination (CheckM [14]). The genome was annotated using Prodigal (15) and the arCOG database (16, 17). A total of 1,930 coding sequences, 47 tRNAs, and 5 clustered regularly interspaced short palindromic repeat (CRISPR) arrays were identified. Based on a RAxML concatenated phylogenetic tree of 16 ribosomal proteins and a 16S rRNA gene comparison using EzBioCloud (18) (98.4% 16S rRNA gene similarity to

Received 3 January 2018 Accepted 29 January 2018 Published 22 February 2018

**Citation** Reysenbach A-L, Donaho JA, Hinsch TM, Kelley JF, Kouba K, Podar M, Stott MB. 2018. Draft genome sequence of a novel *Thermofilum* sp. strain from a New Zealand hot spring enrichment culture. Genome Announc 6:e00005-18. <https://doi.org/10.1128/genomeA.00005-18>.

**Copyright** © 2018 Reysenbach 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 Anna-Louise Reysenbach, [reysenbacha@pdx.edu](mailto:reysenbacha@pdx.edu).

*T. uzonense*), *Thermofilum* sp. NZ13 probably represents a new species in the genus *Thermofilum*.

Like its relatives, the *Thermofilum* sp. NZ13 genome encodes numerous peptidases, glucosidases, and an alpha amylase, suggesting that the organism can grow on a range of peptides, polysaccharides, and starches. Additionally, *Thermofilum* sp. NZ13 shows a large reduction in biosynthetic genes, indicating that, similar to *T. pendens* and *T. uzonense*, it likely relies on the environment for most amino acids, purines, and cofactors. The genome encodes a large number of ABC transporter genes, with a high proportion being membrane transport proteins. This draft genome from a *Thermofilum* sp. from New Zealand illustrates that globally, species in this genus share very similar ecological niches.

**Accession number(s).** The *Thermofilum* sp. strain NZ13 draft genome data have been deposited in NCBI GenBank under accession number [NDWX00000000](https://doi.org/10.1093/genome/10.1128/genomeA.00726-13). The version described in this paper is the first version, NDWX01000000.

## ACKNOWLEDGMENTS

We thank Yitai Liu for assisting in establishing the enrichment cultures and Kristen Briley and Jennifer Meneghin for assistance in the initial stages of the project. Thanks also go to Ryan Skoric for assistance with analysis and Emily St. John for help with initial reviews and editing. We thank the Tikitere Trust (Whakapoungakau 24) for its ongoing support and sampling permission at Hell's Gate (Tikitere).

This work was funded by the National Science Foundation (grant DEB 1134877 to A.-L.R. and M.P.). M.B.S. was supported by the Geothermal Resources of New Zealand research program at GNS Science.

## REFERENCES

- Burggraf S, Huber H, Stetter KO. 1997. Reclassification of the crenarchaeal orders and families in accordance with 16S rRNA sequence data. *Int J Syst Bacteriol* 47:657–660. <https://doi.org/10.1099/00207713-47-3-657>.
- Zillig W, Gierl A, Schreiber G, Wunderl S, Janekovic D, Stetter KO, Klenk HP. 1983. The archaeobacterium *Thermofilum pendens* represents, a novel genus of the thermophilic, anaerobic sulfur respiring *Thermoproteales*. *Syst Appl Microbiol* 4:79–87. [https://doi.org/10.1016/S0723-2020\(83\)80035-6](https://doi.org/10.1016/S0723-2020(83)80035-6).
- Toshchakov SV, Korzhnikov AA, Samarov NI, Mazunin IO, Mozhey OI, Shmyr IS, Derbikova KS, Taranov EA, Dominova IN, Bonch-Osmolovskaya EA, Patrushev MV, Podosokorskaya OA, Kublanov IV. 2015. Complete genome sequence of and proposal of *Thermofilum uzonense* sp. nov. a novel hyperthermophilic crenarchaeon and emended description of the genus *Thermofilum*. *Stand Genomic Sci* 10:122. <https://doi.org/10.1186/s40793-015-0105-y>.
- Anderson I, Rodriguez J, Susanti D, Porat I, Reich C, Ulrich LE, Elkins JG, Mavromatis K, Lykidis A, Kim E, Thompson LS, Nolan M, Land M, Copeland A, Lapidus A, Lucas S, Detter C, Zhulin IB, Olsen GJ, Whitman W, Mukhopadhyay B, Bristow J, Kyrpides N. 2008. Genome sequence of *Thermofilum pendens* reveals an exceptional loss of biosynthetic pathways without genome reduction. *J Bacteriol* 190:2957–2965. <https://doi.org/10.1128/JB.01949-07>.
- Dominova IN, Kublanov IV, Podosokorskaya OA, Derbikova KS, Patrushev MV, Toshchakov SV. 2013. Complete genomic sequence of “*Thermofilum adornatus*” strain 1910bT, a hyperthermophilic anaerobic organotrophic crenarchaeon. *Genome Announc* 1:e00726-13. <https://doi.org/10.1128/genomeA.00726-13>.
- Stetter K. 1986. Diversity of extremely thermophilic archaeobacteria, p. 39–74. *In* Brock TD (ed), *Thermophiles: general, molecular, and applied microbiology* John Wiley & Sons, New York, NY.
- 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>.
- Peng Y, Leung HCM, Yiu SM, Chin FYL. 2012. IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28:1420–1428. <https://doi.org/10.1093/bioinformatics/bts174>.
- Wu YW, Tang YH, Tringe SG, Simmons BA, Singer SW. 2014. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* 2:26. <https://doi.org/10.1186/2049-2618-2-26>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
- 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. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27:2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>.
- Dick GJ, Andersson AF, Baker BJ, Simmons SL, Thomas BC, Yelton AP, Banfield JF. 2009. Community-wide analysis of microbial genome sequence signatures. *Genome Biol* 10:R85. <https://doi.org/10.1186/gb-2009-10-8-r85>.
- 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 GL, 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>.
- Makarova KS, Sorokin AV, Novichkov PS, Wolf YI, Koonin EV. 2007. Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. *Biol Direct* 2:33. <https://doi.org/10.1186/1745-6150-2-33>.
- Makarova KS, Wolf YI, Koonin EV. 2015. Archaeal Clusters of Orthologous Genes (arCOGs): an update and application for analysis of shared features between Thermococcales, Methanococcales, and Methanobacteriales. *Life (Basel)* 5:818–840. <https://doi.org/10.3390/life5010818>.
- Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies. *Int J Syst Evol Microbiol* 67:1613–1617. <https://doi.org/10.1099/ijsem.0.001755>.